SCIENCE AND PROFIT: ESSAYS ON THE BIOTECHNOLOGY INDUSTRY

A Dissertation Submitted In Partial Fulfillment Of The

Requirements For The Degree Of

DOCTOR OF PHILOSOPHY

IN MANAGEMENT

by Kartik Kumaramangalam

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE,

UNIVERSITY OF LONDON

June 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Following the opening three chapters, which survey the literature related to knowledge, economic growth and firm behavior, I focus on the biotechnology industry to understand how firms translate basic scientific ideas into profitable ventures. I find that this industry is characterized by two unique stylized facts: first, firms publish the results of their scientific research openly in peer reviewed journals, and two, they collaborate with universities quite intensively. I explore the private-public nature of biotechnology innovation in three separate papers.

In my first paper[1], I find that collaborative research with academics improves research quality for biotechnology firms. My results indicated that biotechnology firms seek alliances with high status academics with established research reputations so as to gain publication in highly ranked journals which is one measure of research quality. One of the major policy implications of my paper was that support for public science should be strengthened and collaboration across the private-public divide should also be encouraged. However, collaboration should not be conflated with co-option or appropriation of roles, i.e., the public sector should not be encouraged to emulate private sector functions. A strong independent public science nexus is crucial for private biotech firms, otherwise valuable signals of research quality may be compromised (as we shall see below).

The results of my first paper elicited another important question – namely, why should firms publish the results of their research openly in the first place? I address this question in my second paper, by developing an open-science framework of innovation which argues that while R&D expenditures reveal the commitment of a firm's resources to innovation and patents record the completion of R&D activity, a firm's stock of scientific papers signals the quality of its innovative efforts. In biotechnology, quality of research is a valuable signal and publishing peer-reviewed articles allows firms to convince investors and potential collaborators of the worth of their ideas. This proposition is tested using unique data of U.K. biotechnology firms during the years 1988-2000. The findings indicate that research publications bring real financial gains to biotechnology firms and that, on average, publishing fourteen scientific papers in academic journals has approximately the same impact on a firm's market value as obtaining a single patent. Furthermore, papers which are highly cited, particularly by pharmaceutical firms, have a greater impact on market value.

In a third theoretical paper, I show that biotechnology patents can be treated as credence goods – goods/services that require expert opinion to determine quality -- insofar as the market for biotechnology patents is characterized by an asymmetry of information between buyers and sellers. This informational asymmetry is a result of uncertainty associated with biotechnology patents. The chief causes of uncertainty are the legal substance of the patent document itself, the technological and commercial uncertainty associated with patent value and variable quality in screening new innovations at the

[1] This paper was awarded the Best Student Paper Award in the Technology and Innovation Management Division (TIM) and was the only student finalist for the Carolyn Dexter Award at the Academy of Management (AOM) 2005 Meetings.

patent office when granting patents. Despite these limitations, firms continue to patent in increasing numbers. Thus, the market has evolved mechanisms to more accurately ascertain "true" patent value. These mechanisms, that I label, credence verifiers, include publishing scientific papers in peer reviewed journals and the practice of clubbing patents in patent portfolios. Studying how the market ascertains the value of patents has implications for the theory and reality of patenting behavior; and by conceptualizing biotechnology patents as credence goods, this paper makes an interdisciplinary contribution (combining law and economics) to understanding the incentives that drive innovation.

Acknowledgements

I have come to understand that the PhD process is very much that of an apprenticeship rather than a clinical course of studies, and in that sense I am deeply indebted to both my supervisors, Prof. Rafael Gomez and Prof. Kretschmer for their guidance, intellectual support and friendship. Looking back on my last few years, I find that I have enjoyed my doctoral studies very much, and a large part of that joy is derived from working with and learning from my supervisors. Thank you.

I would also like to thank Prof. Grant Lewison at City University, particularly for his help in obtaining and using the bibliometric data used in my thesis.

The Interdisciplinary Institute of Management at the LSE has been my "home away from home" for the last few years, and it's been an excellent one! I would like to particularly thank Prof. DeMeza, Prof. Reyniers, Prof. Barzelay and Prof. Rothe for their support. I would also like to thank Daniel Peppiatt, Sharon Halkyard and Pete Mellows.

I cannot separate my PhD experience from the time I have spent with my friends in London. Particularly, I would like to thank, in no particular order (stand up!): Sameer Singh, Eunhee Lee, Daniel "paratrooper" Lau, Andrew "all-in" Kim, Karun Cariappa, Sam Srinivas, Aashish Velkar, Pieter Lambrecht, Badri Natarajan, Emiko Onoki, Avinash Rao, "Charlie" Xie Gang, Keith Postler, Siri Terjesen, Dominic Chai, Jonathan Rowson, Shiva Thambisetty, Pia Singh, Nitya Mohan, and Pranav Vora.

Finally, I would like to thank Prof. Phanish Puranam and Prof. Dietmar Harhoff for their constructive suggestions. My thesis is much improved by them.

Any flaws in this work are my own.

This thesis is dedicated to my parents, with love and appreciation.

Kartik Kumaramangalam, London, January 2006.

Contents

Introduction

The biotechnology industry can trace its origins to a single scientific paper.

In 1953, in a laboratory at Cambridge University, Francis Crick and James Watson discovered the double helix structure of the DNA molecule. In a classic "race to the finish", Watson and Crick submitted a single page paper to the prestigious journal Nature, starting with "We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.)." and ending with the subtle British understatement: "It has not escaped our notice that the specific pairing mechanism that we have postulated immediately suggests a possible copying mechanism for the genetic material." (Watson and Crick, 1953). This landmark achievement for which they were awarded the Nobel Prize for Chemistry in 1962, triggered a renaissance in the biological sciences. However, it would be several decades before this revolution in the "basic" sciences trickled downstream into commercial application. By this time the scientific action had moved away from the ivy clad halls of academic England to the more bohemian and entrepreneurial environs of 1970s California.

It would be nearly twenty-five later, in 1976, that two scientists working on the cutting edge of the genetic sciences at the University of California San Francisco are convinced by an adventurous venture capitalist to commercialize their ideas. Genentech, the firm they found, successfully engineers a synthetic version of the hormone Insulin. On the fourteenth of October 1980, Genentech goes public offering one million shares of stock for 35 US (1980) dollars a share. By the afternoon, the offering is sold out and by the

end of the day Genentech's stock makes market history by hitting a high of 89 US (1980) dollars, a record for an initial public offering (IPO) at the time (Wall Street Journal, 2005).

Spurred on by Genentech's success, a host of start-ups follow and the biotechnology industry begins to boom. By 2002, the global biotechnology industry comprises of approximately 4,300 publicly listed companies posting revenues of nearly $42 billion (2002) US dollars (DTI, 2005). The majority of firms are still based in the U.S. and offer a positive vision of wealth creation and better living standards based on scientific progress that attract countries across the world to emulate the U.S. model. However, success does not come easy. For every biotech wonder, a hundred others fall by the wayside. What makes some biotech firms more successful than others?

Although, a large literature has emerged devoted to the strategic analysis of the biotechnology industry, it primarily focuses on industry structure and the strategic nature of R&D to the exclusion of a more nuanced model of scientific innovation. Specifically, the bulk of the literature relies on a "linear" view of scientific innovation (Rosenberg, 1990). This view assumes that basic scientific knowledge is created by universities and public research institutions and exists in the economy as a public good. Firms then freely "borrow" from this common pool of knowledge and develop technological products by making purposeful investments in research and development (R&D). Often times, there are several firms simultaneously racing towards the same goal. The "winners" of this R&D race are then awarded restricted monopoly rights to their innovations, thus providing incentives for further R&D investment.

However, this "waterfall" view of scientific innovation has been shown to be inadequate (Rosenberg, 1990). The line between basic "upstream" scientific research and "downstream" technological development is often blurry. Firms, particularly science based high technology firms engage in basic scientific research as well as commercial development. Sometimes, firms specialize purely on scientific research and license their intellectual property. These high technology firms further mimic open science norms by regularly publishing the results of their research in peer-reviewed scientific journals (Hicks, 2002). In doing so, they seemingly engage in irrational behavior.

The classic model of scientific innovation presupposes that firms will choose to reveal as little information about their activities while engaged in an R&D race with their competition (Arrow, 1962; Dasgupta and David, 1994). But yet, we observe that firms regularly reveal information about their R&D program in peer-reviewed journals. Why should they do so? A number of theorists have suggested that high technology firms adopt open science norms in order to develop routines and skills that allow them to utilize effectively the advances in publicly funded research (Cohen and Levinthal, 1990; Cockburn and Henderson, 1998). There is also evidence to suggest that adopting open science norms confers labor cost advantages as scientists are often willing to accept a lower wage in exchange for permission to continue publishing scientific papers and thus, maintaining their links with open research (Stern, 1999). Also, work by Zucker and Darby (1998) shows that "star" scientists play a key role in the success of biotechnology firms. These scientists often retain their links to academic institutions and thus participate in the dual worlds of public and private research. Finally, it may be that

publishing scientific papers acts as advertising signal and conveys information to the market and a firm's competitors of its research trajectory and quality.

Thus, while this dual private-public behavior has been well documented, particularly for firms in the biotechnology industry (Murray, 2002; Henderson and Cockburn, 2003), the process by which these firms navigate the separate domains of public and private research remains unexplored. Further, there is a specific lack of empirical literature that investigates the private-public nature of knowledge assets. This thesis attempts to fill this gap by exploring the interaction between public and private science in the U.K. biotechnology sector.

This thesis is structured as follows. The first three chapters provide the background literature review that motivates this thesis. In **Chapter 1**, I review the literature linking knowledge and economic growth. However, the proximate cause of growth in the economy is directly linked to the action of firms and in **Chapter 2** I review the various theories of the firm giving special attention to the role of knowledge. In **Chapter 3**, I discuss various issues related to measurement of knowledge within the firm.

Chapters 4, 5, 6 are my substantive analytical chapters which focus on the incentives that drive knowledge creation in the biotechnology industry. Specifically, these chapters address the following questions:

Chapter 4. What motivates private-public collaboration? A substantial and growing body of work points to the increasing value of private-public interaction for the

performance and growth of high technology science-based firms and industries. However, research on the effect of this interaction on the resulting quality of a firm's scientific output is scarce. In Chapter 4, I address this gap by comparing the effect of private-public collaboration on research quality using unique data from the UK biotechnology sector from 1988-2001. My findings indicate that collaborative research with academics improves research quality, although the status of the academic researchers and the nature of the biotechnology firm in question are significant factors in determining the strength of positive effect private-public collaboration has on research quality. These results suggest that biotechnology firms should seek alliances with high status academics with established research reputations. A major policy implication is the need to strengthen support for public science and encourage collaboration across the private-public divide.

Chapter 5. Why Do Firms Make Private Knowledge Public? Firms regularly publish the results of their scientific research in peer-reviewed journals. However, given classical models of scientific innovation that assume firms reveal as little information about their R&D activities as possible, the motivation for making private knowledge public is unclear. In Chapter 5, I develop an open-science model of innovation which argues that while R&D expenditures reveal the commitment of a firm's resources to innovation and patents record the completion of R&D activity, firms' publication of scientific papers signals the quality of its innovative efforts. In biotechnology, quality of research is a valuable signal and publishing peer-reviewed articles allows firms to convince investors and potential collaborators of the worth of their ideas. This proposition is tested using unique data of U.K. biotechnology firms during the years

1988-2000. My findings indicate that research publications bring real financial gains to biotechnology firms and that, on average, publishing fourteen scientific papers in academic journals has approximately the same impact on a firm's market value as obtaining a single patent. Furthermore, papers which are highly cited, particularly by pharmaceutical firms, have a greater impact on market value.

Chapter 6. Why are scientific papers a credible signal of research quality of biotechnology firms? In Chapter 6, I show that biotechnology patents can be treated as credence goods insofar as the market for biotechnology patents is characterized by an asymmetry of information between buyers and sellers. This informational asymmetry is a result of uncertainty associated with biotechnology patents. The chief causes of uncertainty are the legal substance of the patent document itself, the technological and commercial uncertainty associated with patent value and variable quality in screening new innovations at the patent office when granting patents. Despite these limitations, firms continue to patent in increasing numbers. Thus, the market has evolved mechanisms to more accurately ascertain "true" patent value. These mechanisms, that I label, credence verifiers, include publishing scientific papers in peer reviewed journals and the practice of clubbing patents in patent portfolios. Studying how the market ascertains the value of patents has implications for the theory and reality of patenting behavior; and by conceptualizing biotechnology patents as credence goods, this paper makes an interdisciplinary contribution (combining law and economics) to understanding the incentives that drive innovation.

In my final chapter, I conclude the thesis with a summary of my main results with their implications for theory and practice and also generate some ideas for future work motivated by this thesis.

CHAPTER 1

Knowledge and economic growth

1.1. Introduction.

The purpose of this chapter is to review theories of economic growth, with a particular view to emphasize the role of knowledge in generation sustainable economic growth and development. The central argument that I will develop in this chapter is that *ideas* are the central proximate cause of growth and productivity. In the remainder of my thesis, I explore the notion of how we can more precisely quantify and define ideas, particularly those that are produced by scientific effort. While in this chapter, I show that new ideas and knowledge are central in explaining increasing returns to scale and continued growth per capita, there is a gap in our understanding of how to value knowledge and how it can be applied by firms. This is a gap that is addressed by my thesis.

1.2. Economic models and their ideas

Classical writers such as Mill and Marx speculated that standards of living would not rise indefinitely unless advances in technology served to augment the productivity of resources (Arrow, 1962). This proposition received analytical support from the neo-classical growth theorists, who elaborated a model of growth based on capital accumulation. If production of output is characterized by diminishing returns to the accumulated factors, the incentive to invest may disappear in long run in the absence of

productivity gains. The fact that investment has continued for more than two hundred years since the industrial revolution suggests that technical change has played a major role in the growth process.

The systematic relationship between output and productivity growth rates and a number of economic variables suggest moreover that technological progress probably is not a purely random process but rather are guided by market forces. Early writers on the sources of technological change saw scientific discoveries as a primary, stimulating force behind innovation. Since scientific advances largely reflect the interests and resources of a community of researchers operating outside profit sector of the economy, a scientific basis for most industrial innovation would remove technological progress from the realm of economic analysis. But Schmookler (1966) took exception to this view of the way that technologies evolve in his influential study of almost a thousand inventions in four different studies. Schmookler argued in a great detail that it is the expected profitability of inventive activity, reflecting conditions in the relevant factor and product markets, that determines the pace and direction of industrial innovation. Schumpeter (1942) had expressed a similar more than twenty years earlier. If Schumpeter and Schmookler are correct, then it would not be surprising to find productivity growth related to an economy's structure and policies, or to find variation in growth experiences in different parts of the world. This paper outlines key economic concepts that underlie this intuition and presents empirical evidence that support the theory.

1.2.1. The Solow model: the role of technology

The Solow model is built around two equations, a production function and a capital accumulation equation.

The production function is assumed to have the Cobb-Douglas form and is given by:

$$Y = F(K,L) = K^{\alpha} * L^{1-\alpha}$$

where α is some number from 0 and 1.

Firms in this economy pay workers a wage, w, for each unit of labor and pay r in order to rent a unit of capital for one period. To generate sustained growth in per capita income in this world, we need to introduce technological progress. We modify the production function so that:

$$Y = F(K,AL) = K^{\alpha}(AL)^{1-\alpha},$$

where A is the "technology variable". Technological progress occurs when A increases over time – a unit of labor, for example, is more productive when the level of technology is higher.

An important assumption of the Solow model is . that technological progress is exogenous. Instead of modelling carefully where technology comes from, we simply recognize for the moment that there is technological progress and make the assumption that A is growing at a constant rate:

$$\frac{\dot{A}}{A} = g \Leftrightarrow A = A_0 \cdot e^{gt},$$

where g is a parameter representing the growth rate of technology.

The crucial implication of this model with technology reveals that technological progress is the source of sustained per capita growth. According to the Solow model, higher investment rates and lower population growth rates allow more capital to accumulate per worker and thus increase labor productivity. However, without technological progress, per capita growth will eventually cease as diminishing returns to capital sets in. technological progress, however, can offset the tendency for the marginal product of capital to fall, and in the long run, countries exhibit per capita growth at the rate of technological progress. A more subtle explanation can be found by examining the transition dynamics of the growth trajectories. Economies can grow at rates different from their long-run growth rates. An economy with a capital-technology ratio below its long-run level will grow rapidly until the capital-technology ratio reaches its steady state level. However, it is very crucial to note that in Solow's model technological improvements arrive exogenously at a constant rate, g, and differences in technologies across economies are unexplained. For a deeper understanding for the sources of growth we will need to examine economic models of knowledge and their impact on technological change.

1.2.2. The Romer model: the economics of knowledge.

Instead of assuming that growth occurs because of automatic and unmodeled (exogenous) improvements in technology as in the Solow model, Romer's model focuses on understanding the economic forces underlying technological progress. The

key contribution of this work is to explore the economics forces that underlie the production of ideas and its consequence for growth.

Basic model

The Romer model endogenizes technological progress by introducing the search for new ideas by researchers interested in profiting from their inventions. By modelling technological change as driven by research and development (R&D), the model seeks to understand how technological frontier is continually pushed forward.

The aggregate production function in the Romer model describes how the capital stock, K, and labor, L_y, combine to produce output, Y, using the stock of ideas, A:

$$Y = K^{\alpha}\left(AL_y\right)^{1-\alpha},$$

where α is a parameter between 0 and 1.

For a given level of technology, A, the production function exhibits constant returns to scale in K and L_y. However, when we recognize that ideas (A) are also an input into production, then there are increasing returns.[2]

The key equation that is new relative to the neo-classical model is the equation describing technological progress. In the neo-classical model, the productivity term A grows exogenously at a constant rate. In the Romer model, growth in A is endogenized. According to Romer, $A(t)$ is the stock of knowledge or the number of ideas that have been invented over the course of history up till time t. Then, $\dot{A}$ is the number of new

[2] The presence of increasing returns to scale results fundamentally from the non-rivalrous nature of ideas.

ideas produced at any given point in time. In the simplest version of this model, $\dot{A}$ is equal to the number of people attempting to discover new ideas, L_A, multiplied by the rate at which they discover new ideas, $\bar{\gamma}$:

$$\dot{A} = \bar{\gamma} \cdot L_A$$

Labor is used either to produce new ideas or to produce output, so the economy faces the resource constraint:

$$L_A + L_Y = L$$

Further,

$$\bar{\gamma} = \gamma \cdot A^{\phi},$$

where δ, ϕ are constants.

In this equation, $\phi > 0$ indicates that the productivity of research increases with the stock of ideas that have already been discovered; $\phi < 0$ correspond to the situation in which the most obvious ideas are discovered first and subsequently ideas are increasingly difficult to discover. Finally, $\phi = 0$ indicates that the tendency for the most obvious ideas to be discovered first exactly offsets the fact that old ideas may facilitate the discovery of new ideas – i.e. the productivity of research is independent of the stock of knowledge.

Thus, in Romer's model, the long-run growth rate of this economy is determined by the parameters of the production function for ideas and the rate of growth for researchers. Intuitively, this means in order to generate growth, the number of ideas must be

expanding over time. This occurs if the number of researchers is increasing. More researchers means more *ideas*, sustaining growth in the model.

It is interesting to compare this result to offset of population growth in the neo-classical growth model (the Solow model described earlier). These, for example, a higher population growth reduces the level of income along a balanced growth path. More people means that more capital is needed to keep K/L constant, but capital runs into diminishing returns.

Here, new growth theory (Romer, 1990) and other theorists[3] emphasize that ideas are very different from other economic goods. Ideas are non-rivalrous: once an idea is invented, it can be used by one person or one thousand people, at no additional cost. In particular, the non-rivalry of ideas implies that production will be characterized by increasing returns to scale. The incentive to create new ideas depends on the profits that an inventor can be expected to earn (the private benefit), not on the entire social benefit generated by the idea. Whether or not an idea gets created depends on the magnitude of the private benefit relative to the invention costs. It is easy to see, then, how ideas that are socially very valuable may fail to be invented if private benefits and social benefits are too far apart. Patents and copyrights are legal mechanisms that attempt to bring the private benefits of invention closer in line with the social benefits. The developments of such institutions and of property rights more generally, are thus crucial to sustained economic growth.

[3] Other notable contributions to the literature on R&D-based growth models include Grossman and Helpman (1991) and Aghion and Howitt (1992). These models are sometimes called Schumpeterian growth models, because they were anticipated by the work of Schumpeter (1942) in the late 30's and early 40's.

1.3. Accounting for Prosperity

We have seen in the Solow model that sustained growth occurs only in the presence of technological progress. Without technological progress, capital accumulation runs into diminishing returns. With technological progress, however, improvements in technology continually offset the diminishing returns to capital accumulation. Labor productivity grows as a result, both directly because of the improvements in technology and indirectly because of the additional capital accumulation these improvements make possible.

In 1957, Solow published a second article, *Technical Change and the Aggregate Production Function*, in which he performed a simple accounting exercise to break down growth in output into growth in capital, growth in labor, and growth in technological change. This "growth accounting" exercise begins by postulating a production function such as

$$Y = BK^{\alpha} \cdot L^{1-\alpha}$$

where B is a Hicks-neutral productivity term.[4] Taking logs and differentiating this production function, one derives:

$$\frac{\dot{Y}}{Y} = \alpha \frac{\dot{K}}{K} + (1-\alpha)\frac{\dot{L}}{L} + \frac{\dot{B}}{B}$$

This equation states that output growth is equal to a weighted average of capital and labor growth plus the growth rate of B. this last term, $\dot{B}/B$, is commonly referred to as total factory productivity growth (TFP). Solow, as well as economists such as Edward

[4] In fact, this growth accounting can be done with a much more general production function such as $B(t)F(K,L)$ and the results are very similar (Solow, 1957).

Denison and Dale Jorgenson who followed Solow's approach, have used this equation to understand the sources of growth in input (see Solow, 1957). Using this approach on historical data on the United States, Solow (1957) revealed that accounting for growth due to capital formation and that due to growth in the labor force: still left a significant share that remains unexplained by growth in inputs to the production function. One interpretation of this "residual" growth is that it is due to technological change (see Romer, 1989). In an influential paper published in 1992, *A Contribution to the Empires of Economic Growth*, Gregory Mankiw, David Romer, and David Weil evaluated the empirical implications of the Solow model and concluded that it performed very well.[5] Further the fit of the model was greatly improved by extending it to include human capital – that is, by recognizing that labor in different economies may possess different levels of education and different skills. Investigating the relationship in Romer's (1992) model of endogenous growth attributable to levels of knowledge stocks, we run against some difficulty. As some fundamental level it is difficult to measure both the inputs to the production function, the ideas themselves.

However, data such as R&D, the number of scientists and engineers engaged in an economy and patent counts are common measures used to proxy the stock of "knowledge" in an economy. Various proxies for human capital, such as the literacy rate (Azariades and Drazen, 1990; Romer, 1989) and the school enrolment rate (Baumol et al 1989, Barro, 1989) correlate positively with real GDP growth. Romer (1989) finds a positive correlation between the number of scientists and engineers employed in

[5] Mankiw, Romer and Weil allow an economy to accumulated human capital in the same way as physical capital – by foregoing consumption (see Romer, 1990).

research and the growth rate of output in a sample of 22 of the most developed countries.[6]

There are several well known problems that arise in interesting results from growth accounting exercises. First, GDP growth may not accurately measure growth in economic output because increases in the quality and variety of goods and services available to consumers are only imperfectly related in the national income accounts. The measurement of the contribution of new and improved varieties to real output growth requires the implementation of sophisticated index number procedures. It is generally believed that reported price indexes often underestimate the economic benefits from product innovations (e.g., see Griliches, 1973; Bresnahan, 1986; Trajtenberg, 1990), in which growth accounting will understate the extent output growth attributable to technology. In separating the contribution of "knowledge stocks" to output growth, knowledge (i.e., the accumulation of R&D investment less estimated depreciation) is heated as an ordinary, accumulated input, along with tangible factors such as capital and labor. Of course one cannot directly observe the reward paid to the knowledge stock as most of the returns are hidden in data on corporate profits. So an independent estimate of the rate of return is estimated econometrically using cross-sectional data on firms or industries. However, only the "direct returns" to R&D (i.e., the returns that accrue to the firm or industry that conducts the R&D) are captured by these methods. Various empirical work (for a review see, Mansfield et al 1977; Scherer, 1982) has shown that research effort often generate sizeable spill over benefits. Further, those external

[6] Griliches (1979) reviews the approaches that have been used to measure the contribution of R&D to economic growth, and discusses the methodological and data problems associated with each one → more empirical reviews.

benefits may not be concentrated in a single industry but rather may be spread to many sectors in the economy.

1.4. Concluding Remarks

New growth theory clearly places knowledge centre stage as the primary driver of economic growth. As suggested by the Solow model: growth ceases unless the technology of production increases exponentially. The Romer model reveals that it is the production and dissemination of *ideas* that makes this possible. Particularly, it is the *non-rivalrous* property of ideas that causes increasing returns to production. The presence of increasing returns to scale means that we cannot model that economics of ideas using perfect competition.

Firms must be able to charge prices greater than marginal cost to cover the one-time expense required to create an idea. It is this wedge between price and marginal cost that provides the economic "fuel" for the engine of growth. The non-rivalrous property of ideas thus requires institutions such as intellectual property rights to provide incentives for firms to invest in the production of new knowledge. Thus, we find that management is always the proximate cause of growth and productivity. In other words, even with Adam Smith's invisible hand, a competitive system produces optimal results because a competitive market is always rewarding well-managed firms at the expense of poorly managed firms. One distinguishing feature therefore, between a good and a bad firm may be the knowledge that they possess. In the next chapter I explore this idea more fully.

Chapter 2

Knowledge and the firm

2.1. Introduction

While the previous chapter set the stage by establishing the role of knowledge and innovation in economic growth, the search for the internal determinants of innovative activity is in its infancy. The task of this chapter is to survey various knowledge-based theories of the firm. I begin by reviewing various classical theories of the firm and outlining their shortcomings. It goes on to describe more recent theories of the firm, particularly those built on knowledge-based perspectives. The chapter concludes by summarizing the main conclusions of the existing literature while suggesting possibilities for future development. The analytical goal of this chapter is to establish the importance of knowledge as a resource in firm capability and sets the stage for my empirical research in the subsequent chapter on how firms in the biotechnology industry build and sustain their competitive advantage.

2.2. Theories of the firm

The cornerstone of formal economic theory is a price mechanism. An economic theorist thinks of the economic system as a being coordinated by the price mechanism wherein price movements direct decisions of production and consumption through a series of market transactions. Traditionally neo-classical economics worked within a strict dichotomy - in the long run all factors of production (i.e. all resources) are variable and

there is no distinctive role for intuitions such as firms. In the short run, when at least one factor, usually capital, is fixed, the firm has a role but this is usually confined to making an optimal price/output decision with regard to prevailing factor and product prices. Thus the neo-classical theory of the firm, at least in its standard manifestations, turns out to be something of a misnomer for a theory of the short run behaviour of markets.

One of the earliest theorists to grapple with this fundamental dichotomy presented by the neo-classical of the firm was Coase (1937). In his classic article, "The theory of the firm", he poses the question asked earlier by Adam Smith, "How is the cooperation of the vast numbers of people in counties all over the world, which is necessary for even a modest standard of living, to be brought about?". Coase argues that there are two possibilities for the economic organization of productivity – contracts with independent contractors over the market and contracts with employees organized within a business firm; and whichever is cheaper will be adopted.

In Coase's words, "there would be an optimum of planning since a firm, that little planned society, could only continue to exist if it performed its coordination function at a lower cost than this same function could be performed by another firm". Thus it may be profitable to establish a firm if the cost of sing the price mechanism exceeds that of organizing the allocation of factor inputs internally within a firm. This will be so if the workings of economic markets are not costless. The most obvious cost of "organising" production through the price mechanism is that of discovering what the relevant prices

are[7]. Further the costs of negotiating and concluding a separate contract for each exchange transaction which takes place on a market must also be taken into account.

While it is true, that contracts are not eliminated when there is a firm, they are greatly reduced. Also, a firm allows for long-term contracts for the supply of some factor input, thus reducing the need for repeated short term market based transactions. There is considerable evidence to support this argument. For example, the classical economic analysis of a vertical integration hinges on the idea that vertical integration is a monopolizing device that is employed to create barriers to entry that extend or protect monopoly power. Coase (1937) claimed that vertical integration can be understood more clearly if one took into account transaction costs rather than focusing on classic monopoly and cartel economic analyses. He argues that businessmen bring a part of the process of production within the firm, rather than arranging for it by contract with other producers, when the cost of co-ordinating the firm's inputs by market transactions exceeds the cost of co-ordinating them hierarchically.

Following on the fundamental work of Coase, theorists (Williamson, 1975, Demsetz, 1982) conceptualized a theory of transaction costs on the basis of employment contracts within the firm being more efficient then contracting on the external market. Thus, this stream of research is concerned with the costs of search, metering and monitoring associated with the transfer of resources and products across markets and within organizations. Particularly, asset specificity with its implications of small numbers bargaining, co-dependence and vulnerability to hold-up has emerged as the most

[7] This cost may be reduced but not eliminated by the emergence of specialists who sell this information.

important impediment to contractual relations among opportunistic agents. Further, there is now a vibrant empirical tradition that reports strong support for the role of transaction cost drivers in determining a wide variety of institutional outcomes (Shelanski and Klein 1995).

This stream of research that seeks to explain the existence of firms in light of their apparent irrelevance in the neo-classical system has come to be classified more broadly as organizational economics. This school of thought has developed as a collection of partially overlapping schools including evolutionary theory (Nelson and Winter, 1982), property rights theory (Demsetz, 1974), transaction cost economics (Williamson, 1975) and the positive theory of agency (Fama and Jensen, 1983).[8]

However, the two branches of organizational economics most directly concerned with firm organization are transaction cost economics and positive theory of agency. Both approaches share the assumption of opportunism and each defines an efficient set of institutional arrangements as one that minimizes the sum of organizational and production costs. Particularly, the positive theory of agency identifies costs of monitoring and bonding as necessary in making viable economic relations between a self-interested, opportunistic agent and principal. It suggests that organizational evolution will favour the emergence of arrangements that minimize the sum of these

[8] Detailed comparisons between several or all of them are available elsewhere. For an extensive review see Mahoney and Pandian, 1992. Note: Organizational economics is interchangeably referred to as "new-institutional economics" (Williamson, 1975).

incurred costs plus deadweight costs that arise from potentially useful collaborative relations rendered non-viable by the agency problem (Fama and Jensen, 1983).[9]

Although both transaction cost minimization and agency cost minimization have strong methodological similarities, each has tended to be applied to somewhat different problems. Transaction cost economics has been used extensively to explain the location of the appropriate boundaries for the firm. Thus decisions on vertical integration are the "make/buy" choice, the use of franchising or company outlets the use of joint- versus wholly-owned ventures have been analyzed extensively in transaction terms. Since these are key strategic decisions of managers, much of this work has appeared within the strategy literature (see Mahoney and Pandian, 1992). By contrast, much work with the principal theory of agency framework concentrates upon the important agency relationship between stockholders and managers. It typically centres on the success of market and non-market governance arrangements in securing an alignment of interest between the two groups.

While those theories dealt with the fundamental paradox of the neo-classical theory of the firm (i.e. why should it exist) by taking a static equilibrium approach, they fail to address a central concern of strategy research – how do firms build and sustain competitive advantage? To do so, newer theories have placed the accumulation and deployment of knowledge as a key strategic asset. For example, recent work by Kretschmer and Puranam (2005) examines how firms secure competitive advantage via managing their firm boundaries. In tandem, newer theories have also arisen that place

[9] It is distinguished from the more formal research of principal-agent theory that typically looks to derive an optimal contract, one that maximizes the principal's welfare whilst securing the agent's continuing participation.

the accumulation and deployment of knowledge as a key strategic asset. I discuss these theories and their implications in the following section.

2.3. Knowledge-based prescriptions in the search for rent

2.3.1. The search for rent

According to Penrose (1959) and Rumelt (1984, 1987) the primary purpose of a firm is to create, exploit and defend sources of economic rent, where rent is defined as return in excess of a resource owner's opportunity costs (Tollison, 1982). The field of strategy research is concerned with this question. Broadly speaking, there are two sets of ideas that contribute in this debate. The first set of ideas come under the rubric of industrial organisation theory and can trace their roots in the early work of the economics of imperfect competition by Chamberlain (1933) and was ultimately popularized by Michael Porter in his now classic book, *Competitive Strategy (1990).*

The field of industrial organization can arguably be divided into two camps; the Harvard school after Bain (1954, 1968) and the Chicago School (Stigler, 1968). While both those firms emphasize that the rent behavior of firms is largely determined by the "structure" of competition, they differ on their views of the sustainability of these rents.

The sustainability of rents in the IO framework is supported by the concept of entry barriers at the industry level and mobility barriers at the strategic group level (Caves and Porter, 1977; McGee and Thomas, 1986). These entry and mobility barriers are a private

collective asset of an industry's (strategic group's) incumbents, and investments to augment these assets are subject to free-riding and under provision. These rent generating barriers can take the form of investments that entail high exit barriers and high switching (Porter, 1986); high sunk cost investments that advantage incumbents (Baumol, Panzar and Willig, 1982); legal restrictions to entry (Stigler, 1968); economies of scale that might combine with imperfect markets (Bain, 1968) or learning and experience curve advantages that are kept proprietary (Lieberman, 1987; Spence, 1981).

The Chicago school views questions whether economies of scale, advertising and R&D expenditure can ever be a sustainable barrier to entry (Demsetz, 1974, 1982; Stigler, 1968) whereas the Porter-Bain framework of the Harvard school posits sustainable monopoly rents. However, neither school grapples directly with firm heterogeneity (such as differential worker ability and managerial strategies) as a source sustainable competitive advantage focusing instead on industry level effects on firm performance.

A contrasting set of ideas come under the heading of "the resource based view of the firm" (RBV) suggest rents are owed to the internal structure of assets within the firm[10]. An asset can be physical or it can be intangible (and indeed it can be even impossible to isolate and understand). Further development of RBV has led to the notion that assets that can be imitated or duplicated cannot yield rent owing to the forces of competition and imitation. To simplify the case: with imitability rents disappear; without it, rents continue. Further, RBV theorists suggest that knowledge may be the key source of rent,

[10] In addition rents can accrue from start-up costs and sunk cost effects which prevent existing firms from adopting new and better technologies. Therefore rents are not only related to RBV theories but also the theories of procrastination and slow diffusion of ideas and technology (see Penrose, 1959).

and the process by which knowledge is created and utilized in organizations may be the key inimitable resource that could create sustainable rents (Spender and Grant, 1996).

2.3.2 The "Resource-based view" (RBV) model and Knowledge

The notion that firms are fundamentally heterogeneous, in terms of their resources and internal capabilities has long been at the heart of strategic management beginning with Penrose's classic work, *The theory of strategic growth of the firm (1959).* Those firms which are superior relative to those of rival in terms of organizational competencies and resources achieve superior returns if they are matched appropriately to environmental opportunities (Barney, 1991). Those ideas may be thought of as the basic principles upon which resource-based research continues to build. However, the major contribution of the resource-based model is that it explicitly seeks to theoretically explain persistent differences in inter-firm profitability that cannot be attributed to differences in industry conditions. Indeed several econometric and empirical studies have indicated that firm-effects are substantial (Mueller, 1986; Hansen and Wernerfelt, 1989; Rumelt, 1991). The resource-based model is a theoretical complement to this work.

In its search for firm specific resources that could sustain competitive advantage, the resource based view places "knowledge" centre stage. Winter (1988) holds that business firms are fundamentally organizations that know how to do things and that they perform their functions as "repositories of knowledge" (pp. 175). Prahalad and Hamel (1990), in developing the concept of core competencies emphasize that core competencies are the

collective learning in the organization, especially their ability to coordinate diverse production skills and integrate multiple streams of technologies.[11] Further, Kogut and Zander (1992) comment that the crucial contribution of the RBV is the realisation that knowledge underlies the set of capabilities that sustain firms' competitive advantage. Lippman and Rumelt (1982) formally link the existence of privately held knowledge, in the form of causal ambiguity, to a firms' ability to learn above-normal returns through productive activity. Thus it is the existence of knowledge in the hands of a small number of firms that creates the market imperfections necessary to generate rents for the firm. Put another way, it is proprietary knowledge that creates comparative knowledge for the firm (Cyert et al, 1993).

It can thus be argued that most firms rely on knowledge, be it process, product or otherwise, in order to deliver a sustainable competitive advantage in the marketplace. Oftentimes, the ability to convey this information to the marketplace or to competitors can be advantageous. In contract theory, signalling is the idea that one party (called the agent) conveys some meaningful information about itself to another party (called the principal). The seminal work in this area is Spence's job-market signalling model (Spence, 2002) where employees successfully signal the level of their skill by acquiring education rather than entering the job market earlier on (thus forsaking wages for education).

Like the job market, there are numerous other examples in which inequalities in access to information upset the normal market for the exchange of goods and services.

[11] Also see Teece et al (1990): "...It is not only the bundle of resources that matter but the mechanisms by which firms learn and accumulate new skills and capabilities, and the forces that limit the rate and direction of this process." (p. 11)

Signalling theory solves this problem by proposing that two parties could get around the problem of asymmetric information by having one party send a signal that would reveal some piece of information to the other party thus bridging the informational gap. Thus, the market evolves a signalling equilibrium in which the sender signals honestly and the receiver trusts that information.

This insight is particularly relevant in knowledge intensive industries such as science based firms (ex., biotechnology firms). Typically, these firms, in their early years, do not have strong profits or products. Instead, they have a portfolio of patents backed by R&D. In such circumstances, it is crucial for the firm to accurately convey to the market (be it potential investors, the stock market, or potential customers) the worth of their knowledge. How might firms do this? Later in this thesis, I argue that by publishing scientific papers firms signal the worth of their knowledge to the marketplace. Scientific publications are peer reviewed and hence act as an independent verifier of knowledge that the market can trust as an honest and accurate account of a firm's knowledge base.

Placing knowledge centre stage still leaves the problem of how we identify and measure such knowledge resources. Much of the earlier empirical research focused upon R&D expenditure as inputs into knowledge creation and patents as outputs. Although patents are very satisfactory indicators of knowledge creation in terms of being documented knowledge whose novelty has been verified by a meticulous, legalistic research process, not only are patents a very partial measure of production of organizational knowledge, but patenting is itself a strategic choice. Griliches (1990) noted that industries varied widely in their propensity to patent. At the firm level, success at patenting does not necessarily correspond to success in translating patents into competitive advantage.

Likewise Narin, Noma and Perry (1987) have shown that much patenting activity is defensive and strategic, not fitting into the simpler model of the patent as a straightforward transitional step between R&D and production. Strategic patenting may be less to do with building up knowledge than with sealing off an area from exploitation by other, so adding value to patents already held.

Broadening ideas about what constitutes knowledge within the firm has been associated closely with the work of Michael Polanyi (1962) on tacit knowledge and Nelson and Winter (1982) on organizational routines. This work has been especially influential in directing attention to knowledge which is embodied in individual and organizational practices that cannot be readily articulated. Such knowledge is of critical strategic importance because, unlike explicit knowledge, it is both inimitable and appropriable. However, these very variables which are most theoretically interesting are those which are at least identifiable and measurable. Polanyi (1962) reaches into this contradiction with his notion of tacit knowledge, implying that a knowledge-based theory of the firm differs from all previous theories in that it must grasp the potentially unknowable.

Despite the epistemological difficulties that pinning down definitions of knowledge entails, there is a growing body of literature that explores how firms create and exploit knowledge in their search for rents. It has been shown that despite the "weightless" (Quah, 1992) nature of ideas and their potential to diffuse widely, the production and spread of knowledge tends to be spatially concentrated (Jaffe, Trajtenberg and Henderson, 1993; Almeida and Kogut, 1999; Krugman, 1983). Thus, innovative high technology firms exploring new areas of technology tend to cluster geographically tied

to specific institutional variables. These clusters arise out of "spillovers" that result in the production of new knowledge given that for new technology based firms often working on the cutting edge of their field; proximity among partners such as other new technology based firms, universities and research institutes allows for face-to-face interaction and unplanned communication which is often a crucial factor for success[12] (Jensen and Thusby, 2001). Further, a great deal of knowledge is created and transmitted in the context of communities of individuals linked by common identification to their work. The role of these "epistemic" communities in generating knowledge has been studied in the context of physicists and molecular biologists (Knorr-Cetina, 1999)[13]; researchers engaged in the design of a new technological artefact (Garud and Rappa, 1994); and aerospace engineers (Knorr-Cetina, 1999)[14].

These scholars suggest that the "stickiness" of knowledge is not so much a function of its underlying degree of tacitness, but reflects its embedded-ness in epistemic communities who share common cognitive frames, norms of communication, and rules governing the creation, validation, and selection of new ideas and artefacts. Thus the strategic dimension of knowledge flows is conditioned by the social organization of the individuals who collectively generate the knowledge. Firms accrue dependable advantages when social identification within the firm is strong and firm-specific organizing principles guide the development and application of new knowledge (Kogut and Zander, 1992). But the identification of employees with wider work-based

[12] Reflecting the "tacit" nature of knowledge (Polanyi, 1962; Dosi, Teece and Winter, 1992).
[13] Epistemic cultures is Knorr-Cetina's term to describe communities identified with the process of creating knowledge itself, e.g., research scientists.
[14] The idea of epistemic communities is echoed in technological communities to describe populations engaged in the construction of complex, interdependent technologies (Knorr-Cetina, 1999; Garud and Rappa, 1994).

communities increases the porosity of firm boundaries around the creation and transmission of knowledge. Brown and Duguid (2001) point out that "as a result of the two strands of identification – one in the organization, one in the network – members of a particular community are likely to have divided loyalties... which strand communities favour or disfavour will have a significant effect on the direction in which knowledge flows".

The problem of knowledge creation and exploitation in the context of work-based communities has recently resurfaced in studies in studies of science-based firms, which primarily depend on research scientists to innovate. Rather there being clear distinctions between proprietary knowledge created by rent-seeking firms and the "open" knowledge of academic communities, there is evidence to show that there are complex interdependencies linking their co-evolution (Murray 2002). Firm scientists collaborate with university scientists to "remain in the flow" of scientific information and improve the productivity of their own research (Cockburn and Henderson, 1998). Conversely, university scientists (often "star scientists") are a strong driver in the biotech industry as they start-up or link closely with biotech firms while maintaining their employment in academe. Given this intense cross-fertilization it is interesting to note conflicting tensions between the logic of the rent-seeking firm and the norms of the epistemic communities to which the researchers belong. The mechanisms by which knowledge is jointly created by these overlapping communities of science and technology are just beginning to be explored and their future elucidation hold significant promise for the theory of the firm.

2.4. Conclusions

This chapter began by reviewing the classical theory of the firm and its shortcomings. In particular, the role of management and firm heterogeneity paved the way for a more realistic theory of the business firm. Following on Coase's (1937) seminal work developed by transaction cost theorists; we reviewed various theories in organizational and industrial economics that attempts to build a theory of the firm with the explicit motivation of understanding how heterogeneous firms build and sustain competitive advantage. Of these theories, the resource based view is seen as the most promising with the emphasis on the role of knowledge as a rent bearing resource. While there is significant empirical evidence to support its overall thesis, work lies ahead in delineating the specific mechanisms by which firms create and deploy resources in their search for economic rent. Particularly, challenges lie in unlocking the processes that govern knowledge-based resources. As a first step, it is necessary to understand how knowledge can be measured at the level of the firm. I expand upon this topic in the next chapter.

Chapter 3

Measuring Knowledge

3.1. Introduction

Economic theory has firmly placed "knowledge" centre-stage as the key economic asset that drives long-run economic performance. Yet "knowledge", "innovation" and "technological change" are elusive notions, difficult to conceptualize and even harder to measure in a consistent, systematic way. Thus while economists from Adam Smith on have amply recognized their crucial role in shaping the process of economic growth, our ability to study these phenomena has been rather limited. The last several decades have seen a number of pioneering attempts to overcome these measurement problems and gather data that can be used for the systematic empirical analysis of technological change. This paper reviews the progress made in this direction. Section 3.2 reviews the use of patent data focusing on its use in determining corporate performance in high technology industries. A major contribution of this thesis is to extend the analysis of innovation to include bibliometric measures of scientific publications as a measure of the science-base in high technology, science-heavy industries such as biotechnology. I review the use of bibliometric measures of innovation in Section 3.3. Section 3.4 concludes.

3.2. The use of patent data

3.2.1. An overview

The origins of the quantitative analysis of technological change lie in the immediate post WW2 period. The path-breaking findings of Abramowitz (1956) and Solow (1957) that there was a large "residual" of aggregate productivity growth that could not be explained by capital accumulation opened up a whole new and exciting research frontier. In parallel, and responding to the challenge posed by the productivity black box identified in those studies, empirical microeconomic analysis of the underlying phenomena of invention and innovation were also undertaken. A landmark volume, edited by Richard Nelson (1962). *The Rate and Direction of Inventive Activity*, brought together these early lives of enquiry and set the agenda for future work in this area. Nelson's volume (1962), best known perhaps for the classic paper by Kenneth Arrow that formalized the market failure inherent in research, contains also a less-cited but visionary paper by Simon Kuznets[15] on the difficulties of measuring the results of the results of the inventive process. Kuznets' paper raised many of the issues that permeate the study of technological change to this day. He discussed the problems of defining and measuring the magnitude of inventions; the relationship between the technological and economic significance of an invention; the distinction between the cost of producing an invention and the value it creates; and the consequences of the highly skewed distribution of invention values.

[15] Kuznets, S., 1962. Inventive activity: Problems of definition and measurement. In Nelson, R.R. (ed), The Rate and Direction of Inventive Activity. Princeton University Press.

A parallel pioneering stream of literature can be seen in the work of Jacob Schmookler (1966). Schmookler methodically went through (non-computerized) patent records to compile hundreds of the time-series of patent totals by industry, going back over a century. He also gave careful attention to the methodological issues arising from the use of these data, particularly the difficulty of identifying patents with particular industries based on their technological classification by the patent office. Using these data, Schmookler provide strong evidence for the role of market forces in shaping the rate and direction of inventive activity. More important in the long run, he demonstrated that patent statistics, though perhaps cumbersome to accumulate and subject to issues of misinterpretation, provide a unique source of systematic information about the incentive process.

Building on those early pioneers, in the late 1970's, Zvi Griliches took advantage of the computerization of the US Patent Office records, to launch a major research initiative on the innovation process. This research produced important conceptual developments in modelling the research process and the role of patents over the next ten years. In an important step that brought theory closer to the empirical world, Griliches (1979) and Griliches and Pakes (1984) extended and refined the concept of "the knowledge production function", a stochastic relationship in which current R&D investment, the firm's existing stock of knowledge, and the knowledge from other sources combine to produce new knowledge. Patents can be viewed as a noisy indicator of the success of this stochastic knowledge production process, with the "propensity to patent" – the ratio of patents to the unobservable knowledge production – possibility varying over time and

institutions. Griliches (1979) also suggested that the possibility of excess social returns in research should be explicitly modelled in relationship to flows of knowledge between and among different economic agents.

Schankerman and Pakes (1985, 1986) took another original track, using information on fees paid for the renewal of patents in European countries. These data allowed them to estimate the distribution of (private) patent values, as induced by the frequencies of renewal and the magnitude of renewal fees at every stage. This line of research provided firm empirical evidence on the extent of heterogeneity in patent values, and also stimuli for further research using novel aspects of patent data (Pakes and Simpson, 1989).

3.2.2. Patents

A patent is a document, issued by an authorized governmental agency, granting the right to exclude anyone else from the production or use of a specific new advice, apparatus, or process for a stated number of years. The grant is issued to the inventor of this device or process after an examination that focuses on both the novelty of the claimed item and its potential utility. The right embedded in the patent can be assigned by the inventor to somebody else, usually to his employer, a corporation, and/or sold to or licensed for use by any interested party. This right can be enforced only by the potential threat of or and actual suit in the courts for infringement damages. The stated purpose of the patent system is to encourage invention and technical progress both by providing a temporary monopoly for the inventor and by forcing the early disclosure of the information

necessary for production of this item or the operation of the new process.[16] The standard of novelty and utility imposed on the granting of such a right is not very high. Although there is some variation across countries and through time, roughly speaking two out of three applications are eventually granted (Griliches, 1990). Thus the number of patents is currently in excess of 6 million, and the flow is over 150,000 patents per year (as of 1999-2000).

Each patent produces a highly structured document containing detailed information on the innovation itself, the technological area to which it belongs, the inventors (e.g., their geographical location), and the organization (if any) to which the inventors assign the patent property right. Patents also include references or citations to previous patents and to the scientific literature. Unlike bibliographic citations, patent citations perform an important legal function, in helping to delimit the patent grant by identifying "prior art" that is not covered by a given patent grant.

Several problems crop up in using patents for economic analysis. The first is primarily a technical problem. How does one allocate patent data organized by firms or by substantive patent classes into economically relevant industry or product groupings? The second problem is fundamentally much harder to resolve: it refers to the obvious fact that patents differ greatly in their technical and economic significance. Many of them reflect minor improvements of little economic value. Some of them, however, prove, extremely valuable. How does one then attach "weights" to patents that capture this heterogeneity?

[16] There is an extensive literature on the relative merits and demerits of the patent system as a means of encouraging innovation. However, this discussion is outside the scope of this paper. See Oddi, 1996.

In resolving the first problem, one has to face the inherent ambiguity in the task. Do we want to assign the invention to the industry in which it was made ("origin"), to the industry that is likely to produce it, or to the industry that will use the resulting product or process and whose productivity may benefit thereby (destination or industry of "use")? Consider, as an example, the case of a new plow invented in a chemical firm's research laboratory as part of its project on new combined fertilizer and tillage systems. It depends on what questions is to be asked of the data. If we want to study the returns to R&D expenditures we may wish to count it in the chemical industry whence the money came to develop it. If we want to analyze the impact of technological change on the rate of investment, on the sale of new equipment, we may want to count it in the form of equipment industry. If we are interested in its effects on measured productivity we are more likely to count it as being relevant to agriculture.

The analysis of firm-level data; for example, relating patents to R&D investment and the subsequent fortunes of the firms where they had been originally developed; presents its own problems. The extensive diversification of many firms and also the various merger waves create severe technical problems in trying to use the patent data even at the individual firm level. What is noted on the patent is the name of the organization to which it has been assigned. This organization can easily be a subsidiary or a separate division of a larger company. Further, a company may change its name and/or may merge. In high technology industries such as biotechnology; patents are often licensed to third parties or owned by individual inventors or universities who license them to

biotech firm and/or pharma. This overlapping network of ownership makes it difficult to assign patents to individual firms in a straightforward way.

The second, more fundamental difficulty arises because the economic significance of individual patents is so variable making it hard to estimate the average value of patent rights, the average value of the invention represented by a particular patent, and the dispersion in both these concepts. Looking at patents as indicators of success of the underlying inventive activity, we are mainly interested in the second concept (i.e. average value of the invention represented by a particular patent). The available data, however, are mostly informative only about the first: the value associated with the differential legal situation created by the possession of the patent.

There are basically three sources of data on this topic: 1. Results of direct surveys of patent owners or assignees about past returns and the potential market value of those rights, 2. The valuation implicit in the decision, whether to pay a fee to renew the patent, 3. Econometric analyses of the relationship of some other value-denominated variable, such as profits of stock market value, to the number of patents (An example is the use of patent numbers as a proxy for "intangible" capital in stock market value of the firm regressions). It is this last line of research that this paper will focus on.

The use of stock market values as an "output" indicator of the research process has one major advantage. All other indicators of success, such as profits or productivity, are only likely to reflect it slowly and erratically. On the other hand, when an event occurs that causes the market to re-evaluate the accumulated output of a firm's research

endeavours, its full effect on the expected present vale of a firm's future net cash flows is recorded immediately. This, of course, need not equal what will eventually materialize. The downside of this type of measurement is the large volatility in stock market measures. The needle might be there but the haystack can be very large.

The simplest market value model starts from the market valuation identity, with the market value of the firm proportional to its physical ("tangible") and intangible capital, the latter being in the part the product of its past R&D investments and possibly also reflected in its accumulated patent position (Griliches, 1981; Ben-Zion, 1984; Cockburn and Griliches, 1988). The model, in its simplest form, can be written as:

$$V = q(A + gK) = qA(1 + g\,K/A) \qquad \text{(Equation 3.1)}$$

where V is the market value of the firm, A is the current replacement cost of its intangible assets, K is the level of intangible ("knowledge") capital and g is the relative shadow price, and q is the current premium or discount of market value over the replacement cost of tangible assets.[17]

Writing q as $exp(a+u)$; where a represents individual firm differences in average valuation due to the exclusion of other unmeasured capital components or market position variables, taking logarithms, and approximating $log(1+x) = x$, we can rewrite the estimating equation as:

$$\ln Q = \ln(V/A) = a + q\,K/A + u \qquad \text{(Equation 3.2)}$$

[17] This equation would hold exactly in a world in which all assets were fully traded in the same market. More generally, such an equation is valid in a multi-capital setting only under very stringent conditions, such as the linear homogeneity of the profit function. See Wildasin (1984) and Hayashi and Inoue (1990) for more discussion.

where the dependent variable is now defined as the logarithm of Tobin's Q (Wildasen, 1984).

Using different measures of current and past patents and current and past R&D expenditures as proxies for *K*, it is possible to estimate this kind of equation. It has been observed that for the most part, R&D is the "stronger" variable. The evidence for additional information in the patent variable varies from sample to sample and depends on which variables are included in the equation (Griliches, 1990). Further, patents are estimated to contain a significant noise component (a component whose variance is not related to either the R&D or the stock market rate of return. This noise component accounts for only a small fraction of the large differences in the number of patent applications of different firms (about 25% in study by Pakes, 1985) but plays a much larger role among the smaller fluctuations that occur in the patent applications of a given firm over time (about 95%). Similarly, the effect of unexpected increases in patents on market value is highly variable. Nevertheless, there is still some information in the time-series dimension. If we were to observe, for example, a sudden large burst in the patent applications of a given firm, we would be quite sure that events have occurred to cause a large change in the market value of its R&D program. By the same token, smaller changes in the patent application of a given firm are not likely to be very informative.

The timing of the response patents and R&D to events that change the value of a firm's R&D effort is quite similar. One gets the impression from the estimates that such events cause a chain reaction, including an increase in R&D expenditures far into the future,

and that firms patent around the links of this chain almost as quickly as they are completed, resulting in a rather close relationship between R&D expenditures and the number of patents applied for. Perhaps surprisingly, Pakes (1985, 1995) finds no evidence that independent charges in the number of patents applied for (independent of current and earlier R&D expenditures) produce significant effects on the market's valuation of the firm. Hence it is not possible to distinguish between demand shocks (where demand shocks are defined as events that cause increases in patenting only through the R&D expenditures they induce, and technology or supply shocks that may have direct effect on patents as well as indirect effect via induced R&D demand.

It is not at all obvious whether one can separate "demand" from "supply" factors in this area. Patent data is thus useful if one were willing to assume that independent, "unanticipated" shifts in the level of patenting by firms represent shifts in technological opportunities and not responses to external shocks. That is, the identifying assumption is that the "news" component in the patent statistics reflects technological "news", the information that a particular line of research has turned out to be more (or less) fruitful or easier (harder) than expected when the decision to invest in it was originally made. Changes in technological opportunity are thus identified with "abnormal", "unexpected" bursts (or declines) in the number of patents applied for.

Several implications of this analysis are immediate. Id patent statistics contain additional information about shifts in technological opportunity, then they should be correlated with current changes in market value above and beyond their current relationship with R&D and they should affect R&D levels in the future, even in the

presence of the change in market value variable because the latter variable is measured with much error: in other words, patents should "cause" R&D (Pakes, 1985). The available evidence on this point is not very encouraging. Griliches (1981) did find a significant independent effect of patents on the market value of firms, above and beyond their R&D expenditures, but Pakes did not detect a significant influence of lagged patents on R&D in the presence of lagged R&D and the stock market rate of return variables. Nor did Hall, Griliches and Hausman (1986) find future R&D affecting current patenting as the "causality" argument might have implied. Griliches, Hall and Pakes (1990) replicate some of Pakes' computations on a larger sample (340 firms) and expand his equation system to add equations for sales, employment, and investment. Their results indicate that the addition of the latter variables is helpful, in the sense that fluctuations in their growth rates are related to fluctuations in both the growth rate of R&D and the stock market rate of return and hence should help in identifying the relationship we are interested in. but the expansion of the sample to include many small firms with low levels of patenting deteriorates significantly the informational content of this variable, raising its noise to signal ratio, and making it hard to discern a feedback from the independent variability in patenting to any of other variables.

In order to improve the "signal strength" of patent data, more recent approaches have capitalized on the computerization of patents to more fully exploit patent information. As noted previously, patent documents also contain citations to other previous patents. Advanced computational techniques now allow the search for all subsequent citations of

a particular patent.[18] The potential significance of patent citations can be inferred from the following citation:

"During the examination process, the examiner searches the pertinent portion of the "classified" patent file. This purpose is to identify any prior disclosures of technology ... which might anticipate the claimed invention and preclude the issuance of a patent; which might be similar to the claimed invention and limit the scope of patent protection ...; or which, generally reveal the state of the technology to which the invention is directed ... If such documents are found they are made known to the inventor and are "cited" in any patent which matures from the application ... Thus, the number of times a patent document is cited may be a measure of its technological significance"

- OECD Report, Benchmarking Science-Industry Relationships, 2002, p. 87.

Moreover, there is a legal dimension to patent citations, since they represent a limitation on the scope of the property rights established by a patent's claims which are enforceable in a court of law. Equally important, the process of arriving at the final list of references, which involves the applicant and his attorney as well as the examiner, does generate the right incentives to have all relevant patents cited, and only those (see Campbell and Nieves, 1979). Thus, the presumption that citation counts are potentially informative of something like the technological importance of patents is thus well grounded.

[18] It is important to note that this process sets up what is classically known as the "inverse problem". To count the citations a patent receives; it becomes necessary to check all subsequent patents issued to check if the patent in question has been cited.

The question is whether citation counts may also be indicative of the (ex-post) value of the innovations disclosed in the cited patents.[19] While this question can be only resolved empirically, one can examine the theoretical arguments that underlie this claim. Most patents cited are referenced in patents issued within the same narrowly defined field of innovation as the cited patent. The very existence of the those later patents attests to the fact that the cited patents opened the way to technologically successful line of innovation. Moreover, it presumably attests also to economic success (at least in expected value terms), since those subsequent patents are the result of costly innovation efforts undertaken mostly by profit seeking agents. Given that citations to a patent are counted for a period of years following its issuance, there should be enough time for the uncertainty regarding the economic value of the innovation to resolve itself. Thus, if citations keep coming, it must be that the innovation originating in the cited patent had indeed proven to be valuable.

We have previously seen that in the relationship between patents, patent citations and the stock market valuation of firms, patent counts add little to market value after R&D is included in a Tobin-Q type equation.

However, a significant relationship is found between citation-weighted patent stocks and the market value of firms (Hall, Jaffe and Trajtenberg, 2001). The market premium associated with citations appears to be mostly to the high increase in value as citation intensity increases. Hall et al (2001) showed that after controlling for R&D and the un-weighted stock of patents, there is no difference in value between firms whose patents

[19] This clearly need not be the same as technological importance: the latter could be thought of as having to do with the supply side of innovations, whereas value obviously reflects a market equilibrium.

have no citations, and those firms whose patent portfolio has approximately the median number of citations per patent. There is, however, a significant increase in value associated with having above-median citation intensity, and a substantial value premium associated with having a citation intensity in the upper quartile of distribution. Further, it is interesting to note that self-citations (i.e., citations from patents assigned to the same firm) are, on average, associated with about twice as much market value as citations from others (Jaffe and Trajtenberg, 2000). This supports the argument that self citations, because they represent subsequent building on the invention by the original firm, are indicative of the firm's capturing a larger share of the overall social value of the invention. Thus, the evidence shows that both social and private values are increasing in the citations intensity, apparently with increasing returns, and that a high rate of self-citation is indicative of a larger fraction of social returns occurring to the innovating firm.

More recent work has gone beyond stock market valuations in estimating patent value. Harhoff et al (1999) have confirmed the relationship between patent citations and "value" using survey-based measures of the value of specific important inventions. In the study conducted by Harhoff and his co-authors 1999, economic value estimates were obtained on 962 inventions made in the United States and Germany. Patent value was found to hinge on two factors – 1) patents that are renewed to full term expiration in environments such as Germany (which have highly progressive annual maintenance fees) are more valuable and more highly cited than patents which are allowed to expire before running to full term, and 2) within the relatively exclusive cohort of full-term

patents, citation frequency rises with economic value, although with considerable noise in the relationship.

Recent work by Jenny Lanjouw and Mark Schankerman (1999) explores the information content of patent citations relative to other indicators also derived from patent data, and examines the relationship of these measures to other economic variables. They construct composite measures of patent "quality" based on the number of citations received, the number of citations made by the patent, the number of claims in the patent, and the number of countries in which patent protection is sought ("family size"). They show that this measure is related to the likelihood of patent renewal and patent litigation, and to measures of the economic significance of a patent to its owner. Finally, they show that the quality-adjusted rate of patenting by firms exhibits a more stable relationship to firm's R&D expenditure than simple patent counts.

There is clearly room for further work on the meaning of and relationships among these different indicators of quality importance, and value. An important issue are the inter-relationships among the technological significance of an invention, the spillovers that it generates for future innovators (firms), and the value of the invention to its owner (firm). It remains to be seen whether the different measures of patent quality can shed light on these issues (beyond the self-citation effect mentioned above). One aspect of this is variations in patent "size", in the sense of different uses or applications for a single idea, as distinct from the intrinsic significance of the idea. Jaffe and Trajtenberg (2001) have shown that citations exhibit an interesting geographic pattern: initial localization that fades over time. But there is much more that could be done to further

explore these patterns. How important are "border" effects (continents, countries, regional districts) as distinct from physical distance? Does language matter systematically? How about historical, social and cultural connections? For example, recent work by Hu and Jaffe (2001) shows that Korea is "closer" (in terms of frequency of patent citations) to Japan than it is to the United States, and Korea is much closer to Japan than Taiwan is to Japan. These relationships appear to be consistent with patterns of institutional and historical connections in these pairs of economies.

Another idea raised by Jaffe and Trajtenberg (2001) is endogenous obsolescence. A patent that is highly cited is presumed "important", but it would also seem that the accumulation over time of many patents building on a given invention would eventually make it less valuable, at least in the private sense. In principal, it should be possible to implement a dynamic model of the process that might be able to shed light on the rate of private obsolescence of knowledge, and how that varies across different technologies or industries, as well as over time.

Finally, patents contain references to scientific literature as well as to other patents. Work by Fiona Murray (2002) has shown that there is a "co-evolution" of scientific and technical knowledge. In the next section, this chapter reviews recent work being pioneered in exploring how these interlinked networks of science and technology overlap.

3.3. Using Bibliometric Measures to map Innovation

Recent work, building on the early theorizing of Nathan Rosenberg (1982), has shown that technological and scientific knowledge co-evolve in overlapping networks. A recent paper by Murray (2002) makes a detailed examination of how two worlds overlap and interact. Murray's methodology examines the field of tissue engineering and uses patent analysis to create a map of people and institutions that develop the key ideas. The patent data reference both "prior art", patented ideas that the inventor has built his or her work upon and "forward citations", later inventions built on the ideas in the current patent. The study then extends its analysis to references of scientific papers cites on the patent to further understand the link of the patented idea to "public" scientific knowledge.[20] The study revealed that scientific and commercial progress arise in two distinctive yet overlapping networks – one predominantly scientific; the other a blend of individuals and institutions in science and business. The findings suggest that the links between the two communities are shaped by key scientists who engage in the practices of both in the range of activities that span patenting, consulting, advisory board membership, sponsored research licensing, joint publication and entrepreneurial ventures. According to Murray, these synaptic activities play a significant role in translating scientific progress into technological (and commercial) knowledge.

However, the use of scientific references listed on a plot as an indicator of the scientific knowledge it links to, is not without its methodological problems. Campbell and Nieves (1979) argue that due to their specific legal functions, citations in patents are less valuable as signals of cognitive debt than citations in journal papers. Thus, a major

[20] The study encompassed the detailed analysis of 76 patents and 158 peer-reviewed scientific papers, supplemented by in-depth interviews.

concern is that scientific references are arbitrarily chosen to complete a set of legal formalities, favouring a tendency to cite by role rather than relevance.[21]

Two recent studies, Meyer (2000) and Tijssen et al (2000), debate the relevance and reasons of scientific references in patents. Meyer investigated the reasons for citing scientific literature through a case study of ten patents in the field of nanotechnology. Among his major conclusions were that non-patent references cited on a patent do not accurately reflect their cognitive contribution to the invention. Thus in the case of references to scientific literature, there is little direct "antecedent" relationship between cited paper and citing paper. However, scientific references do line the patent to a wider body of scientific knowledge, even there is no explicit causal link. This conclusion is supported by Tijssen et al (2000) whose study on a set of Dutch USPTO (United States Patents and Trademark Office) patents confirm that non-patent citations are likely to indicate a cognitive debt but not necessarily a causal relationship between the patented technology and scientific knowledge.

Another measure of science and technology interaction are scientific publications by industry. As can be expected, traditionally companies patent more than they publish, and conversely university researchers publish more than they patent.[22] Despite this skew[23], a considerable number of papers result from scientific co-operation and are

[21] Patel and Pavitt (1997) report that business firms are granted about 80% of all patents and many of the remaining 20% go to individual owners or small businesses.

[22] Within the patent system, the legal responsibility of the applicant is to describe the prior art and the "originality" of the patents (Meyer, 2000). This conditionality requires the applicant to describe the background of the patent in such a way as to show that the claimed invention relates to, but is innovatively different from what was already public knowledge.

[23] According to De Solla Price (1965), science and technology differ substantially in their central activities due to the different ultimate objectives that motivate these activities. Scientists publish to

often co-authored by scientists who work in both academia and industry. Firms publish scientific research in order to signal their attractiveness as a research environment in the labour marker as well as to maintain links with the cutting edge of academic knowledge (Hall 2000). This behaviour is particularly relevant to "science-based" industries such as biotechnology where the distinction between university and corporate laboratories are often blurred (Hicks and Katz, 1996).

Other research supports the notion that industrial publications are an indicator of competitive advantage that contributes to the overall performance of the firm. In their study of pharmaceutical firms, Decarolis and his colleagues use scientific publications as one of several measures (products under development is another example of these measures), to track the performance of individual firms. Citation counts of papers written by a firm's researchers appear to be positively correlated with a firm's overall performance (Decarolis and Leeds, 1999).

However, the analyses of industry's or business firm's scientific output suffer from shortcomings generally associated with bibliometric indicators. Hicks and Katz (1996) outline these shortcomings: The key drawback of bibliometric data as an output signal of scientific activity is that scientific knowledge rapidly becomes common knowledge and thus key results are often referred to in the literature without citation. Conversely; often scientific work is not always acknowledged by contemporaries of rival "camps" (Price, 1965). Further scientific fields vary intensely and thus the probability of being cited varies as well. This is further accentuated by the fact that the propensity to publish

---

maximize their visibility and recognition whereas the technologist's objective is to construct or design a proprietary artefact from which he can extract rent.

and cite varies significantly between scientific disciplines (Katz,1996). Lastly; scientific activity does not proceed on a linear curve but rather in "bursts" of scientific advance (Kuhn, 1962) and subsequently both the value of scientific papers and the number of citations they garner do not follow a linear rate over time making time-series comparison problematic. Lastly, scientific papers represent only one output of scientific activity. More informal exchanges that form the bulk of the "tacit" knowledge-base that moves scientific progress forward goes un-detected in bibliometric measures.

3.4. Conclusion

This chapter reviewed the different ways by which knowledge can be measured focusing on the firm as a unit of analysis. The primary data available to do so is patent data as well as bibliometric data of scientific publications by firms. Of these, the former has been more thoroughly explored. The latter measure, while still in its infancy of use points to an important methodological signpost; increasingly scientific and technologically knowledge are interwoven in an intricate matrix and understanding of how scientific knowledge spurs industrial innovation and finally, economic growth. While the work of Murray and others has opened up this rich and exciting area of research: this work is still in its infancy. Extending and integrating this work from the "frontline" of industrial innovation to fundamental epistemological notions of knowledge is a challenge that has not been fully realized. It is this interdisciplinary co-mingling of theory and empiricism that is the focus of this thesis.

Chapter 4

Private-Public Collaboration and Research Quality: Do Firms Produce Better Quality Research with Greater Academic Collaboration?

4.1. Introduction

As seen in Chapter 1, economic models of technological change have traditionally assumed the relationship between basic scientific research and technological invention to be unidirectional (Romer, 1990). Publicly funded universities and research institutes advance basic science by producing ideas. This body of knowledge is freely accessible in scientific journals and has the attributes of a public good. Business firms, drawing from this resource, engage in research and development in the search for technological inventions. These inventions, protected by patents, then generate rents for the firm (Rosenberg, 1988). However, recent literature exploring the interaction of science and technology has altered this fairly simplistic picture by highlighting the blurred boundaries between scientific research and technological invention (Murray, 2002; Rosenberg 1988).

Biotechnology is a science-based industry (Meyer-Krahmer and Schmoch, 1998) characterized by collaboration between academic[24] institutions and firm-based scientists (e.g. Blumenthal et al., 1996; Gittelman and Kogut, 2003; Liebeskind et al., 1996; Zucker and Darby, 1997). On the one hand, traditional repositories of scientific knowledge such as universities and research institutes have shifted along the basic-

[24] In this paper, the terms "academic," "public" and "open" are used interchangeably to refer to scientific research performed outside of business firms.

applied spectrum and are increasingly involved in commercializing their science via patenting, licensing and high technology spin-offs (Hicks, 2002). Parallel to this trend, high technology firms have adopted "open science" academic norms and now routinely engage in basic scientific research and employ scientists who regularly publish in scientific journals. In doing so, industry scientists often collaborate with university researchers in solving particular scientific problems.

Explorations of the co-mingling of scientific and technological networks in biomedicine indicate co-evolution via interlinked networks of scientists that bridge the private-public divide (Murray, 2002). More recent work by Murray (2004) highlights the critical role of academic inventors' social networks to entrepreneurial firm growth. Zucker et al. (1998), in their study of the US biotechnology industry, show that university-based star scientists played a key role in the birth and growth of the industry through dual roles as entrepreneurs and research scientists. See Carayol (2003) for a survey of the literature on both firm and scientist motivations, a typology of collaboration and the matching process and Mora-Valentin et al (2004) for the different success factors of firms and research organizations. Thus, there exists a substantial and growing body of work that points to the increasing value of private-public interaction in the evolution of science and technology and in the performance of firms and industries.

What is still missing, however, is research that delves into the effects of this private-public interaction on the resulting quality of scientific output. By quality, I refer to the impact of research on the academic community as measured by citation counts of the article and the journal. This constitutes an important gap in the literature since the

specific mechanisms by which private-public collaboration might improve firm, industry or even region-level performance are not well known. Clearly, improvements to research quality brought about by private-public collaboration is one obvious channel by which firms and industry may benefit from co-mingling with academic expertise. It is this assertion, therefore, that is in need of further exploration.

This chapter attempts to answer the question of whether private-public collaboration produces better quality research by comparing the effect on research quality of private-public collaboration using unique data from the UK biotechnology sector from 1988-2001. Specifically, I test whether the quality of scientific research undertaken by UK biotechnology firms is improved by more intensive collaboration with academic institutions. My findings suggest that collaborative research involving greater shares of academic participation does indeed improve research quality, although the nature of the biotechnology firm in question is an equally important factor in determining how strong a positive effect private-public collaboration has on research quality.

This study makes a number of contributions. First, it strengthens arguments for the tacit nature of knowledge (Polanyi, 1967), particularly in science-based industries such as biotechnology (Pisano, 1997). If knowledge in the biotechnology sector were not tacit in nature, then it would not matter who collaborates with whom. To the extent that it *is* tacit, then the strength of the quality leap when private sector science works with public sector science will be comparatively large.

Second, the gains from collaboration are evidence not only of the tacit nature of knowledge but also its heterogeneity across private and public sectors. If scientists in both sectors possess largely the same set of skills and technical know-how, I would not expect the strong gains from collaboration that I see. This suggests that the "markets of knowledge" in the academic and private sector are sharply heterogeneous. Thus the scientific labour market can be thought of as composed of two distinct markets: one composed of academic scientists and another of scientists employed in industry, sorted on the basis of individual preferences. (For example, Stern (1999) suggests that academic scientists could trade off salary for scientific autonomy by working in academia.) Modern public policy increasingly ignores this distinction by encouraging universities to raise capital by working as de-facto external research and development (R&D) labs for business firms (Dasgupta and David, 1994). This study extends this market driven approach to science by highlighting the gains that result from collaborating *across* the divide not by eliminating the borders altogether. For a deeper discussion of the policy debate see Hicks and Katz (1996). Finally, this study answers recent calls for research on private-public collaboration (Carayol, 2003; Mora-Valentin et al., 2004) and the usefulness of scientists' knowledge (Murray, 2004; Sorenson and Fleming, 2004).

This chapter is organized as follows: Section two provides a literature review of the competing incentives for "open science" versus that of private R&D, the issues that arise in measuring their interaction and the UK biotechnology industry. Section three describes an empirical model used to estimate the impact of private-public collaboration on research quality. Section four provides a description of the data and variables.

Section five discusses results with tests of robustness. Section six concludes with implications for R&D policy and practice and suggestions for future research.

4.2. Literature Review

4.2.1. Open Science Versus Private R&D

Economic debate on the role of publicly funded science begins with the observation that there is a substantial gap between social and private returns to R&D, particularly in the case of "basic" research. Many researchers have identified externalities arising from the public good aspects of knowledge as the source of the gap (Arrow, 1962). The idea that the inability of profit maximizing firms to appropriate the full economic returns from R&D is likely to lead to under-investment in research relative to the social optimum remains the basis for substantial public support of R&D in general, and basic scientific research in particular (Arora et al., 1995).

This broad logic could explain the scientific division of labor wherein universities (largely subsidized by the government) working under the "open science" paradigm, publish scientific articles that are freely available in publicly accessible journals. Firms on the other hand, drawing from this "public" body of scientific know-how, invest in R&D, mostly in areas of applied research, in order to secure valuable patents that would generate rent.

Recently this "waterfall" model has come under criticism, with the distinction between university "open" science and firm level R&D increasingly seen as a blurred one (Murray, 2002). The biggest problem for the "waterfall" model is that some for-profit firms organize their research in ways that mimic the practices found in universities or publicly funded organizations (Cockburn and Henderson, 1998; Dasgupta and David, 1994; Gittelman and Kogut 2003). In particular, large pharmaceutical and biotechnology firms have been found to rely heavily on collaboration with academic scientists to improve research productivity and regularly publish scientific articles in the open science paradigm (Paula, 1996; Rosenberg, 1990). These findings challenge the traditional understanding of the distinction between public and private science and raise the question of why would companies allow their work (or the work of their scientists) to be freely available?

4.2.2. Why Do Private Firms Adopt Public-Science Norms?

Several strategic advantages have been identified with the private sector adoption of open science, including the development of absorptive capacity (Cockburn and Henderson, 1998), labor costs reduction (Stern, 1999), and enhancing firms' competitive position in a patent race (Lichtman et al., 2000; Parchomovsky, 2000).
Drawing upon Cohen and Levinthal's (1990) "absorptive capacity" argument, Cockburn and Henderson (1998) suggest that firms use pro-open-science incentives to develop routines and skills that allow them to utilize effectively the advances in publicly funded research. The interviews with senior scientists and management of pharmaceutical firms conducted by Cockburn and Henderson (1998) indicated that firms strived to develop

such a capacity by recruiting and rewarding scientific employees based on their standing in the hierarchy of public-sector science and by encouraging them to actively engage themselves in the academic community.

Ties to the academic community have also been found to underlie the innovative activities of biotechnology firms (Zucker et al., 1998). Assuming university-affiliated scientists prefer research projects that will lead to publications, adoption of open science may help firms attract high-quality academic collaborators. In addition, Stern (1999) suggested that there might be a labor cost advantage associated with pro-publication firms. Based on the analysis of offers accepted by a sample of postdoctoral job applicants, he showed that scientists are willing to accept a lower wage in exchange for permission to keep up with research in high quality basic science. Game theorists have suggested that publishing discoveries may be used by firms engaged in a patent race to establish a convincing history of prior art in competing claims of discoveries (Lichtman et al., 2000; Parchomovsky, 2000).

While there is a significant and growing body of literature that investigates the theoretical forces driving the private-public interaction in the production of scientific knowledge, the empirical literature is less well developed. Though measuring the research output of "open science" and its impact on the rest of the economy presents enormous challenges, both quantitative and qualitative estimates suggest that the rate of return to basic research is probably quite high. Direct quantitative estimates rate of return on the order of 25 to 40 percent (Adams, 1990; Griliches, 1994; Mansfield, 1989). Also studies of university-industry relations have found a positive effect of

university research on private sector R&D (Jaffe, 1986; Mansfield, 1989). However, most of these models focus on the training and education function of the university sector rather than its research output.

There is also a stream of literature in economic history that points to the critical role of public sector research in laying the foundation for technological advances that have had enormous impact on the economy (see, for example, David et al., 1992). An example of this is the US pharmaceutical industry, one of the most science-intensive sectors of the economy and one where public support for research has been very substantial (Cockburn and Henderson, 1996). Further, the biotechnology industry was incubated within academic science and close links between academic science and industry continue to be commonplace in this industry. While it seems clear that the industry's rapid rate of technological change and impressive economic performance rests on a foundation of long term publicly funded investments in basic science (Comroe and Dripps, 1976; Ward and Dranove, 1995), attributing specific tangible payoffs to these investments is difficult. Furthermore, there has been no work that I could identify in the literature measuring the impact of private-public collaboration in producing scientific knowledge on research quality, a question that I address using evidence from the UK biotechnology industry.

4.2.3. How to Measure Private-Public Collaboration?

Unpacking the relationship between "open science" and industry R&D poses several methodological challenges. Pioneering work by Narin and Rozek (1988), Narin and

Olivastro (1992) and Penan (1996) used citation patterns very successfully to trace interaction among researchers and across organizational boundaries. However, citation analyses present a number of difficulties. Citation is often highly ritualized, occurs with variable and often very long lags, and may represent negative as well as positive acknowledgement of previous research. Another more direct measure of interaction is joint-authorship of papers. Often, scientists working in industry collaborate with academics on research publications. This pattern of co-authorship probably represents the strongest empirical record of the interaction of private firms and academic science (see Zucker and Darby, 1995; Zucker, et al., 1998 and Liebeskind et al., 1996).

Furthermore, as it is often pointed out, citation in the age of the word processor and computerized databases is extremely cheap and easy. By contrast, as many researchers can testify from personal experience, joint authorship is costly in terms of effort as well as other resources. In order to be willing to collaborate on a paper, all authors must be willing to incur these costs, making an instance of co-authorship a stronger empirical signal than a citation. Further, I believe that co-authorship is also evidence of a qualitatively different kind of interaction than does citation. Joint authorship often reflects joint research, which is an opportunity for the exchange of *tacit* knowledge (Katz and Martin, 1997). By contrast, citation may be seen as an acknowledgement of the exchange of codified knowledge. Citation also refers to old knowledge, whereas co-authorship reflects generation and exchange of new or current knowledge. Thus while citations can often be an impersonal referencing to existing knowledge, I see co-authorship as evidence of joint problem solving and something that represents a much more significant investment on the part of the firm.

Naturally, there are some difficulties with this interpretation. Clearly, co-authorship does not capture the entire range of active knowledge exchange among scientists. Academic researchers and industry scientists often read each others' work, correspond informally, listen to conference presentations, serve on professional committees together and so on, which all may serve as legitimate conduits for knowledge exchange. Further, co-authorship may also reflect a variety of things other than exchange of information and joint problem solving. It may be offered as a quid pro quo for supplying information or resources such as money or research materials (Cockburn and Henderson, 1996). It may also serve as a way to acknowledge intellectual debts and in the physical and biological sciences to list laboratory directors or other senior project leaders as authors on papers which they may have had very little involvement in writing (Murray, 2002). Notwithstanding these issues, this study proceeds on the assumption that co-authorships represent evidence of a significant investment on the part of the firm in developing connections to publicly funded "open science" research.

4.2.4. U.K. Biotechnology Industry

Biotechnology is the industrial application of biological processes. The biotechnology industry has been a source of both controversy and excitement over the last two decades. Whilst the origins of modern biotechnology lie in scientific advances made in the UK, such as discovery of the double helix by Watson and Crick in 1953 and the development of monoclonal antibodies by Milstein and Koehler in 1978, it was in the U.S. that commercial biotechnology emerged as an industry in the late 1970s. The UK was the next country to follow when commercial biotechnology was developed

following a governmental push-policy led by the new conservative government under Margaret Thatcher during the 1980s (see DTI, 2005).

The early lead established by the U.S. continues till today. However, since the 1980s, the United Kingdom (UK) has established itself as the European market leader in the biotechnology industry and remains the second largest in the world market, after the United States. Its leading position in biotechnology is based on a long-established reputation for excellence in research in the biosciences in its universities, teaching hospitals and research institutes. Scientifically, the primary strengths of the industry include therapeutics and human health applications with a particular emphasis on genomics and gene therapy, stem cell research, and drug discovery. In 2002, the UK Parliament voted to allow therapeutic cloning and end a ban on stem cell research using human embryos. As a result, additional funding to support research is expected from both government and private sources and the UK is now positioned to lead the world in stem cell research. The UK pharmaceutical industry is also the fourth largest in the world, accounting for 12% of the world market (House of Commons, 2003).

Biotechnology in the UK is characterized by strong multinational corporations and a vibrant SME sector reliant on mergers and partnerships with large pharmaceutical firms. As of 2003, the UK biotechnology industry comprised of 455 companies employing a total of 9,600 people in research and development and generating .3.6 billion in revenues (DTI, 2005). Some comparative statistics are provided in Table 4.A.

Table 4.A. The Biotechnology Industry in Selected Countries in 2002.

	UK	Germany	France	USA
Revenue (£M)	2860	665	515	16,099
Employees	23,650	14,408	unavailable	141,000
No. companies	481	430	330	1467
No. public companies	46	17	3	380

Source: Data collated by author from OECD, 2002; House of Commons, 2003; DTI, 2005.

It is through the emergence of scientific discoveries, primarily from academic institutions, that potential products and techniques for commercial biotechnology are identified. It is also at the level of basic research that government involvement and funding is at its most visible; despite the involvement of charitable foundations, government funding of basic research remains the bedrock of the industry. That the UK was able to make an early start in commercial biotechnology is in no small part due to its traditional strength in research in the biosciences. UK universities and research institutes have established themselves at the forefront of biotechnology research; for example, the MRC Laboratory of Molecular Biology in Cambridge (and its forerunners) alone has had 13 Nobel Laureates as members of its faculty.

In the UK, government expenditure on biomedical research and has fall in real terms and private investment plays an expanding role in funding public domain science (Lewison, 1997). On the private side, business expenditure on R&D in the UK in the late 1990s was 7.2 percent, above the US (6%) and Japan (2.4%), but below Germany (9.7%) and Belgium (10.6%) (OECD, 2002). However, the UK government has recently announced plans to reverse years of under investment in higher education and basic research (House of Commons, 2003).

One area where the UK has a clear advantage over countries such as Germany and the USA is the regulatory framework within which biotechnology research is conducted. A major constraint on the development of biotechnology in Germany in the past, for instance, was the restrictive regulatory regime that made certain areas of research very difficult. Although regulation has been liberalized recently, areas like stem cell research are still far more strictly regulated in other countries as compared to the UK. In contrast, the UK can be seen to have a comparatively liberal framework of regulation. Many scientists and that I interviewed informally as part of my research based in the US were envious of the relative freedom enjoyed by British scientists. The prospect of tight restrictions being imposed on biotechnology research is seen by scientists as a real threat and it appears that they are prepared to move to avoid these restrictions. For example, I was told of a prominent biotechnology research team that had abandoned Massachusetts for California for this reason. If tighter restrictions were imposed more generally in the USA, the UK would presumably prove a reasonably attractive alternative and a migration of research expertise could take place. However, the more liberal regime in the UK cannot be taken for granted and countries such as Singapore

now form a genuine alternative. Regulation of aspects of biotechnology research has recently been discussed at European Union level and public opinion in a number of member states favors stricter regulation — or even outright bans — on some types of research. Whilst the UK has so far preserved its right to impose its own regulatory regime, pressure for the setting of standards at European level is likely to continue. Policies vary across the EU and it remains a contentious issue for some member states.

Whilst excellence in research is a necessary condition for a flourishing biotechnology industry, it is not a sufficient one. Ultimately, success depends on commercialization of good science. The key factors that underlie successful commercialization in biotechnology are 1) entrepreneurial culture among scientists, 2) presence of venture capital, 3) support of public markets for biotechnology firms and 4) government support (such as choosing the correct regulatory framework).

While culture is a difficult variable to capture, in my interactions with several scientists across a range of institutions indicate that in the university environment going the "commercial route" is frowned upon. This is in contrast to the US where commercial activity by leading edge research scientists is the norm.

The biotechnology industry relies on venture capital for its survival. In a sector dominated by firms with little in the way of tangible assets to act as collateral for loans — their intellectual property and the know-how of their staff are their primary assets — and which require a considerable quantity of money to sustain themselves, venture

capital provides the main route of funding. The venture capital support for the UK biotech sector is relatively (relative to other European countries) robust and raised £392 million in equity investment which was more than any other European country (House of Commons, 2003).

The conventional route by which venture capitalists have sought to exit their biotechnology investments has been either through trade sales or through IPOs (Initial Public Offering) which is the preferred route. The UK has a robust financial market that is welcoming to biotechnology firms. Further, since 1993, under pressure from the growing industry, the London Stock Exchange agreed to make special provisions for biotechnology companies allowing them a fast-tract to secure a listing on the exchange.

4.3. Empirical Model

I seek to answer the following question: Is the quality of scientific research by UK biotechnology firms improved by collaboration with academic institutions? In other words, does the quality of collaborative research (between biotechnology firms and academic institutions) as measured by citations and journal prestige tend to be of higher quality with increasing shares of academic participation?

In order to test this proposition I use a dataset of scientific publications in biotechnology that include at least one author from a biotechnology firm. However, this set of papers does not include purely academic papers, i.e. those set of papers with no biotechnology presence. Thus, in order to specify a robust econometric model, a study would have to

control for any bias that might exist from this sample selection. Specifically, it requires the teasing out of a prior question – namely, what characteristics mark out papers that involve some public and biotechnology firm interaction rather than those composed purely of academic scientists? Contingent on this selection, I then test for the effect of increased academic participation on research quality.

4.3.1. Model Specification

This research begins with a simple specification:

$y_i = X_i\beta + \varepsilon_i$ (Equation 4.1)

where y_i is some measure of research quality (defined in Section 4.) and X is the vector of characteristics that include co-authorship, research level and other predictors of quality. However, as noted before, the papers included in the sample are those that include at least one biotechnology author and therefore may not be a random sample of research publications. This "selectivity" may bias the co-efficients in a standard OLS estimation of Eq.(2) (Green, 1981). A solution to this sample selectivity bias was found by Heckmann (1979). First, I write down a participation equation:

$T_i = 1(Z_i\gamma + \varepsilon_{0I} > 0)$ (Equation 4.2)

where Z includes variables that predict whether a paper has biotechnology involvement or not. Note that Z and X may include common variables. Thus a paper is included in the sample if $Z_i\gamma > \varepsilon_{0I}$. The selectivity problem is made apparent by taking expectations of Eq. (1) over the sample of papers with at least one biotechnology author:

$$E[y_i | X_i, T_i = 1] = X_i\beta + E[\varepsilon_{1i} | \varepsilon_{0I} > -Z_i\gamma] \quad \text{(Equation 4.3)}$$

If ε_0 and ε_1 are jointly distributed I can write

$$\varepsilon_{1i} = \frac{\varepsilon_{0,1}}{\sigma_0^2}\varepsilon_{0i} + \upsilon_i \qquad \text{(Equation 4.4)}$$

where υ_i is uncorrelated with ε_{oi}, $\sigma_{0,i}$ is the covariance between ε_{oi} and ε_{1i}, and σ_0^2 is the variance of ε_{0i}. This last observation is pertinent, because I can now write:

$$E\left[\varepsilon_{1i}\middle|\varepsilon_{0i} > -Z_i\gamma\right] = \frac{\sigma_{01}}{\sigma_0}E\left[\frac{\varepsilon_{0i}}{\sigma_0}\middle|\frac{\varepsilon_{0i}}{\sigma_0} > \frac{-Z_i\gamma}{\sigma_0}\right]$$

$$= \frac{\sigma_{01}}{\sigma_0}\frac{\phi(Z_i\gamma/\sigma_0)}{\Phi(Z_i\gamma/\sigma_0)} \qquad \text{(Equation 4.5)}$$

where $\phi(\cdot)$ is the standard normal density and $\Phi(\cdot)$ its cumulative distribution function. It is now evident why OLS estimates of Eq. (4.1) may be biased. In particular, the last expectation in Eq. (4.3) may not be zero. Selectivity bias is said to occur whenever σ_{01} is not zero. Heckman (1979) noted that the problem with using Eq. (4.3) is that $\hat{\beta}$ is generally biased owing to the presence of an omitted variable, where the quantity (usually referred to as the *inverse Mills ratio*) $\frac{\phi(Z_i\gamma/\sigma_0)}{\Phi(Z_i\gamma/\sigma_0)}$ is the omitted variable. If this variable were now to be included in the original specification as:

$$y_i = X_i\beta + \frac{\phi(Z_i\gamma/\sigma_0)}{\Phi(Z_i\gamma/\sigma_0)}\tilde{\sigma} \qquad \text{(Equation 4.6)}$$

then consistent estimates would be straightforward. An estimate of σ_{01}/σ_0 can be read off as the coefficient $\tilde{\sigma}$ on the inverse Mills ratio. Essentially, Heckman (1979) noted that such a model could easily be estimated with the following steps: First to run a probit of the treatment on the vector Z to obtain estimates of γ/σ_0. Then, these estimates are used to construct the inverse Mills ratio and finally, OLS regressions can then be run on Y using the *estimated* inverse Mills ratio as an additional regressor.

Standard errors are more complicated because the resulting model is heteroskedastic and uses estimated values. Merely adjusting the standard errors for heteroskedasticity will not be adequate in general, because such a correction fails to account for the loss in precision resulting from estimates of the inverse Mills ratio instead of actual values. One solution to obtain the correct errors is to use a maximum likelihood solution in the empirical estimation. This approach is used in the model. See Greene (1981) for a fuller treatment.

4.4. Data and Variables

4.4.1. Data Collection

The dataset of scientific articles analyzed in this analysis is sourced from the Research Output Database (ROD) that contains a record of published research outputs for all of UK biomedicine and for 32 biomedical sub-fields during the years 1988-2001. These data are sourced from CD-ROM versions of both the Science Citation Index (SCI) and the Social Sciences Citation Index (SSCI) produced by the Institute for Scientific Information in Philadelphia. The SCI is an excellent source because it covers a broad range of basic and applied scientific journals. Moreover it lists up to 255 authors and addresses for each publication, unlike other databases, which only include the institutional affiliation of the first author (e.g. Compendex, INSPEC and Biosis). Thus it is possible to construct data regarding the number of authors and addresses on individual papers, leading postcode areas on the paper, domestic and international

collaboration details as defined by author's addresses. Further, the SCI also lists any research funding sources.

The methodology whereby UK biomedical papers are identified and downloaded from the SCI and the SSCI is discussed in Lewison and Paraje (2004). Briefly, all papers with a UK address in biomedical and relevant social science journals are included, as are those with a biomedical address keyword in other journals. The dataset can reliably be said to contain the majority of UK biomedical papers published during 1988-2001 and comprises (no. of) 355 183 individual scientific papers (Lewison et al., 2003). Of these, 2 915 papers listed at least one author from a biotechnology firm as a contributing author (as identified by the address field) and are included in the analysis. The number of firms in my dataset is 203.

4.4.2. Variables

4.4.2.1 Dependent Variables

Research Quality. Research quality is the dependent variable and the most difficult and contentious item to measure. In the broadest terms, research quality may be defined in terms of how funded research feeds through to the welfare of society through health and wealth creation. Wealth might be reflected in the development of, for example, new pharmaceuticals, diagnostic regents or other medical technology. Health benefits, however, are manifest in better patient care and preventive measures based on regulation or advice. Therefore, the routes to health creation may be considered in terms of research papers leading to, for example, new techniques for diagnosis and treatment, improved medical education and training or better clinical care based on clear evidence-

based guidelines and recommendations (See Lewison et al., 2003). For impacts on wealth creation, one could examine rent seeking activities of firms as linked to their research and publishing activity. However, in my analysis I seek to determine how the community of researchers engaged in biomedical research might benefit from private-public collaboration and hence research quality is defined as the effect of research on other researchers. Thus, it may reflect the importance or quality of the research qua research but it is not necessarily an indicator of clinical utility or wealth creation.

If one defines research quality in this way, one can assess the importance of a paper by several measures – of which the two strongest indicators are its actual impact as determined by citation counts to the individual paper and its potential impact as judged by the journal in which it is published. The first and second measures are complementary but not identical, although papers in high impact journals tend to have more citations. Both are useful indicators and show how a paper has been judged by two different readerships: the general body of researchers and, in the second case; a journal editor plus a few specialized reviewers.

Research quality is signaled by citations (Trajtenberg, 1990). In the first estimate of research quality, the number of citation counts to individual papers over a five year period, from the year in which the paper is published is totaled and denoted by CIT. Previous research indicates that most citations occur in the first five years (Trajtenberg, 1990). The citations are tracked across scientific publications present in the ISI Science Citation Index (SCI). Only citations from original research articles were included, eliminating meeting notes, review articles and book reviews, to more accurately gauge

genuine impact. However, a better system may be to estimate citation scores received by individual papers on a log scale and use the transformation:

$$\log citations = 1 + 2\log_{10}(citations + 1) \qquad \text{(Equation 4.7)}$$

where *citations* is a five-year citation count for the individual paper. This variable is LOGCIT. Thus when $citations = 0$, then LOGCIT$= 1$, and for a high prestige journal like Nature, it equals about 5. This roughly corresponds to the relative subjective weight that scientists and scientific administrators give to papers in different journals in the two instances, which is not 100:1 but more nearly 5:1 (Lewison et al., 2003). I use this definition as a baseline estimation but also run the regressions with the simpler definition of citation counts as a test of robustness.

Journal quality. The quality of the journal is measured by an impact-weighted citation index (Laband and Piette, 1994). This index depends on the citations per character generated by articles in these journals, and weights the citations according to the citing journal. Thus the quality measure used is equivalent to an *expected* impact-adjusted citations measure. (See Laband and Piette, 1994, for a fuller discussion of the impact weighted citation measures).

To create a research quality variable based on journal quality (denoted PIC – Potential Impact Category), I use a simple system in which journals are put into four categories (which take values between 1-4) based on the mean number of citations to papers published in the journal in a given year that are received in the year of publication through the fourth year after publication. Journals with mean citation scores greater than 20 are classed PIC=4; 11 to 20 as PIC=3; 6 to 11 as PIC=2 and less than 6 as PIC=1.

These values were chosen so that approximately 10% of biomedical papers from the UK would have PIC=4, 20% would be PIC=3, 30% would be PIC=2 and 40% would be PIC=1 (Lewison et al., 2003). These measures were used in preference to "raw" impact factors (average number of citations in a given time period to papers published in a journal) because they are more likely to reflect the perceptions of scientific administrators and medical researchers. In two separate polls (Lewison et al., 2003), scientific administrators and researchers in the biomedical sciences, voted the relative importance of papers in "excellent" journals about four to five times that of papers in "ordinary" journals, and that of paper in "good" journals about two to three times of the latter.

4.4.1.2 Independent Variables

Collaboration. As discussed earlier, the primary measure of private-public collaboration is co-authorship, the joint authoring of research papers by individuals at business firms and from academia. Two measures of co-authorship are used: one aggregated at the level of the institution and another at the individual. First, the count of biotechnology firms present on a research paper (denoted by BIOTECH), and the number of collaborating academic institutions (ACAD). The number of *individual* authors from biotechnology firms is counted and denoted (AUT_BIOTECH), as are the number of authors from academia (AUT_ACAD). There are two measures of the impact of academic contribution on a biotechnology research paper: first, an estimate of the academic contribution at the institutional level as a simple ratio of academic institutions to the total number of institutions (denoted by ACADRATIO_INST). I then construct

the ratio of academic authors to the total number of authors (ACADRATIO_AUT) to capture the academic contribution to a publication at the *individual* level. There is also evidence that papers with greater number of total authors are associated with increased potential impact due to the both citation networks of individual authors as well as a self citation effect (Lewison and Devey, 1999). To control for this, I count the total number of authors on a paper (AUT). Further, to gauge diminishing returns for number of authors on an academic paper, I calculate the square of AUT (AUTSQ). The idea here is that too many "cooks" may spoil the broth.

Research level. There is evidence that research of a more "basic" nature tend to receive more citations than those of a more "applied" character (Cockburn and Henderson 1996; Lim, 2004). Therefore, a method is needed to capture the "basic" or "applied" nature of the research article. The prevailing trend in the literature is to use classification schemes based on the journal in which the article is published (Lewison et al., 2003; Lim, 2004). However, the price paid by this approach is the inability to capture heterogeneity among papers within each journal. This is a problem in a cutting edge area of research such as biomedical technology. Often, there is a blurring of boundaries between "basic" and "applied" papers and high prestige journals such as *Nature* or *Science* are highly multidisciplinary and cover both basic and applied topics. To overcome this limitation, I use the categorization developed by Lewison and Paraje (2004), based on the presence of one or more of about 100 "clinical" or "basic" words in the titles of papers in a given journal and year. This gives a decimal number between 1.00 (most applied) and 4.00 (most basic).

Institutional Affiliation. The institution to which an academic belongs can influence the quality of his research. Thus, academics of high status institutions might deliver greater quality gains in private-public collaboration than those in lower status institutions. In order to control for this effect, I need to record the status scores of institutions to which academic scientists might belong. Fortunately, research publications record post-codes (codes of letters and digits used as part of the postal address to sort mail) on the paper. A single publication usually yields several postcodes. Each postcode then corresponds to an address for a collaborating institution (either biotechnology or academic). While it is possible that two institutions might share a single postcode, I assume that each institution is identified by a unique postcode. This assumption stands up to scrutiny, as the number of postcode areas generated is equal to the total number of collaborating institutions in the sample. This generates a list of 160 UK postcode addresses active in biomedical research. Further, I record a measure of the research quality associated with each postcode address over time. This captures the intuition that institutions have different "status" that could be mirrored in the research quality of the papers associated with it. Following the definition of CIT (citations) outlined earlier, I estimate the mean number of citations of all the papers published in a year that include the relevant postcode in the address field. This information is then used to construct LEADACAD_GEOGCIT and LEADAUT_GEOGCIT to capture the status of the lead *academic* author and the lead author[25] respectively as measured by their institutional affiliation (Note: the lead academic author need not necessarily be the lead author).

[25] The lead author is the first author listed on a research paper who by convention is the seniormost or "lead" scientist on the research team.

International Collaboration. There is evidence to suggest that foreign collaborations can impact positively impact on research quality (Zucker and Darby, 1995). To control for this, the presence of foreign collaborators present on a publication is estimated by counting the number of foreign addresses (FOR).

Author Characteristics. Perhaps the strongest predictor of research quality are the individual characteristics of the scientists that author a research paper. In his classic work on the sociology of scientists, Merton (1968) outlined the "Matthew Effect" where, scientists of higher prestige tend to draw resources such as research funding, grants and international collaborations towards themselves, which gives them further advantages thus strengthening their status in self reinforcing loop. Thus status can be a powerful predictor of research quality as measured by impact (citations and journal prestige). To capture the "status" of authors present on a paper, as before, I record the total number of citations received by and individual author as AUT_CIT, for all his publications. In my database, this resulted in 9,445 citations. In order to estimate the average impact of an author as measured by journal impact, I calculate the mean PIC (potential impact category) of the author (AUT_PIC). These variables are constructed by calculating the respective means for that set of papers on which a particular author is present. Particularly, I record LEADAUT_CIT as well as LEADAUT_PIC for the lead author of a paper and LEADACAD_CIT for that of the lead *academic* author. LEADACAD_CIT and LEADAUT_CIT can be the same, although not necessarily so.

4.4.1.3 Control Variables

Whether or not a particular research effort receives support from government funding bodies could influence its potential impact. The Science Citation Index (SCI) records funding sources present on a paper, and which is captured by a simple categorical (yes/no) variable for the presence of the key funding bodies (denoted by FUNDING). These are (in order of total number of papers funded): Medical Research Council (denoted by MED), Biotechnology and Biological Sciences Research Council (BBSRC), the Department of Health (HEALTH), Ministry of Agriculture, Fisheries and Food (MAFF), National Environment Research Council (NERC), British Council (BC), Engineering and Physical Sciences Research Council (EPSRC), Overseas Development Administration (ODA) and the Economic and Social Research Council (ESRC). Funding acknowledgements were also found to the Agricultural and Food Research Council and the Science and Engineering Research Council. However, both these bodies been largely taken over by the BBSRC and are recorded as such. Finally, to account for firm effects, I record the firm variable (FIRM) by whom the principal biotechnology author is employed. There are 114 firms present in the dataset. The year of publication (YEAR) is also captured to account for learning effects in doing research by biotechnology firms. It could be, for instance, that there is an overall increasing trend in research quality over time for the firms in the dataset. A full listing of the key variables and their definitions can be found in Table 4.1.

4.5. Results

4.5.1. Descriptive Statistics

Table 4.2 provides some basic descriptive information on the analytical variables of interest. On average, the papers received 11.47 citations. There were 5.6 total authors for each paper with an average of just 1.08 from industry and 4.51 from academia. The average count of biotechnology firms and academic institutions is 1.03 and 1.94 respectively. On average, the lead academic author has 23.65 citations, the lead author institution 19.28 citations, and the lead author 19.63 citations. The research is generally more applied than basic. Fifty-six percent of the total sample includes some foreign collaboration.

In Tables 4.3 through 4.5, the descriptive statistics found in Table 4.2 are split by research level, research quality and academic contribution. The data in Table 4.3 suggests that the more basic the nature of the research, the higher is the research quality. There is some evidence in the literature to support this tendency. It has been documented that in many fields of research, papers that concern fundamental or basic research tend to have greater impact than papers of a more applied nature (Lewison et al., 2003; Lewison and Devey, 1999).

In Table 4.4, it is interesting to note that there is no discernable increase in the ratio of academic collaboration as I move from applied to basic research level. This is not so counter-intuitive, however, since this dataset, by definition, has only selected papers which have at least one co-author from a biotechnology firm that presumably is interested in more applied work to begin with.

Finally, Table 4.5 shows some descriptive evidence that increased academic contribution on a research paper is correlated to higher research quality. The ratio of academic co-authors rises from 67 percent to 78 percent as one moves from lowest to highest research quality publications, meaning that more academic input is positively correlated with private-public research quality. I formally test whether this result is robust to the inclusion of control variables, firm fixed effects and different model specifications.

4.5.2. Stage 1: Controlling For Selection Bias

In this section, I control for any biases that might arise in my study by excluding research publications with pure academics. There are several ways in which I can do solve this problem. The most straightforward method would be to include all the research publications including those authored purely by academics in my dataset. However, to do so, I would have to individually sort out 355, 183 scientific papers that exist in an unprocessed format from the original database. This is a very labor and time intensive exercise that is beyond the scope of my thesis. Also, in this thesis, I am more interested in the actions of *firms* and how they might improve their research quality by collaborating with academics. Thus, I limit my dataset to include only those papers that have at least one biotech author.

To overcome any potential biases, two approaches are most commonly used: 1. to find instrumental variables, 2. to use an econometric adjustment to the baseline regression.

Of these two methods, the use of instrumental variables is usually preferable. However, good instruments are hard to find. Given my own dataset, I found it hard to justify the use of instrumental variables that do not influence outcomes in my regression equation. Thus, in this chapter, I use an econometric technique used widely in studies of this kind (the Heckman selection model) to control for any biases that might arise. Below, I discuss the motivation for the inclusion of these variables in my Heckman selection equation.

Dependent variable

The dependent variable (denoted as BIAS) is 1 in the presence of a biotechnology firm on the research paper, and 0 otherwise.

Independent variables

Status of lead author. High status authors are more likely to form partnerships across institutional divides. Scientists at biotechnology firms tend to collaborate primarily with high status authors from academia. When biotechnology firms collaborate with academia they usually do so if the incentives are high. Thus they undertake collaboration with academic scientists to obtain crucial know-how from cutting edge science, or to gain credibility (in the eyes of investors, financial markets and large pharmaceutical firms) by association with a prestigious scientist or research laboratory. Both these incentives will direct them to work with academic scientists of high prestige (Cockburn and Henderson, 1998; Stern, 1999). By the mechanism of homophily (Lazarsfeld, 1972), high status academic scientists, in turn, are more likely to collaborate with high status scientists at biotechnology firms. Also, high status academic scientists are more likely to have links with industry and form research

partnerships across the private-public divide (Murray, 2002).Thus, the status of the authors on a paper could be powerful predictor of biotechnology presence. In the specification, the status of the lead author as a key dependent variable is included and recorded as LEADAUT_CIT (the average number of citations received by a paper authored by the lead author), LEADAUT_PIC (the average journal impact achieved by the author) and LEADGEOG_CIT (the average citation score of the papers produced by the institution to which the lead author belongs).

Research Level. It is reasonable to assume that papers of a more applied nature might be more likely to have biotechnology presence than more basic papers, given the applied nature of the science pursued at biotechnology firms. The research level (RL) is included in the selection equation.

Government funding. The UK government funds public science via a grant system through a variety of funding agencies. These agencies typically encourage private-public collaboration as part of the UK government's effort to strengthen the knowledge based economy. I found evidence of this intent in the literature (Lewison et al., 2003) that suggests although funding is not strictly contingent on collaboration, it is a strongly encouraged by informal practice as well as formal declaration. Publications produced by funded research programs are required to list sources of funding that support their research. I record this variable as FUNDING (0/1).

The empirical specification (See Table 4.6 for a listing of variable definitions and some basic descriptives) of the selection Equation 4.4 is:

$$Prob\ (T_i = 1) = \alpha + \beta_1 LEADAUT\_CIT_j + \beta_2 LEADAUT\_PIC_j$$
$$+\beta_4 LEADAUT\_GEOGCIT_j + \beta_4 RL_i + \beta_5 DFUNDING_i$$

(Equation 4.8)

The subscript i denotes paper characteristics, j denotes author characteristics and k denotes institutional characteristics.

The regression was run on the entire set of scientific publications (i.e. including those including no biotechnology involvement) in the database (355 183 papers). Table 4.7 shows results of the PROBIT estimation of Equation 4.8. The status of the lead author and the presence of government body influence biotechnology participation in research. This result is consistent with the existing literature and the empirical specification.

4.5.3. Stage 2: The Effect Of Co-authorship On Research Quality

In the second stage, I exclude research publications that have no biotechnology presence from the dataset leading to a sample of 2915 papers.

4.5.3.1 The Baseline Estimate

In order to measure the impact of private-public co-authorship on research quality, I begin with a simple baseline specification:

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i$$
$$+ \beta_5 LEADACAD\_CIT_j + \beta_6 LEADACAD\_GEOGCIT_k$$
$$+ \beta_5 DFUNDING_i$$

(Model 4.1)

This specification captures the effect of academic input (ACADRATIO_AUT) on the quality of the research paper (LOGCIT), i published in year t, controlling for total number of authors (AUT), the presence of foreign collaborators (DFOR), the level of research on the basic-applied spectrum (RL), the status of the lead academic author as measured by his individual citation score as well as his institutional affiliation and the presence of funding bodies (DFUNDING). Table 4.8, column (1) presents OLS estimates for Model 4.1. In column (1) I find statistically significant evidence that higher levels of academic involvement result in higher research quality, controlling for the total number of authors, research level foreign collaboration, author status and funding presence.

4.5.3.2. Are There Diminishing Returns To Total Number of Authors?

A question not answered by Model 4.1 is at what point (if at all) do diminishing returns set in with increased total authorship (i.e., do too many researchers spoil the broth?). This is estimated in Model 4.2 below. Specifically, to test for diminishing returns of total number of authors, I add a squared term for number of authors, AUTSQ, and specify the following:

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_4 AUTSQ_i + \beta_5 RL_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i$$

(Model 4.2)

After removing 29 outliers (.009% of sample) in the total number of authors (some papers had upwards of 60 authors), there is a significant curvilinear effect associated with adding co-authors and research quality. The threshold number of total authors, after which additional authors result in decreasing quality, can be calculated from the co-efficients for AUT and AUTSQ in Table 4.8, column (2). This value turns out to be 11. Thus, including more than eleven authors actually *decreases* quality of the resulting publication. It is important to note that although the co-efficient for ACADRATIO_INST falls by a third, it remains significantly positive.

4.5.3.3. Time Trends

There is some evidence to suggest (Cockburn and Henderson, 1994) that, with time, firms get better at doing research. Controlling for these learning effects, I modify Model (4.1) to include year dummies (YEARDUM):

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i + \beta_9 DYEAR_i$$

(Model 4.3)

From the Table 4.8, column (3), I can see that the model specification, as judged by the R-squared, improves on this addition. The individual year dummies, however, reveal no overall rising time trend in research level nor do they dampen my ACADRATIO_INST coefficient.

4.5.3.4. Does Controlling For Firm Effects Improve My Results?

Finally, the firm dummies FIRMDUM are included in the model specification (Model 4.1):

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i + \beta_9 DFIRM_i$$

(Model 4.4)

This is an important addition as not all private involvement is of the same quality. By not including some control for the firm from which the private researcher is drawn from, I may be excluding valuable information which would improve my estimates of the effect of academic co-authorship. Indeed, given that I assumed that co-authorship is not random, it is likely that the best firms attract better and more numerous academic collaboration. From Table 4.8, column (4), I can see that the model specification is further improved, with an R-squared that is double my baseline results. Moreover, as expected, the effect of academic co-authorship increases in size and significance with the inclusion of firm dummies. I also estimated the same model with standard errors clustered around individual biotechnology firms, which only marginally refined the results.

4.5.3.5. Are There Diminishing Returns To Academic Participation?

In my baseline regression, I allow for team composition with at least one biotech author excluding those papers that are authored purely by academics. Thus, I do not test for a combined academic-industry team to have better citation performance than a purely academic one. While this is an important question, the unit of my research is the biotech firm and the action these firms might take. Thus, I am more concerned with how biotech

firms might improve research quality by collaborating with academics. I control for the fact that my database excludes papers by pure academics by using a Heckman selection model to control for potential biases. I find that adding academic authors improves research quality of papers authored by industry.

One question raised by this result is whether there is a limit to this effect? Does adding more academic authors continuously improve research quality or is there a ceiling on the gains to be had from academic involvement? In Table 4.8a, I split my academic authorship variable (AUT_ACAD) into 3 categories with increasing levels of academic participation[26]. I find that the adding a single academic author results in a significant jump in research quality. However, after this initial jump, the gains from adding more authors increases at a decreasing rate. This effect is visually presented in Fig. 4.8a.

In Table 4.8b, I pin down this decreasing gains effect more precisely by adding a squared term (ACAD_RATIOSQ) to my baseline regression (Model 4.1).

Thus, I run the regression:

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_4 ACADRATIO\_AUTSQ_i + \beta_5 RL_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i$$

(Model 4.5)

From Table 4.8b, Col (2), by using the coefficients for ACAD_RATIO and ACAD_RATIOSQ, we can calculate the threshold number of academic authors after

[26] 1. Pure Industry: no academic authors, 2. Low: 0 < AUT_ACAD <= 1, 3. Medium: 1 < AUT_ACAD >= 4, High: AUT_ACAD > 4.

which adding more academic authors result in decreasing quality. I calculate this value as 4. Thus, we find adding academic authors do *not* result in better research once this threshold has been crossed. In the previous section (4.5.3.2) we find that there are diminishing returns to total authors as well once this number crosses 11. Thus, we find, in fact, that once the share of academic participation crosses a third (36.35% to be precise), the gains drop off.

To investigate the effects of team composition in more detail, I re-run my baseline regression by defining ACAD_RATIO as a categorical variable. To do so, I break my ACAD_RATIO variable (which ranges from 0 to 1) which is the share of academics to total number of authors into three groups with the cutoff ratios as < 0.33, 0.33 – 0.67, > 0.67 and include dummy variables for each (the omitted reference category is a pure industry team with no academics). The results are shown in Table 4.8c. I find that even having low academic participation results in significant gains in research quality. These gains rise slowly with medium level of academic participation. Interestingly, high levels of academic participation seem to lower the effect of academic participation on research quality which is consistent with the result I obtained previously by including a squared term for academic participation in the baseline regression.

4.5.3.6. Impact of Academic Input on Private-public Research Quality Split by Research Level

The baseline is estimated (Model 4.1) after splitting the data set into three research level (RL) categories: Basic, Medium and Applied. Results are shown in Table 4.9. The

results in the Applied category closely mirror those for the entire sample. This is not very counter-intuitive as "applied" papers dominate the sample. However, when I run the estimation on the "basic" papers, I find an even stronger advantage in collaborating with academia. This result is borne out by the intuitive idea that academic scientists are more likely to possess expertise and skills in basic rather than applied fields and are therefore most likely to boost private-public research quality on topics of basic research.

4.5.4. Alternative Definitions of Key Variables

4.5.4.1. Different Measures of Research Quality

As discussed earlier, measuring research quality is highly contentious. A broad definition of research quality would include effects on wealth creation and health benefits. However, for the purposes of this paper, I defined research quality as the impact of research on other researchers. Even this measure, can be specified in alternative ways. In the baseline estimation (Model 4.1), I measure research quality as a logarithm transform (Equation 4.7.) of the number of citations received by the publication over a five year period from the year of publication (LOGCIT). Alternatively, I could have used a simple five year citation count to measure quality. I do so in estimating the equation:

$$CIT_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i$$

(Model 4.6)

Table 4.10, Col (2) shows that while the basic interpretation remains unaltered, specifying the model with the logarithm transform nearly doubles the R-square of the model, justifying its use as the baseline definition.

As noted earlier, the status of the journal in which a paper is published is an alternative measure of research quality. Lewison and Devey (1999) suggest that journal prestige captures the quality of a research paper as judged by a journal editor and a few specialized reviewers as opposed to the general body of researchers. Lewison et al. (2003) argue that this makes it a more stable measure of research quality as it might reflect more consistent standards of judging quality, rather than reflect the scientific trends and fads that could influence citation scores. To check whether this impacts my earlier results, the baseline regression is estimated with the dependent variable as the potential impact category (PIC) of the research paper. Thus,

$$PIC_{i,t} = \alpha + \beta_1 ACADRATIO\_AUT_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADGEOGCIT_k + \beta_8 DFUNDING_i$$

(Model 4.7)

In Table 4.10, Col (3), I find that while the R-squared is roughly equivalent to the baseline estimate (using raw citation scores), it is not as powerful as using the logarithm transformation of the citation score (Model 4.1). This might be true as noted earlier that logarithm measures of citation scores might more closely reflect research quality as judged by the scientific community. The results show that quality gains by increased academic participation are more keenly reflected in this measure. Also, it is interesting

to note that using the potential impact category (PIC) of the journal is roughly equivalent to using raw citation scores (Model 4.8). The two measures are seen to be roughly equivalent. This could mean that papers in high status journals get cited more often *because* they appear in them or alternatively, because high status journals are good filters of high quality research.

4.5.4.2. Measuring Co-authorship At The Institutional Level

In the baseline estimate, the principal measure of co-authorship was aggregated at the level of the institution. I found that, other things being equal, research quality increases with increasing shares of academic participation as measured by the ratio of academic authors to the total number of authors. However, one could specify co-authorship at the level of the institution and record shares of academic participation as the ratio of academic institutions present on a publication to the total number of institutions (industry + academy). I re-specify the baseline regression with this altered definition of co-authorship as:

$$LOGCIT_{i,t} = \alpha + \beta_1 ACADRATIO\_INST_i + \beta_2 AUT_i + \beta_3 RL_i + \beta_4 DFOR_i + \beta_6 LEADACAD\_CIT_j + \beta_7 LEADACAD\_GEOGCIT_k + \beta_8 DFUNDING_i$$

(Model 4.8)

Table 4.10, Col (4) shows that while the results are robust to this measure of co-authorship, the R-square is higher when co-authorship is measured at the institutional level. This is because there could be more than a single author from a particular

institution and adding authors from the same institution may not significantly influence research quality.

4.5.5. Multicollinearity, Endogeneity And Remaining Biases

The main biases that need to be addressed by the use of my empirical design are: any selection bias that might exist by excluding research papers purely authored by academics, multicollinearity[27] between my exogenous variables and endogeneity between my dependent and independent variables.

I address the selection bias in an earlier section (Section 4.5.2.) and control for this bias by using a Heckman selection model.

I test for multicollinearity, and report the cross-correlation of exogenous variables in Table 4.11. The results suggest that my estimates do not suffer much from multicollinearity. With the exception of the correlation between funding support and research level, no correlation appears to be high enough to render the obtained z-values unreliable[28]. This strengthens the claim for accuracy of the reported standard errors in my analysis.

[27] Leung and Yu (1996, p.201) explain that "models with few exclusion restrictions, a high degree of censoring and a low variability among the regressors, or a larger error variance in the choice equation, can all contribute to near collinearity between the regressors and the inverse Mills ratio, rendering the two-step estimator ineffective".

[28] Note that multicollinearity has no impact on the coefficients. Even if two or more variables were collinear, the estimated coefficients would remain unbiased. Therefore, our inferences are not hampered by collinearity problems (Greene, 1997).

Secondly, there could be a problem of endogeneity between the measured dependent variable, research quality, and the central independent variable, private-public collaboration. It could be possible that the level of research quality impacts the level of collaboration, instead of the other way around. For example, projects of higher research potential might attract funding from a variety of sources, both private and public, and hence reflect a greater degree of collaboration. However, disentangling this causality is difficult. I have tried to do so, by controlling for individual author characteristics, such as their individual status and the institution to which they belong, the presence of international collaborations and support from governmental funding bodies.

To achieve a true ex-ante separation of variables, one possibility would be to use more sophisticated matching methods or the use of an experimental design that would control for such two-way causality. To do so, I would either have to find a naturally occurring experiment where I am certain of the causality or to construct a more elaborate dataset where I may "match" industry authors seeking academic collaboration against a set of otherwise similar industry authors that do not. Though this empirical framework lies outside the scope of this thesis, I hope to incorporate these ideas in future work with the dataset, investigate the matching methods literature more closely in an effort to more precisely identify my main effects.

4.6. Conclusion and Implications for R&D Process and Policy

In this chapter, I have used unique data on UK biotechnology research output from 1988 to 2001 to investigate the impact of private-public collaboration on research quality.

Specifically, I tested whether research quality of private-public research output is improved by increased academic input. I find clear evidence for this in the data that is robust to the inclusion of control variables, firm fixed effects and different model specifications. To some extent, my results mirror studies in the US in which star scientists have been found to impact firm research productivity (Zucker and Darby, 1995).

This study also suggests that science and technology policies geared to "corporatize" academic research might be misplaced. The results indicate that researchers in academia and industry might possess very distinct skills developed over a path-dependent process. Collaborating across this divide then provides real gains for industry which might be lost if the academic "market" were to lose its unique identity. Recent research on biotechnology firms by Gittelman and Kogut (2003) and Murray (2003) support this contention that collaboration between academics and industry is crucial to the success of the biotechnology industry. By uniquely measuring the quality gains to be had from collaboration, the research deepens our understanding of this process. As noted earlier, this research provides evidence for the tacit nature of knowledge in the biotechnology sector and further suggests that the public and private sectors have heterogeneous knowledge bases which could explain the quality increase as a result of collaboration.

My results offer a number of implications for R&D process and policy for firms and institutions. First, for firms, this study suggests that firms in the biotechnology sector produce better quality research, as measured by journal impact and citations, by collaborating with academic researchers.

Even discounting "club effects", these quality gains can provide clear benefits to biotechnology firms. However, irrespective of whether club-effects, as opposed to "true" improvements in research quality, are the cause of getting published in prestigious journals, getting published. in and of itself. might bring private firms other benefits. For example, getting published in top journals might allow young biotechnology firms to more easily raise money, attract better human capital and form alliances with large pharmaceutical firms. Collaboration with academia is one avenue by which biotechnology firms can achieve this goal. The greatest impact of collaboration is felt in "basic" research although the results hold across the "basic-applied" spectrum. Further, the research indicates that these quality gains are accentuated by collaborating with more than one laboratory and across disciplines, particularly seeking out alliances with academics with established research reputations. Firms may wish to consolidate their academic relationships to a smaller number of high performing university researchers and build long term relationships.

The theoretical literature reviewed in Section 4.2 suggests that a variety of competing mechanisms might be at work to explain these quality gains to be had from private-public collaboration. However, it must be noted that instead of "true" quality gains that result from collaboration, an alternative mechanism could be that academic co-authorship may act as a form of "club membership" required for getting published in more prestigious journals. Further, if club effects do influence research quality, it may be that academic authors at prestigious universities and research institutes can more easily publish in prestigious journals. In fact, in the empirical estimation, this study

finds evidence for this as a strong predictor of research quality is the status of the lead author..

On the policy side, the study suggests a number of implications for government efforts at spurring biotechnology research and industry. Publication is an important means of increasing the rates of technical innovation (Sorenson and Fleming, 2004), a key focus of the UK and other national governments' policies.

As UK government spend on R&D lags that of other countries with strong biotechnology competence, namely the US, it is important that these funds be invested wisely. The Research Assessment Exercise (RAE) is a census designed to assess the research quality of universities and colleges across the UK. The RAE measures the research of individuals and aggregates these to the institutional level. The results are used by higher education funding bodies to distribute public funds for research, including Quality Related (QR), the most important source of funding for universities due to its size and block nature. Since 1997, the Research Councils have invested over £7 billion in 130 higher education institutions in the UK (HM Treasury, 2004).

My results suggest that mechanisms may be useful as predictors of higher quality research. To encourage higher quality outputs, it would seem that public funding to lower-performing individual academics and institutions might be better directed to higher performing academics. Just as market forces have driven consolidation among biotechnology firms over the past few years, it may be that the UK government can only afford to fund a smaller number of truly high performing research institutions. The

additional resources provided to these laboratories could fund research students and post-doctoral fellows who would receive top level training. From the university perspective, those institutions seeking high rankings in the next RAE cycle (2008) may wish to consider setting aside resources to attract and retain highly productive research scientists.

While this study dealt with research from biotechnology firms based in the UK, my results offer policy implications for other countries keen to develop biotechnology competence at academic and firm levels. In particular, national governments may wish to encourage their ambitious biotechnology firms to seek collaborations with the most high status academics, at home or abroad.

My findings suggest several directions for future research. First, future studies could examine how research quality impacts firm performance and through what mechanisms. It could be interesting to explore how research quality of these collaborations is linked to government and industry funding. Further work could also examine this trend in other disciplinary fields characterized by academic-firm research collaboration and co-authorship, such as business management.

Finally, I wish to acknowledge the limitations of this study. I did not take into self citations. Also, co-authorship can entail any number of different forms of collaboration which I do not take into account.

Table 4.1
Variables for Analysis of Collaboration Effect on Research Quality

Variable Name	Definitions
Research Quality	
Citation score (CIT)	Total number of citation count to individual papers over a five year period from the year of publication.
Potential Impact Category (PIC)	Potential impact of a research publication. Ranked variable from lowest to highest quality (1 to 4).
Log Citation Scores (LOGCIT)	Uses $\log CIT = 1 + 2\log_{10}(CIT + 1)$ to more accurately gauge relative citation scores as perceived by researchers.
Collaboration	
Total authors (AUT)	Number of authors listed on a paper.
Academic authors (AUT_ACAD)	Number of authors from academia listed on a paper.
Biotechnology authors(AUT_BIOTECH)	Number of industry authors listed on a paper.
Biotechnology firms (BIOTECH)	Number of biotechnology firms present on a paper.
Academic institutions (ACAD)	Number of academic institutions present on a paper.
Academic author ratio (ACADRATIO_AUT)	Ratio of academic authors to total number of authors per paper.
Academic institution ratio (ACADRATIO_INST)	Ratio of academic institutions to total number of institutions per paper.
Square of total authors (AUTSQ)	The square of total number of authors.
Research Level	
Research Level (RL)	Research level on basic/applied scale of a research paper. Ranked from 1 (most applied) to 4 (most basic).
Institutional Affiliation	
Status of Lead Academic Author: Institutional Affiliation (LEADACAD_GEOGCIT)	Mean citation score of institution to which the lead academic author belongs.
International Collaboration	
Foreign addresses (FOR)	Number of foreign addresses listed on a paper.
Foreign dummy (DFOR)	Presence of foreign addresses listed on a paper. 1 if foreign presence and 0 otherwise.
Author characteristics	
Author citation score (AUT_CIT)	Mean citation score of all publications by a particular author.
Status Of Lead Academic Author: Citation Score (LEADACAD_CIT)	Mean citation score of lead academic author.
Other	
Funding body presence (FUNDING)	Presence of UK government funding body listed on a paper. 1 if funding present and 0 otherwise.
Year (DYEAR)	Year of publication dummy.
Firm (DFIRM)	Firm dummy.

Table 4.2
Descriptive Statistics

Variable name	Mean	SD	Min	Max	Obs
Research Quality					
Citation Score	11.473	3.02	0	432	2915
Log Citation Score	3.192	0.321	1	6.27	2915
Potential Impact Category	2.467	1.032	1	4	2915
Collaboration[a]					
Total authors	5.600	1.798	1	11	2915
Industry authors	1.086	0.302	1	4	2915
Academic authors	4.514	1.795	0	7	2915
Biotechnology firms	1.03	0.008	1	3	2915
Academic inst.	1.942	0.118	0	6	2915
Academic author ratio	0.722	0.209	0	0.98	2915
Academic institution ratio	0.394	0.105	0	5	2915
Square of total authors	45.791	13.54	1	7	2915
Research level					
Research level	3.067	0.902	0	3.98	2915
Author characteristics					
Status Of Lead Academic Author: Citation Score	23.653	2.494	5	432	2915
Status Of Lead Academic Author: Institutional Affiliation	19.285	1.863	17	76	2915
Status Of Lead Author: Citation Score	19.638	1.903	7	432	2915
Other					
Foreign addresses	0.566	1.306	0	21	2915
Year	-	-	1988	2001	2915

[a] on an individual paper

Table 4.3
Research Level by Category and Research Quality

Research Level	Research Quality		Obs
	PIC	Citations	
Applied	2.02	11.02	2062
Medium	2.24	11.38	659
Basic	2.58	12.07	194

Table 4.4
Research Level by Category and Academic Contribution

Research Level	Academic Contribution		Obs
	Institution Ratio [a] (mean)	Author Ratio [b] (mean)	
Applied	0.31	0.70	2 062
Medium	0.38	0.73	659
Basic	0.37	0.72	194

[a] Ratio of academic institutions to total number of institutions (biotechnology firms and academic institutions) present on a paper.
[b] Ratio of academic authors to total number of authors (biotechnology and academic) present on a paper.

Table 4.5
Research Quality and Academic Contribution

Research Quality PIC 1-4 (Number of citations)	Academic Contribution		Obs
	Institution Ratio [a] (mean)	Author Ratio [b] (mean)	
1. Less than 6	0.29	0.67	631
2. 6-11	0.35	0.71	850
3. 12-19	0.37	0.73	874
4. 20	0.41	0.78	560

[a] Ratio of academic institutions to total number of institutions (biotechnology firms and academic institutions) present on a paper.
[b] Ratio of academic authors to total number of authors (biotechnology and academic) present on a paper

Table 4.6
Variables For Controlling Selection Bias

Variable Name	Definitions	Mean	STD	Obs.
Dependent Prediction of biotechnology participation (BIAS)	Selection variable: 1 if a paper is likely to be included in the sample, 0 otherwise.	-	-	-
Independent Status of lead author a)Citation Score (LEADAUT_CIT)	Mean citation score of all publications of the lead author on the paper. Derived from AUT_CIT.	17.657	2.347	355, 183
b)Potential Impact Category (LEADAUT_PIC)	Mean PIC (potential impact category of journal) of lead author.	2.962	0.287	355, 183
Research Level (RL)	Publication Research Level	3.621	0.915	355, 183
Government funding (DFUNDING)	Presence of UK government funding body listed on a paper. 1 if funding; 0 none.	-	-	-

Table 4.7
Probit Participation Estimation

Dependent Variable: BIAS (0/1)	Coefficient	Standard Error
Lead Author Status: Citation Score (LEADAUT_CIT)	0.302**	0.027
Lead Author Status: Institutional Affiliation (LEADAUT_GEOGCIT)	0.513***	0.049
Lead Author PIC (LEADAUT_PIC)	0.335**	0.051
Research Level (RL)	0.285*	0.016
Funding Body Presence (DFUNDING)	0.113**	0.011
$\log(\lambda)$	0.584**	0.035
- Log Likelihood	2.423	
Wald χ^2	89.37	
F – Statistic	0.606	
Obs	3,55,183	

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level

Table 4.8
The Effect of Academic Input on Private-Public Research Quality

Dependent Variable: Quality of Research (LOGCIT)	Heckman OLS Baseline Model 1	Heckman OLS Model 2	Heckman OLS Model 3	Heckman OLS Model 4
Academic Ratio	0.745*** (0.117)	0.715*** (0.129)	0.674*** (0.229)	0.795*** (0.283)
Total Authors	0.043*** (0.034)	0.082*** (0.061)	0.058*** (0.011)	0.037*** (0.005)
Total Authors Squared	--	-0.004*** (0.001)	--	--
Foreign Address Dummy	0.138** (0.072)	0.145*** (0.043)	0.156*** (0.015)	0.182*** (0.054)
Research Level	0.633*** (0.159)	0.516*** (0.169)	0.652*** (0.134)	0.493*** (0.115)
Status Of Lead Academic Author: Citation Score	0.678*** (0.244)	0.613*** (0.184)	0.714*** (0.209)	0.781*** (0.214)
Status of Lead Academic Author: Institutional Affiliation	0.618*** (0.213)	0.527*** (0.104)	0.509*** (0.121)	0.622*** (0.227)
Funding Body Dummy	0.314** (0.115)	0.285** (0.156)	0.313** (0.102)	0.267** (0.054)
Mills Ratio	0.795**	0.862**	0.814**	0.833**
Year Dummies	No	No	Yes	Yes
Firm Dummies	No	No	No	Yes
R-squared	0.27	0.27	0.28	0.42
Observations	2,915	2,915	2,915	2,915

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Standard errors in parentheses.
1. OLS regression with research quality as dependent variable and academic author ratio, total number of authors, number of foreign collaborators and research level as independent variables.
2. There is also a specification as in Model 1 but with standard errors clustered around individual biotechnology firms. The results are similar and available upon request.
3. The R-Squares is presented from the standard OLS regression after correction in the two-step Heckman estimation.

Table 4.8a
Returns To Research Quality On Adding Academic Authors

Level of Academic Participation (No. of academic authors)	**Research Quality** (Citation Score)
[pure industry]	3.442
Low	11.172
Medium	12.243
High	12.342

Fig. 4.8a
Returns To Research Quality On Adding Academic Authors

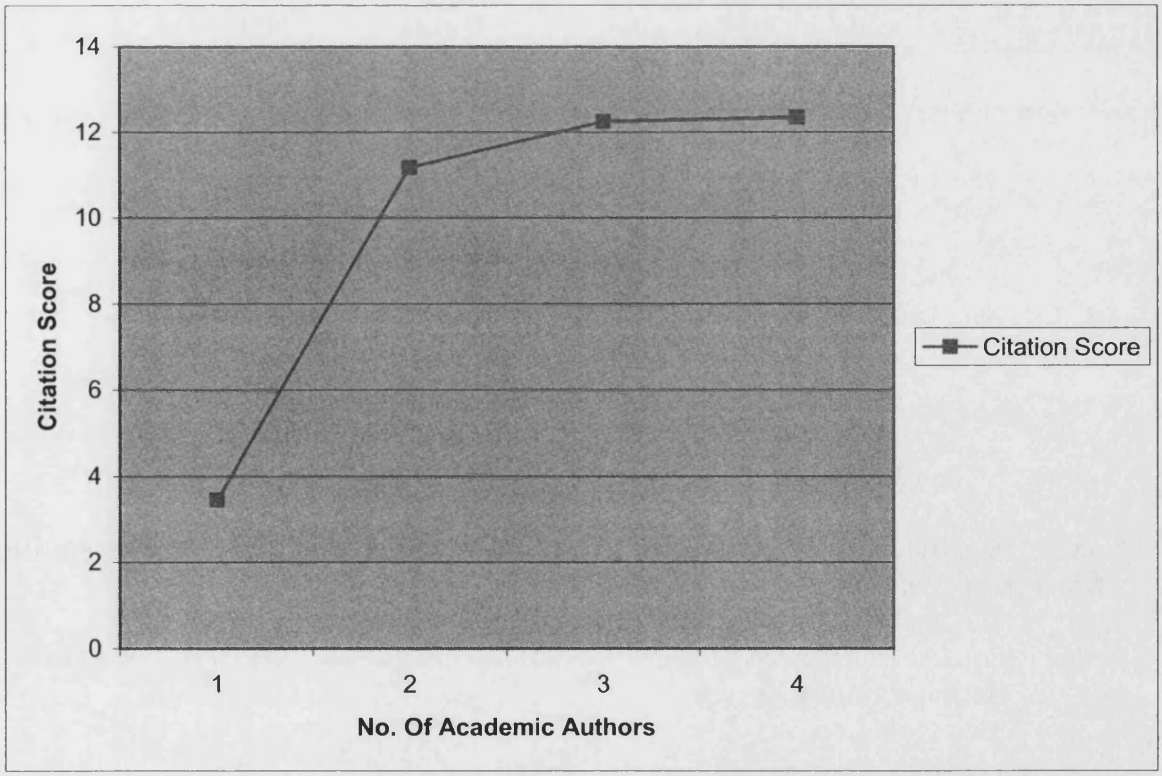

Table 4.8b
The Effect of *Increasing* Academic Input on Private-Public Research Quality

Dependent Variable: Quality of Research (LOGCIT)	Heckman OLS Baseline Model 1	Heckman OLS Model 5
Academic Ratio	0.745***	0.812***
	(0.117)	(0.213)
Academic Ratio squared	--	0.103
		(0.016)
Total Authors	0.043***	0.033***
	(0.034)	(0.002)
Foreign Address Dummy	0.138**	0.179***
	(0.072)	(0.044)
Research Level	0.633***	0.487***
	(0.159)	(0.128)
Status Of Lead Academic Author: Citation Score	0.678***	0.722***
	(0.244)	(0.232)
Status of Lead Academic Author: Institutional Affiliation	0.618***	0.634***
	(0.213)	(0.217)
Funding Body Dummy	0.314**	0.258**
	(0.115)	(0.044)
Mills Ratio	0.795**	0.811**
Year Dummies	No	Yes
Firm Dummies	No	Yes
R-squared	0.27	0.44
Observations	2,915	2,915

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Standard errors in parentheses.
1. Estimation method: Ordinary Least Squares.
2. The R-Squares is presented from the standard OLS regression after correction in the two-step Heckman estimation.

Table 4.8c
The Effect of *Increasing* Academic Input on Private-Public Research Quality

Dependent Variable: Quality of Research (LOGCIT)	Heckman OLS Model 1	Heckman OLS Model 1
Academic Ratio	0.745***	--
	(0.117)	
Dummies for Academic Ratio		
	--	--
[pure industry]		
	--	
Low		0.773***
		(0.137)
	--	
Medium		0.811***
		(0.202)
	--	
High		0.117***
		(.022)
Total Authors	0.043***	0.037***
	(0.034)	(0.005)
Research Level	0.633***	0.493***
	(0.159)	(0.115)
Status Of Lead Academic Author: Citation Score	0.678***	0.781***
	(0.244)	(0.214)
Status of Lead Academic Author: Institutional Affiliation	0.618***	0.622***
	(0.213)	(0.227)
R-squared	0.27	0.42
Observations	2,915	2,915

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Standard errors in parentheses.
1. Estimation method: Ordinary Least Squares.
2. [] refers to omitted reference category.
3. The R-Squares is presented from the standard OLS regression after correction in the two-step Heckman estimation.

Table 4.9
Baseline Estimation Across Research Level

Dependent Variable: Quality Of Research (LOGCIT)	Baseline Estimation (Model 1) Applied (RL)	Medium (RL)	Basic (RL)
Academic Ratio	0.642***	0.682***	0.692***
	(0.197)	(0.214)	(0.223)
Total Authors	0.046***	0.049***	0.069***
	(0.013)	(0.017)	(0.028)
Total Authors Squared	--	--	--
Foreign Address Dummy	0.137***	0.174***	0.147***
	(0.028)	(0.031)	(0.029)
Research Level	0.602***	0.589***	0.493***
	(0.217)	(0.161)	(0.143)
Status Of Lead Academic Author: Citation Score	0.665***	0.581***	0.812***
	(0.201)	(0.124)	(0.267)
Status of Lead Academic Author: Institutional Affiliation	0.613***	0.597***	0.634***
	(0.192)	(0.102)	(0.203)
Funding Body Dummy	0.303***	0.292***	0.213***
	(0.067)	(0.056)	(0.023)
Mills Ratio	0.762***	0.722***	0.745***
R-squared	0.27	0.20	0.38
Observations	2,062	659	194

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Standard errors in parentheses.
1. The R-Squares is presented from the standard OLS regression after correction in the two-step Heckman estimation.

Table 4.10
Tests of Robustness: Estimation Using Alternate Measures of Quality & Collaboration

Dependent Variable: Quality of Research	Standard Baseline OLS Model 1	Dependent as CIT OLS Model 6	Dependent as PIC OLS Model 7	Individual Collaboration OLS Model 8
Academic Ratio	0.745*** (0.117)	0.693*** (0.213)	0.657*** (0.187)	0.639*** (0.313)
Total Authors	0.043*** (0.034)	0.064*** (0.012)	0.102*** (0.078)	0.653*** (0.212)
Total Authors Squared	--	--	--	--
Foreign Address Dummy	0.138** (0.072)	0.197** (0.061)	0.148*** (0.045)	0.141*** (0.038)
Research Level	0.633*** (0.159)	0.643*** (0.168)	0.524*** (0.179)	0.492*** (0.112)
Status Of Lead Academic Author: Citation Score	0.678*** (0.244)	0.432*** (0.124)	0.625*** (0.202)	0.417*** (0.108)
Status of Lead Academic Author: Institutional Affiliation	0.618*** (0.213)	0.538*** (0.133)	0.732*** (0.203)	0.517*** (0.112)
Funding Body Dummy	0.314** (0.115)	0.234** (0.085)	0.287** (0.151)	0.301** (0.067)
Mills Ratio	0.795**	0.662**	0.811**	0.634**
R-squared[b]	0.27	0.14	0.13	0.24
Observations	2,915	2,915	2,915	2,915

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Standard errors in parentheses.
1. OLS regression with research quality as dependent variable and academic author ratio, total number of authors, number of foreign collaborators and research level as independent variables.
2. The R-Squares is presented from the standard OLS regression after correction in the two-step Heckman estimation.

Table 4.11
Cross Correlation of Key Exogenous Variables In Stage 2

	Academic Ratio	Research Level	Foreign Address Dummy	Funding Body Dummy	Status Of Lead Academic Author: Citation Score	Status of Lead Academic Author: Institutional Affiliation
Academic Ratio	1.000***					
Research Level	0.272*	1.000***				
Foreign Address Dummy	0.132*	0.348*	1.000***			
Funding Body Dummy	0.219*	0.555*	-0.049*	1.000***		
Status Of Lead Academic Author: Citation Score	0.277*	0.203*	0.242*	0.432*	1.000***	
Status of Lead Academic Author: Institutional Affiliation	0.331*	0.187*	0.311**	0.407*	0.512**	1.000***

Note: ***significant at 1-percent level of significance; ** significant at 5-percent level of significance; * significant 10-percent level of significance.

Note: A shorter version of this chapter appears in the Academy of Management 2005 Best Paper Proceedings.

Chapter 5

Why Do Biotechnology Firms Make Private Knowledge Public?

5.1. Introduction

An emerging body of research focuses on the importance of partnerships between public organizations and private firms (here after private-public) in the growth of high technology industries (Pisano, 1997; Cockburn and Henderson, 1998; Gittelman and Kogut, 2004; Murray, 2004). In this literature, the traditional distinction between "upstream" basic science produced in the public sector and "downstream" technological development undertaken by private firms is blurred. High technology firms engage in basic scientific research, the traditional preserve of the public sector. These firms sometimes specialize purely in scientific research, licensing their intellectual property rather than directly engaging in commercial development. This is particularly evident in the biotechnology industry (Financial Times, 2005).

Private firms organize their research in ways that mimic the practices found in universities and publicly funded organizations (Cockburn and Henderson, 1998; Dasgupta and David, 1994; Gittelman and Kogut, 2003). Further blurring the distinction between the public and private spheres is the fact that many high technology firms mimic open-science norms by regularly publishing the results of their research in peer-reviewed scientific journals (Hicks, 2002). In doing so, these firms engage in seemingly irrational behavior. This is because the classic economic models of scientific innovation

presuppose that firms will reveal as little information about their scientific activities as possible, especially while engaged in a research and development (R&D) race with their competition (Arrow, 1964; Okuno-Fujiwara et al., 1990; Bhattacharya et al., 1992; Dasgupta, 2000). There is, thus, a decided lack of empirical literature exploring the open-science behavior of firms. We have little understanding of why firms undertake pure science research and then proceed to publish the results from basic research activity. We are also unsure of what benefits, if any, do firms derive from engaging in upstream basic science and making their efforts public?

I address these two questions by first proposing a model of open-science innovation that includes three separate measures of innovative activity: R&D expenditure, patent counts and published scientific publications. I argue that while R&D expenditures signal the commitment of a firm's resources towards innovation, and patents record the completion of R&D activity, the stock of scientific papers, which up to now have never been used as a predictor of financial success, play a crucial role in that they signal the quality of a firm's R&D to a public audience.

Unlike large R&D led conglomerates, biotechnology firms, at least in their early years, are fairly small and rely on speculative venture capital, collaborative and contractual R&D, and the support of capital markets to finance their operations. By publishing scientific papers, biotechnology firms are credited for the quality of their R&D efforts.

The question of how firms benefit from open science is also an important one. In this chapter, I undertake a study of the financial performance of U.K. based biotechnology firms during the years 1988-2000 and focus attention on the stock of peer-reviewed published scientific papers produced by firms. My results indicate that publishing scientific papers brings real financial gains to biotechnology firms and that, on average, publishing fourteen scientific papers in academic journals has approximately the same impact on a firm's market value as obtaining a single patent. In addition, collaboration can result in better research, which in turn leads to better products. In other words, R&D can indeed have causal effects on performance.

I use financial market valuation as a proxy for firm performance for several reasons. This measure is particularly germane in the case of entrepreneurial biotechnology firms, given that they that often lack profits or sales in their early years. Furthermore, equity financing is of particular strategic importance to biotechnology firms (Wall Street Journal, 2005) and market valuation is thus highly relevant in its own right (Financial Times, 2005).

The chapter is organized as follows: Section 5.2 surveys the previous literature on innovation. Section 5.3 builds my theoretical model and Section 5.4 outlines the empirical model used in the estimation. Section 5.5 describes the data set and provides a description of the variables used in my empirical analysis. Section 5.6 presents the results. Section 5.7 concludes by presenting the key implications of this analysis with suggestions for future research.

5.2. Literature Review

5.2.1. Innovation and Measurement

Scholars have noted the role of science and innovation in generating technological change and firm and national wealth (e.g. see Schumpeter, 1942), but have been challenged by the difficulty of *measuring* knowledge flows (Krugman, 1991) and their economic impact (Kuznets, 1962; Cohen & Levin, 1989).

The starting point in measuring innovation is the knowledge production function as formalized by Zvi Griliches (1979) in which firms exist exogenously and then engage in the pursuit of new knowledge as an input into innovative activity. In its classical form, Griliches (1979) model sought to predict the performance of knowledge intensive firms (most often, its market valuation) as a function of various measures of its innovative capability. The empirical estimation of the knowledge function however was problematic from the start primarily due to the difficulty in measuring and obtaining data on inventive activity.

Until recently, the state of the art (Trajtenberg and Jaffe, 2002) in estimating the knowledge function in knowledge intensive industries has used R&D expenditures, patent counts and citation-weighted patent counts as traditional measures of inventive activity. I outline the usage and limitations of these measures below.

5.2.2. Traditional Innovation Measures

R&D Expenditures. The original analyses of inventive activity focused on R&D investments as a proxy for a firms' knowledge capability (Griliches, 1979). As R&D is considered to be the greatest source of new economic knowledge (Cohen & Levinthal, 1990), a stream of research has analyzed the contribution of R&D expenditure to economic impact. However, the use of R&D to measure technological change is limited for a number of reasons. First, quantitative reporting of R&D expenditures is required in only a few countries (most prominently the UK and US). Second, R&D is a measure of innovation input rather than output (akin to measuring how many have worked rather how much is produced). Furthermore, R&D measures reflect only those formal efforts to generate innovation (e.g. budgets for labs, see Kleinknecht and Verspagen, 1989), and do not take into consideration the important informal activities (Santarelli and Sterlachinni, 1990). Finally, R&D labs' budgets are dedicated not only to new innovation, but also to imitation and technology transfer (Mansfield, 1989). Despite these limitations, a consistent finding in the empirical evidence is that R&D expenditures explain a reasonable fraction of market value variation after controlling for ordinary assets (Griliches, 1981; Ben-Zion, 1984; Jaffe, 1986; Hall, 1993b, Blundell et al., 1995; Oriani and Sobrero, 2003).

Patent Counts. A second means of measuring innovation is the use of patent counts. Since the data became publicly available in the 1960s, scholars have used patents and patent counts as a measure of innovation output. Recent developments including the digitization of patent documents in electronic format and the creation of free databases (Hall et al., 2002) have prompted further studies. Patents were initially seen as a solution

to the problems of measurement in R&D expenditure, as patents provide detailed information about the innovation output (e.g. area of technology, inventors), are numerous, and have been reported in the same manner for over 100 years (Hall et al., 2002).

However, patents also suffer from a number of problems. First, simple patent counts do not reflect the heterogeneity in patent quality (Hall et al., 2002). Second, patents are a measure of intermediate output as patents reflect new knowledge of innovation, but not economic value. Most patented inventions do not result in innovation and furthermore, indeed most inventions which do result in innovations are not patented (Gittelman and Alcacer, 2004). The weakness of using patents as an independent measure of innovation has been noted over time (Scherer, 1983; Mansfield, 1989) and is perhaps best articulated by Griliches (1990: 1999) who pioneered their use: "Ideally, we might hope that patent statistics would provide a measure of the (innovative) output. . . the reality, however, is far from it. The dream of getting hold of an output indicator of inventive activity is one of the strong motivating forces for economic research in this area." A consistent finding in the patents literature is that they offer a weaker correlation to market value than R&D (for a thorough review of this literature, see Jaffe and Trajtenberg., 2005).

Citation Weighted Patents. Since the National Bureau of Economic Research (NBER)'s recent public release of citation-weighted patents data (detailed in Hall et al., 2002), a number of scholars have made use of this new measure (Narin et al., 1997; Scherer et al., 1998). These citations identify prior art and provide a map of linkages among inventions,

inventors, scientists, firm and geographic locations. Citation-weighted patents were thought to be a good measure of quality, which would be reflected in market value. Indeed, received citations have been said to be an indicator of the size of the technological footprint of the patent (Hall et al, 2002) and a number of studies demonstrate citation-weighted patent counts to be a measure of research quality (for a review of these studies see Trajtenberg, 1990). More recent research, however has highlighted a number of weaknesses associated with citation weighted patents. First, patents must be considered within the multifaceted legal and institutional environment for which they were designed, and as such are meant to classify and assert the patentability and property rights for an innovation. Patents make claims to intellectual property and citations are often used to scope the claims of earlier citations. Indeed, forty percent of patent citations have been found to have been added by the patent examiner (Alcacer & Gittelman, 2004).

5.2.3. New Measures of Innovation Output

Taken together, although R&D expenditures, patent counts and citation weighted patent counts have been used to explain market value, we have seen that there are numerous problems inherent in this approach. These problems are further compounded by their inadequacy to address the increasingly blurred divide between private-public collaboration. While previous research has described the spillover of knowledge from universities to private firms (e.g. Acs, Audretsch and Feldman, 1992), more recent studies explore the interactions between the two spheres, including collaborative research

and co-authorship of publications (e.g. Murray, 2002; Chapter 4). Indeed, this new mode of knowledge production at the interface of public and private organizations requires new measures of the quality of innovative activities, including scientific publication counts and citation-weighted publications (described below).

Citation-Weighted Scientific Publication Counts. The use of scientific publications in a market value equation of biotechnology firms to value knowledge assets is a key innovation in my research. That firms publish scientific papers in academic journals is well documented (Murray, 2004; Gittelman and Kogut, 2005; Chapter 4). What is less well understood is how publishing papers impacts performance. To the extent of the knowledge of this author, this is the first analysis to include scientific publications in a traditional market value estimation in biotechnology.

The simplest measure of scientific research output is to count the number of papers published by researchers employed in a biotechnology firm in a given time period. However, as in the case of patents, these a simple count measure does not account for quality and the impact of a scientific paper varies tremendously (Meyer, 2000). A range of bibliometric techniques have been developed in classifying, and measuring the impact of scientific papers. In the simplest form, the citations received by a scientific paper is a measure of its impact on the scientific community (for a more extensive discussion see Chapter 4). A deeper discussion of the use of citations follows later in this chapter.

In the next section, I outline a model of Open-Science Innovation that includes these new measures of innovative activity (i.e., paper counts and citation-weighted papers), along with traditional R&D expenditures, patent counts and citation-weighted patents measures.

5.3. Theoretical Framework

In this chapter I propose an open science framework of innovation[29], whereby firms make investments in R&D to develop their innovative capability. En route, they not only obtain patents but also participate in public science by publishing scientific papers. Thus, at any given time, we are able to record three dimensions of a firm's innovative activity:

1. R&D investment stocks
2. Citation-weighted patent stocks
3. Citation-weighted publication stocks

In the remainder of this section, I outline the information contained in each of these measures and then show why the latter dimension, despite having never been used until now as a predictor of financial success, is a crucial independent measure of innovation in its own right.

1. R&D Stocks. At time τ there is a component of a firm's R&D investment (its most recent expenditure) that has not yet had a chance to produce either patents, scientific papers or tangible products. Thus, the total R&D stock observed at time τ has two

[29] A similar model may be found in the community of open-source software. Work by von Hippel (1987, 2001) and others have shown that rivals or competitors in the software labor market can willingly co-operate and trade-know in return for acceptance and status in the community. Similar norms of practice exist in other areas of science including biotechnology.

components: the old R&D whose output can be readily judged, and new R&D that is still unproven[30]. In this chapter, I argue that this stock of unproven R&D reveals the commitment of a firm's resources to innovation. This interpretation is strengthened by the findings of Toivanen, Stoneman and Bosworth (2002) who report, using a fourteen year panel data set, that R&D has a stronger effect in the market value regression in the time period following the release of its R&D budget (i.e., when it first announces the figures publicly). Further, in their annual reports, biotech firms typically use phrasing[31] that signals commitment when reporting their R&D budget for the upcoming year.

2. Citation-Weighted Patent Stocks. The goal of a successful R&D program is to secure valuable patents that translate into products and services that would earn a firm revenues. However, as noted earlier, when we observe R&D stocks at time τ, due to the lag inherent in research work, the stock of new R&D investment has not yet had an opportunity to produce patents.

However, by observing the stock of citation-weighted[32] patents previously accumulated by a firm, we obtain a measure of a firm's past innovative success. Thus, notwithstanding any major structural changes within the firm, it is reasonable to expect that its previous

[30] There is no way, ex ante, to separate old and new stocks of R&D with any degree of accuracy. Thus the separation of R&D stocks into "old" and "new" R&D is a purely notional concept rather than an empirical distinction.

[31] Hedge funds and other financial institutions also use a similar language. For example, when discussing the prospects of a portfolio biotech firm, the Annual Report 2005 of the International Biotechnology Trust interprets its R&D outlay for the following year as "(X firm)...commits $485.34 million dollars to research and development" (International Biotechnology Trust 2005).

[32] As noted earlier citations listed on a patent are primarily listed as markers of 'prior art'. Thus, patent citations play an important legal function rather than accurately record the relative importance of patents. We find that even after weighting patents by the citations they receive, patents remain highly skewed in their economic value (Hall, Jaffe, and Trajtenberg, 2000).

patenting success (represented by its patent stocks) are a guide to the patenting productivity of a firm's new R&D stock.

3. Citation-Weighted Publication Stocks. While patent counts are a measure of innovative success, they are incomplete due to the variability in their economic value. I argue that the citation-weighted stock of scientific papers authored by a firm provides a valuable signal of the future economic value of its patent stock by measuring the quality of its tacit scientific knowledge base.

Scientific papers are read and cited by other scientists. By citing a paper, scientists acknowledge a direct scientific debt. Therefore, the more citations a paper receives, the greater has been its impact in the scientific community. Thus, by weighting publication counts by citations, we obtain a measure for the quality of science produced by a biotech firm as judged by a cohort of fellow scientific experts who are who are usually defined in a niche area of specialization.

Thus, I argue that the three measures of innovation discussed above (R&D, citation-weighted patents and citation-weighted publications) are not interchangeable measures of innovation but rather reveal different dimensions of a firm's innovative effort. If this assertion is true, entering R&D, patents and citation-weighted publications into a single market value equation should reveal separately the information contained by them. Therefore, in the next section, I construct an empirical model on this basis and rigorously test this assertion using unique data from the U.K. biotech industry during the years

1988-2000, in order to flush out the theoretical and empirical implications of this open-science model of innovation.

5.4. Empirical Model

Using firm market value as the measure of firm performance relies on the fact that firms are composed of bundles of assets (both tangible as well as intangible) whose values are determined by trading of the firm's shares on the financial market. The typical model of market value hypothesizes that the market value of a firm is a function of the set of assets that it comprises:

$$V(A_1,A_2,A_3,...) = f(A_1,A_2,A_3,...) \qquad \text{(Equation 5.1)}$$

where f is an unknown function that describes how the assets combine to create market value. If the firm invests in the various assets $A_1,A_2,A_3...$ according to a value maximizing dynamic program, and if the stock market is efficient, the function f will be the value function associated with that dynamic program.

However, in estimating Equation 5.1 we run against a problem: its functional form is not well understood, nor is it easy to compute from an explicit dynamic program or maximization model (Hayashi and Inoue 1991). There are two predominant specifications of the value function, an additively separable linear specification[33] and a multiplicative separable specification of the Cobb-Douglas form.

[33] For theoretical development of this usage, see Hall (2000). For the first empirical usage see Griliches (1981), later developed by Hall (1998).

Of these, the linear specification is more commonly used for two reasons. In the linear, additively separable version, the marginal shadow value of the assets is equalized across firms, while the Cobb-Douglas form assumes that the value elasticity is equalized. While the constant elasticity form tends to fit the data better, a constant shadow value across firms is more defensible from a market efficiency point of view. The better fit provided by the Cobb-Douglas form is because this specification is less sensitive to outliers. However, outliers are important in knowledge-based industries. Thus, to discount the importance of high value patents or highly cited publications weakens the interpretation of the market value equation in the context of high tech.

For these reasons, in my analysis, I adopt the linear specification[34] as was first used by Griliches (1979). This model is given by:

$$V_{i,t}(A,K) = q_{i,t}(A_{i,t} + \gamma K_{i,t})^{\sigma} \qquad \text{(Equation 5.2)}$$

In general, q may vary across firms and time. Thus,

$$q_{i,t} = \exp(f_i + m_t + u_{i,t}) \qquad \text{(Equation 5.3)}$$

where f_i is a firm effect, m_t is the market effect at time t, and $u_{i,t}$ is an independently distributed error term. Hence the term q allows for the fact that the market valuation may vary across firms and time, and that there also may be 'noise' in the valuation.

Substituting Equation (5.2) in Equation (5.3) , I obtain

$$V_{i,t}(A,K) = e^{(f_i + m_t + u_{i,t})}(A_{i,t} + \gamma K_{i,t})^{\sigma} \qquad \text{(Equation 5.4)}$$

Taking logarithms both sides, I arrive at

[34] In contrast, the log-linear model has the Cobb-Douglas form $V_{i,t}(A,K) = q_t A_{it}^{\sigma_t - \alpha_t} K_{it}^{\alpha t}$. I also run the model using this specification, and find, that the broad results remain unchanged. The results are available from the author upon request.

$$\log V_{it} = f_i + m_t + \sigma_t \log(A_{it} + \gamma_t K_{it}) + u_{i,t} \quad \text{(Equation 5.5)}$$

By re-arranging the terms, I can rewrite (5.5) as

$$\log V_{it} = f_i + m_t + \sigma_t \log[A_{it}(1 + \gamma_t K_{it} / A_{it})] + u_{i,t} \quad \text{(Equation 5.6)}$$

which gives,

$$\log V_{it} = f_i + m_t + \sigma_t \log A_{it} + \sigma_t \log(1 + \gamma_t K_{it} / A_{it}) + u_{i,t} \quad \text{(Equation 5.7)}$$

which is the formal model that I estimate in my analysis.

Griliches (1981) takes the $\log A$ term to the left hand side of Equation 5.7 and uses $\log(V/A)$ as a dependent variable. Other studies include the $\log A$ as an explanatory variable (Hall 1993a, Bosworth and Rogers 1998). These variations are all essentially modifications of Equation 5.7 and do not challenge the basic empirical relationship being tested in any fundamental way.

The variable f_i, which captures firm effects, can be proxied by other explanatory variables. Past revenue or profit growth for the firm are sometimes taken to imply high future profits and hence high market values. Firms may also have different risk premiums depending on their exposure to debt and shareholder's equity. I describe set of variables that capture this firm heterogeneity in section 5.5.2.3. I capture the effect of variable m_t by including year dummies. Estimating $A_{i,t}$ is a relatively straightforward affair as data on tangible assets of a firm are readily available from business databases and annual reports (see 5.4.1).

In the market value equation (5.7), the impact of firms' innovation is measured by $K_{i,t}$ which I proxy by R&D expenditures, patent counts and citation-weighted counts of scientific publications as described in Section 5.2.3 and 5.5.2.2. In order to assess the relative quantitative impacts of the three measures of innovation on market value, I need to compute the semi-elasticities[35] for the different stocks. For R&D stocks, I obtain,

$$\frac{\partial \log MKTVAL_{i,t}}{\partial (R\&D/A_{i,t})} = \hat{\gamma}_1 \left(1 + \hat{\gamma}_1 \frac{R\&D}{A_{i,t}} + \hat{\gamma}_2 \frac{PAT}{A_{i,t}} + \hat{\gamma}_3 \frac{PAPERS\_CITES}{A_{i,t}} \right)^{-1}$$

(Equation 5.8)

and similarly I derive the semi-elasticities for the patent and citation-weighted publication stocks.

5.5. Data and Variables

5.5.1. Data

I obtained a list of 113 U.K. biotechnology firms from surveying the financial press and an exhaustive biotechnology industry report prepared by the Department of Trade and Industry (2005). The financial and accounting data of these firms was sourced from the Financial Analysis Made Easy (FAME) database. FAME contains a comprehensive list of 2.9 million U.K. and Republic of Ireland companies. It has a clear advantage over other competing databases such as DataStream by focusing purely on U.K. and Irish companies

[35] This is a partial derivative, holding the other ratios constant.

(including those that are not publicly listed) and offering a greater breadth of information on these firms. FAME includes not only standard financial and accounting data (including R&D expenditures) but also information on the composition of the board and senior management, the holding structure of the company, its main competitors, the most profitable companies in its peer group and the latest news about the company from the business press. This information is very useful for constructing a wider range of performance variables as well as building a deeper intuition for the mechanics of the industry acquired whilst constructing the dataset.

The dataset of scientific papers produced by U.K. biotechnology firms during the years 1988-2001 is obtained from the Research Output Database (ROD) and from a tally of 355,183 individual scientific papers in the original ROD database, I obtain 2,915 papers firms that list at least one author from a biotechnology firm as a contributing author (as identified by the address field). I have described this dataset in more detail previously in Chapter 4.

The patent data are sourced from the computerized database available from the U.S. Patent Office. The decision to use U.S. as opposed to U.K. patents was motivated by the desire to screen out the numerous low value patents taken out each year. Given the strategic importance of the U.S. market, all patents seen as important, are patented in the U.S. (Kretschmer et al., 2005). Using U.S. patents only also sidesteps the problem of duplicate patent counts (awarded in the UK and in the U.S.) since the U.S. patents include global patents registered in the U.S. There are some caveats though: the low cost

of taking out a patent in the US combined with the prevalence of protective patenting even when the actual value of a patent is uncertain might lead to frivolous patenting behavior in the US market and it is not clear that only the best patents are taken out in the US market.

Because of the way in which the patents are organized within the data base, matching the patents owned by a firm to firm level data is a challenging task. Firms patent under a variety of names (their own and those of their subsidiaries) and the Patent Office does not use a unique identifier linking the disparate patentees. In order to obtain a clean dataset of patent ownership, I had to match names of patent owners across a variety of ownership structures. To do so, I used a variety of sources; predominantly the FAME database but also the Who Owns Whom Directory of Corporate Affiliations (2002), in addition to referring to primary sources such as annual reports. Of the 113 firms in my dataset, I corrected ownership data to 67 of them (over 50% of them). The mean number of patents without the ownership correction was 11.72, while the corrected mean stood at 3.321. This is a significant correction in the data and owes to the fact that without cleaning up the ownership data, patents are usually counted more than once as the same patent is shown to be owned by different entities which are really the same. Thus, this is a significant correction to the patent data used in my dataset.

5.5.2 Variables

Tables 5.1 provides a listing of my analytical variables with a brief description of their definitions. Table 2 contains descriptive statistics for my key variables.

5.5.2.1. Market Value of Biotechnology Firms

The market valuation of biotechnology firms is my central variable of firm performance. From 1988-2001, I obtained the daily market value of a firm's issued common stock for those firms that were listed on the London Stock Exchange (for the days the company traded) from the FAME database. Of a total of 113 firms that were originally in my database, I reduced the number of firms to 87 resulting in a total of 1131 observations in order to obtain a balanced panel data for the entire period[36]. I then found the average market capitalization of its issued common shares over the year. To this number I add the market value of the debt (extracted from annual reports) to arrive at an estimate of the total market value of a firm i in year t. (recorded as $MKTVAL_{i,t}$). Of the 1131 observations in my dataset I corrected for various changes that could effect its market valuation. Particularly, important were stock splits, mergers and acquisitions. I did this by running simple tests on my dataset to record sustained dips or increases in market valuation and stock price in a single direction (to weed out the fairly high volatility prevalent anyway in biotech asset prices) . Of the 1131 observations there were around 226 such incidents linked to 83 out of the 87 firms in my dataset. I then matched the observations with the business press (using F.A.M.E, the Financial Times, the Wall Street Journal and other online sources) to take note of mergers or acquisitions that could have

[36] The results remain consistent when I run regressions for the entire panel using an unbalanced panel data specification. The key advantage in using a balance panel data is because is to avoid issues of entry and exit entering into my analysis. I also run my regressions for those firms that were in existence in 1988 between 1992-2001, allowing for 5 years to build knowledge stocks. The results are broadly comparable and I retain my analysis at the level of the balanced panel. The results for these alternative specifications are available from the author upon request.

impacted the stock. I found that of the 87 firms used in my observation, some 11 firms were acquired by other firms in my dataset. I adjusted my dataset to reflect this change. I also tracked another 114 observations to M&A activity related to firms outside my dataset. Of the remaining 101 observations, I tracked down 32 stock splits (these were easier to observe from the data as the stock price usually halved with the underlying market valuation staying the same) and adjusted for these well. The remaining 62 fluctuations remain unverified but can probably be attributed to market volatility or market reaction to actions of biotechnology firms (for example, the success or failure of a drug in the FDA approval process).

5.5.2.2. Measures of Scientific Knowledge

The innovative capability of a biotechnology firm is proxied by three separate measures: R&D investments, patent counts and citation-weighted publication counts. The constructed variables are described below:

R&D Investments. In my sample, R&D investment makes up 40% of a firm's operating expenditure. R&D expenses are an expression of the technical aggressiveness of these firms, which is an important strategic attribute in this industry. For example, the term "burn rate" is used by venture capitalists and industry analysts to describe the high rate of R&D spending per month which is then used to calculate the number of months the firm can survive. Further, in asset light biotechnology firms, firms' R&D investments are its

most visible asset and therefore play a dominant role in its market valuation (Chan et al., 1999).

R&D expenditures are obtained from the FAME database and verified by cross checking against a firm's annual report (disclosure of R&D expenditures is compulsory according to U.K. company law). I record the variable $R\&D_{i,t}$ as the R&D expenditure of firm i in year t.

Patent Counts and Citation-Weighted Patent Counts. As seen earlier, the use of patent data in economic analysis stretches back to the work of Schmookler (1966) and Scherer (1965). The availability of information from the U.S. patent office in machine-readable form in late 1987 spurred greater use of patent information in the economic analyses of innovation and technological growth. Schankerman and Pakes (1986) provide a broad overview of the use of European patent data. A fuller review of the empirical work using patent data can be found in Jaffe and Trajtenberg (2002).

I record variable $PAT_{i,t}$ as the number of patents earned by firm i in year t. I weigh the raw patent counts by multiplying the individual patents by the number of citations (recorded as $PATCITES_{i,t}$) received by the patents from the date it is issued until five years after (i.e., if a patent is issued in 1989, I record all citations till the end of 1994), and record this variable as $PATWEIGHTED_{i,t}$[37].

[37] An alternative weighting scheme is used by Hall, Jaffe and Trajtenberg (2000) in which rather than arbitrarily truncating the citation distribution after a period of five years, they use a weighting algorithm to estimate the "true" citation score over the lifetime of an awarded patent. I run my results using both the

Publication Counts and Citation-Weighted Publication Counts. The original key variable in my analysis is the citation-weighted stock of scientific papers produced by a biotechnology firm. My database includes all scientific papers published in the UK and Ireland, during the years 1988-2000 which include at least one biotechnology author in the address field (see Section 3.1). From this database, I construct a simple count of all scientific papers authored by a biotechnology firm in a given year (denoted by $PAPERCOUNT_{i,t}$). I then weigh this simple count by two different citation measures: first, a simple citation based count, and, secondly, a measure based on the prestige of the journal in which a scientific paper is published. While both these measures capture research quality, they are based on slightly different logic, which is outlined later in the section.

First, I total the number of citation counts to individual papers going forward five years from the date of publication. I denoted this total number of citations received as $CITES_{i,t}$. The citations are tracked across scientific publications present in the Science Citation Index (SCI). Only citations from original research articles were included, eliminating meeting notes, review articles and book reviews, to more accurately gauge genuine impact. I then weigh the publications by the logarithm transformation of the raw citation score:

---

patent as well as the publication citations using this method and find no significant alternation to my results (available upon request) primarily because 80% of the citations to both patents as well as publications are received within the first five years of their lifetime (Lewision et al, 2003) and thus the additional sophistication offers very little empirical traction.

$$\log citations = 1 + 2\log_{10}(citations + 1) \qquad \text{(Equation 5.9)}$$

where *citations* is the total raw citation count to an individual paper. I record this variable as $LOGCITES_{i,t}$. Thus when a paper gets no citations, $LOGCITES_{i,t} = 1$, and when it receives 100 citations, $LOGCITES_{i,t}$ equals about 5. This roughly corresponds to the relative subjective weight that scientists and scientific administrators give to the relative importance of the papers in the two instances, which is not 100:1 but 5:1[38]. Thus, I can then multiply $PAPERCOUNT_{i,t}$ and $LOGCITES_{i,t}$ to obtain an accurate citation-weighted publication count. This variable is recorded as $PAPERS\_LOGCITES_{i,t}$.

An alternate measure of research quality uses the journal in which a paper is published as a proxy for the quality of the article. Articles that are published in higher prestige journals (i.e., journals which contain articles that receive a higher than average number of citations) are judged to be of higher quality than lower ranking journals. It is important to note that the more direct measure of citation counts and the second measure of the journal quality to proxy for the quality of the article are similar measures but not identical. The two measures reveal how the paper has been judged by two different readerships: the general body of researchers in the case of direct citation counts, and, in the second case, a journal editor and a few specialized reviewers.

To create my research quality variable based on journal quality (denoted PIC standing for Potential Impact Category), I record the mean number of citations to papers published in

[38] In two separate polls (Lewison et al., 2003), scientific administrators and researchers in the biomedical sciences voted the relative importance of papers in "excellent" journals about four to five times that of papers in "ordinary" journals, and that of paper in "good" journals about two to three times of the latter.

the journal in a given year that are received in the year of publication through the fourth year after publication. The fortunes of journals rise and fall and by capping citations to five years, an additional benefit is that the citation weight more accurately captures the relative quality of a journal during the time period an article is published. I then use the log transformation of Equation (5.9) on the mean journal citation scores to more accurately reflect the perceptions of scientific administrators and medical researchers (see above). I record the variable $PAPERS\_PIC_{i,t}$ as $PAPERCOUNT_{i,t}$ times $PIC_{i,t}$ as a measure of scientific output controlling for quality as reflected by the relative prestige of the journal in which the article is published.

Constructing Stocks for R&D, Patents and Citation-Weighted Publications. To estimate my market value equation, I calculate the "stock" of R&D, patents and citation-weighted scientific publications accumulated by a firm at a given moment in time. To do so, I need to construct equations that capture the flow of R&D expenditures, patents and papers outlined earlier in a robust stock. I outline my methods below.

The following equation is used to approximate the stock of R&D of firm i in year t (denoted by $KR\&D_{i,t}$):

$$KR\&D_{i,t} = \sum_{k=0}^{\infty} R\&D_{t-k}(1-\delta)^k \qquad \text{(Equation 5.10)}$$

where t denotes the current time, $t-k$ denotes past periods, δ denotes the constant rate of depreciation, and k denotes the number of compounding periods. I fix the rate of depreciation at 15 percent keeping with many previous studies in the USA and UK, notably Ben Zion (1984) who uses UK data.

I construct the patent stock variable of firm i in year t (denoted by $KPAT_{i,t}$) along similar lines as the R&D stock shown previously. Thus,

$$KPAT_{i,t} = \sum_{k=0}^{\infty} PATWEIGHTED_{t-k}(1-\delta)^k \qquad \text{(Equation 5.11)}$$

where, as previously, t denotes the current time, $t-k$ denotes past periods, δ denotes the constant rate of depreciation, and k denotes the number of compounding periods.

As above, the stock of scientific publications of firm i in year t is constructed using

$$KPAPERS\_LOGCITES_{i,t} = \sum_{k=0}^{\infty} PAPERS\_LOGCITES_{t=k}(1-\delta)^k \qquad \text{(Equation 5.12)}$$

(replacing $PAPERS\_LOGCITES_{i,t}$ by $PAPERS\_PIC_{i,t}$, in my alternative estimation of research quality.)

5.5.2.3. Control Variables

The empirical specification includes a host of financial and accounting measures as control variables in order to more completely isolate the impact of firms' knowledge assets on its market value.

To the extent that they are valuation-relevant, the inclusion of control variables allows tighter inferences about the contributions that intangible assets make to market value. There is a number of financial and accounting variables in which financial markets take an interest and which prove to be statistically relevant to firm performance. Information on these are easily found in the firms' annual statements on income, cash flow, and the

balance of assets and liabilities as well as from the primary database FAME. The choice of control variables is based on both statistical relevance as evidenced in the dataset and theoretical consistency. A review of the corporate finance literature (see Ross, Westerfield and Jaffe 1996) led me to the following variables – profit before tax, long term investments, equity and current assets.

5.6. Results

5.6.1. Baseline Estimate

In my empirical model (Equation 5.7), market value is assumed to depend on three measures of innovation: R&D expenditures, citation-weighted patent yield, and on the stock of citation-weighted publications.

Table 5.3, Column 1 presents the estimation of Equation 5.7 with only R&D and patent stocks. Table 5.3, Column 2 presents the full estimation with R&D and patent stocks and citation-weighted publications. I find that adding scientific citation stocks improves the predictive power (as measured by the R-squared) of the model.

Thus, it is clear that each of the measures – R&D, patent counts and citation-weighted publications are not just interchangeable measures of innovation but convey valuable information on the market valuation of a biotech firm on top of what could be inferred from the other variables. For a given level of R&D spending, firms that manage to patent

more will presumably have higher market valuations, and also for firms with similar patent portfolios, higher quality scientific research as measured by citation-weighted publications will generate higher market valuation.

In order to compare the coefficients of the three measures of innovation, I compute the semi-elasticities of the various stocks as described in Equation 5.8 using the estimates $(\hat{\gamma}_1,\hat{\gamma}_2,\hat{\gamma}_3)$ from Table 5.3, Column (2) and present the results in Table 5.4. I find that an increase of one percentage point in the R&D intensity of a firm (i.e., in the ratio of R&D/Assets) leads to a similar increase in market value (about 0.8%, to be precise). An extra patent/million($) of R&D boosts market value by about 2% and an extra citation to a scientific publication authored by an employee of the firm per patent by around 0.013%. Thus, on average, 14 published scientific publications make approximately the same contribution to a firm's market value as a single patented innovation.

This is a very interesting result, highlighting the importance of scientific papers in the market value equation. It also confirms the working hypothesis of my analysis; that scientific papers are an important signal of scientific capability (a signal which is independent of the information contained in patents and R&D expenditures) and will therefore be highly correlated with market value. It is also interesting to note that the statistics in Table 4 make it clear that the ratios (when computed at the mean, median and ratio of totals) are far from normally or even symmetrically distributed, which suggests that some exploration of the distribution of the citations, such as the impact of

highly cited publications might prove insightful. I proceed on this basis in the next section.

5.6.2. The Impact of Highly Cited Papers on Market Value

Table 5.4, which evaluates the impact of citation-weighted publications at the mean, median, and ratio of totals, alludes to the fact that the distribution of citations is very skewed. In fact, about one quarter of all scientific papers get zero citations, and only a few dozen (out of tens of thousands) receive 100 citations or more[39]. This suggests that the average effect that we get in my baseline regression (Table 5.3, Column 2) may not reveal the full extent of the impact of the tail of the citations distribution.

To explore the impact of the citations distribution on the baseline market value equation (Equation 5.7), I break the scientific citations/patents variable up into five groups, where the number of citations received are < 5, 5-6, 7-10 (6 is the median), 11-20, > 20, and include dummy variables for each (the first serves as the base category). As shown in Table 5.5, Column 2, for firms with fewer than the median number of scientific citations per patent, it makes no difference how far below the median they fall: firms with 5-6 citations per patent have only slightly higher value than those with less than 5. However, firms that average more than the median number of citations per patent exhibit a very significant increase in market value: 10% higher if having 7-10 citations per patent, and 35% higher if having 2-3 times the median (11-20 citations per patent). The most

[39] The skewness of the citation distribution is consistent with bibliometric studies of scientific research (Bosworth and Mahdian 1999).

dramatic effect are for those scientific papers that receive more than 20 cites per patent: the market value of these firms is 54% higher than if the firm's paper output had generated 4 or fewer citations.

These statistics are not as surprising as they appear. The role of "star science" in the success of biotechnology firms has been well documented. Research by Zucker, Darby and Armstrong (2002) on the U.S. biotechnology industry for example found that a robust indicator of firm's tacit knowledge capture (and strong predictor of its success) is the number of research articles written by firm scientists in collaboration with "star" scientists, usually working at top universities. Here I find that highly cited scientific publications are more strongly correlated with market success than mediocre ones, and in my previous work (Chapter 4) show that biotech firms have a greater chance of producing highly cited papers by collaborating with "star" scientists, i.e., scientists with established research reputations and usually employed at top research universities. Thus, nurturing and maintaining links with public science such as universities and research institutes, appears to be crucial to the future financial success of biotech firms.

5.6.3. The Impact of Citations by Pharmaceutical Firms on Market Value

Large pharmaceutical firms play a key role in the biotechnology industry. Research has shown that young biotechnology firms benefit greatly from collaborations with "big pharma" (Pisano 1991; Zucker and Darby 1998). These gains can be material, with large pharmaceutical firms making venture capital investments in promising young

biotechnology firms, or, as is more often the case, relational, with young biotechnology firms working jointly on collaborative R&D with large pharmaceutical firms. Scientific citations are a mechanism of expressing intellectual debt and, therefore, I would expect a greater impact to a biotechnology firm's market valuation from a citation by a pharmaceutical firm than from a more "generic" citation.

To capture the difference in these citations, I break up my citation stock into two categories – citations received from large pharmaceutical firms[40] ("big pharma") and citations received from others. Table 5.6, Column 2 displays the results after splitting my citation stocks in this way. I find that the coefficient for pharmaceutical citations is higher than non-pharmaceutical citations by a ratio of nearly 10:1. This suggests that in terms of market value, a pharmaceutical citation is worth nearly ten times more than a regular citation. Thus, my results agree with previous research on the role of large pharmaceutical firms in mentoring smaller biotechnology firms, and that pharmaceutical citations are a strong signal of such activity[41].

5.6.4. The Impact of Interaction Effects on Market Value

[40] I collated a list of 27 pharmaceutical firms from industry databases that I matched with the citation data. My database of scientific publications includes an address field that allows me to identify pharmaceutical citations from the others.

[41] "...we track our citations to keep a pulse on the evolving scientific frontier and when pharmaceutical firms cite our research, we approach them to see if there is any commercial opportunity... in particular, if our work is cited heavily by a particular pharmaceutical firm, a working relationship, either contract R&D or joint product development usually follows..." (Interview with Dr. J. Blumenthal, Celltech, UK, Feb. 2005).

I include three interaction terms of (R&D*Citation-weighted Patents), (R&D*Citation-weighted Publications) and (Citation-weighted Patents*Citation-weighted Publications) in my market value estimation (Equation 5.7). The results are displayed in Table 5.7.
I find that the three interaction terms are positive and significantly correlated. From the coefficients in Col (2), Table 5.7, I find that the presence of patenting and publishing significantly raises the impact of R&D on market value. There are two possibilities, first that R&D acts as a complement to the joint activities of publishing and patenting and we would expect them to appear in a bundle with higher market value. Thus, there is a case to be made that firms that invest in R&D reap the greatest benefit by combining these investments with high levels of patenting and publishing.

However, a weaker form of joint adoption could also originate from a common factor driving the adoption of both – for example, smart firms, those with higher market value, might learn how to manage how to juggle R&D, patenting and papers in an optimal mix. In this case, we would again expect R&D, patenting and publishing to appear in combination with each other, and firms that draw higher market value from the adoption of one will also draw positive benefits from the other.

It is important to note that in the context of the model used in this study, with R&D, papers and patents signaling the financial market of the potential success of the firm, the two observationally identical outcomes generate the same impact on market value.

5.6.5. Tests for Causality

While the results show us that the stock of citation-weighted publications acts as a signal of the quality of firms' R&D, at this stage, the direction of causality is unclear. Does publishing papers improve the quality of a firm's R&D or does the stock of published papers merely reflect the existing levels of R&D quality within the firm? If the latter argument were true, it would suggest that publishing papers are a fundamental component of a firm's innovative program[42] and are a key knowledge asset in their own right.

One possible strategy to establish the direction of causality would be to proxy the knowledge "capabilities" of a firm by some variable and regress this measure against the stock of scientific publications. The simplest outcome of innovative capability are citation-weighted patent counts. Another measure would be its ability to generate profits.

I regress the citation-weighted publication stock against both these variables[43]. In Table 5.8, Column 1, I find that citation-weighted patent counts are weakly correlated with profitability and that citation-weighted publications are strongly correlated with profitability. Further, in Table 5.9, I find that the stock of published papers is weakly correlated with citation-weighted patent counts. As noted earlier in the chapter, because

[42] Drawing upon Cohen and Levinthal's (1989) "absorptive capacity" argument, Cockburn and Henderson (1998) suggest that firms use pro-open-science incentives to develop routines and skills that allow them to utilize effectively the advances in publicly funded research.

[43] I include a time lag of three years to capture the gap in time between patenting and publishing activity and its impact of profitability. A lag of three years gives me the strongest result and is consistent with previous results on R&D investments, patenting and profitability (reference).

of the skewness of patent value (even when weighted by citations), this result is not particularly worrisome[44].

Thus publishing papers appears to be an activity that imparts distinct benefits from patenting but with an independent relationship on both market valuation as well as profitability. However, it may be that publishing good papers might allow a firm to signal that it provides a welcoming environment to do high quality scientific research and therefore attract and retrain high talent scientists to the firm. In other words, there could be a two-way effect between publishing papers and a firms ability to innovate. A true *ex ante* separation of variables needed to establish causality is difficult to establish on purely empirical grounds and is beyond the scope of my analysis. However, in subsequent work I explore the theoretical basis for firms participating in the world of public science.

5.6.6. Tests for Robustness

5.6.6.1. Measuring Research Quality by Journal Prestige

Rather than using citation scores for individual papers, I can proxy the quality of an article by the quality of the journal in which it appears (See Section 5.5.2.2 for an account of the construction of this variable). This is not simply a "cruder" measure than raw citation scores, but rather captures a slightly different perception of the article. It reflects an estimate of its quality as judged by a smaller group of peer-reviewers and journal editors. Also, often times, articles that appear in top ranking journals receive a fair share

[44] A recent paper by Gittelman and Kogut (2003) that regresses high impact papers against high impact patents, indeed report a negative correlation between the two.

of attention and deliver benefits to the authoring firm even though they do not garner many citations.

I run Equation 5.7 using journal quality instead of citation scores to individual papers and show the results in Table 5.10, Column 1. I find that while the R-square remains largely unchanged (see Table 5.3, Column 2), the coefficient of the publication stocks variable is higher. Thus, getting published in high prestige journals has a greater impact on the market value of a firm than the total citations that a paper eventually receives. This is not surprising since investors and potential collaborators regularly scan top journals and by getting published in these journals, biotech firms can more convincingly signal the worth of their ideas.

5.6.6.2. Multicollinearity

Table 5.11 reports the cross-correlation of the key explanatory variables. Not surprisingly, I find that a firm's R&D expenditure is correlated with the number of patents that it might earn and the number of publications that it produces. This result is reasonable and not particularly counter-intuitive. However, the first point to note is that I find that these variables are not *perfectly* correlated. Thus, while R&D and patenting activity are correlated, not all of a firm's patenting or publishing activity can be explained by its R&D expenditure. Thus, while there is an overlap between these measures, they still independently contain useful information.

Second, I find that while R&D expenditure is strongly correlated with a firm's publication output in numerical terms, it is less powerful as a predictor of the number of citations these publications receive. In other words, citation-weighted publications measure an aspect of innovation distinct from R&D expenditures, namely the quality of a firm's scientific innovative activities as perceived by a peer group of experts.

Further, I find that while R&D expenditures (and patent counts) are not strongly correlated with profit before tax, citation-weighted publications are. This also supports the hypothesis that the information contained in the citation-weighted publication stock contains separate and useful information on firm performance, namely a signal of the quality of its innovative efforts.

5.7. Conclusion and Suggestions for Future Research

My analysis uses new data on UK biotechnology research output from 1988 to 2001 to build an open-science model of innovation in biotech that argues that R&D expenditures reveal the commitment of a firm's resources to innovation, patents record its success at innovation, and a firm's stock of citation weighted scientific publications signal the quality of its innovative efforts. The original contribution of my analysis, beyond assembling a new dataset, is to include citation-weighted publications in a market value equation of biotechnology assets and to analyze the strategic import of this variable.

By publishing scientific papers subject to a peer review process, firms are able to convincingly signal the worth of their ideas to investors and other interested third parties such as potential collaborators and employees. This is a valuable signal given the uncertainties of biotechnology research and the findings indicate that publishing papers in academic journals provides real financial gains to firms. On average, publishing fourteen scientific papers in academic journals has approximately the same impact on market value as obtaining a patent. Further, highly cited papers and those papers that are cited by pharmaceutical firms have a much greater impact than average.

At a prescriptive level, then, the implications of these results are straightforward, with the old adage "Publish or Perish" applying equally to biotechnology firms as it does to academic environments. Biotechnology firms need to continually publish scientific papers in order to convince investors and potential collaborators of the worth of their ideas. A firm's market valuation is inextricably linked with this public signal of R&D efforts.

Previous research shows that biotechnology firms can boost their publication record by collaborating with "star" scientists and institutions (see Chapter 4). By star, I mean those institutions and individuals with particularly strong scientific track records. Collaborating with stars produces papers that are more highly cited and that appear in better quality journals. Thus forming links with top academic institutions is a strategic imperative for biotechnology firms. From a public policy perspective, developing and nurturing a strong independent scientific establishment is one of the key ingredients in fostering a

competitive biotechnology industry. Further, arrangements that allow for the free mobility of scientists across the private-public divide are a crucial requirement for ensuring that the private sector can then capitalize on the intellectual and human capital available in the public sphere.

It is important to note that previous research indicates great variance across industries in terms of the propensity to patent and to cite patents (Levin et al, 1987), the count and citation weighted count of publications (Narin and Rozek, 1988) and the economic value of innovation (Acs and Audretsch, 1990) across industries, the findings may not be generalisable outside the high technology industry, or indeed even the biotechnology sector.

Finally, the question of why firms make their private knowledge public is clear. They gain financially from it. Further research on the incentives and outcomes of open-science innovation would be of much practical and theoretical interest.

Table 5.1. Definitions of Key Variables

Variable Name	Definition
1. Market Value	Algebraic sum of the market value of a firm's outstanding common shares and an estimation of the market value of its book debt.
2. Log Market Value	Logarithm of Market Value.
3. R&D Expenditures	R&D expenditure of a firm as recorded in its annual report in a given year.
4. Patent Counts	Total number of patents awarded to a firm by the USPTO in a given year.
5. Citation-weighted Publications Weighted By,	
a. Log Citations	Number of scientific papers authored by a firm multiplied by a logarithm weighted (see Equation 5.1) citation score.
b. Log Pharma Citations	Citations counted only as those received from pharmaceutical firms (using logarithm transform of Equation 5.1).
c. Log Citations to Journal Prestige	Average citations of journal used instead of citations to individual paper. Logarithm transform used (Equation 5.1).
6. Profit Before Tax	Declared profits before tax.
7. Ordinary Dividends	The dividends awarded by a firm to its shareholders the previous year (historic dividends).
8. Long-term Investments	The long-term investments and receivables held by a firm. Strong signal of a firm's cash flow in the absence of regular profits.
9. Current Assets	Firm's current assets that indicate the short-term management of cash flow.
10. Shareholder's Equity	The equity of the firm held by shareholders. Representative of exposure to capital markets and hence risk of financial distress versus growth prospects.

Table 5.2. Descriptive Statistics

Variable name	Mean	SD	Min	Max	Obs.
1. Market Value	117.45	0.034	10.092	4.732	1131
2. Log Market Value	2.069 (millions)	0.002 (millions)	1.004 (millions)	0.675 (billions)	1131
3. R&D Expenditures	11.532 (millions)	0.004 (millions)	1.782 (million)	1.84 (billions)	1131
4. Patent Counts	3.321	0.014	0	37	1131
5. Publication Counts Weighted by Citations as:	3.452	0.089	0	117	1131
5.1. Log Citations	7.112	0.321	0	202.708	1131
5.2. Log Citations By Pharmaceutical Firms	1.148	0.002	0	4.863	1131
5.3. Log Journal Quality	8.192	0.132	0	229.345	1131
6. Profit Before Tax	(millions) 8.783	(millions) 0.0019	(millions) 0	(millions) 432.34	1131
7. Ordinary Dividends	37.42	0.0048	0	732.12	1131
8. Long-term Investments	56.38	0.0061	0	234.32	1131
9. Current Assets	34.34	0.0043	0	261.32	1131
10. Shareholder's Equity	22.39	0.011	1.932	940.34	1131

Table 5.3. Market Value Equation as a function of R&D, Patents & Scientific Citations

Dependent variable: Log Market Value	Column 1 R&D and Patent Stocks	Column 2 R&D, Patent, and Publication Stocks
R&D Stock	1.276***	1.362***
	(0.061)	(0.068)
Patent Stock	0.027***	0.030***
	(0.006)	(0.007)
Citation-weighted Publication Stocks		0.052***
		(0.004)
Control Variables		
Profit Before Tax	0.343***	0.353***
	(0.015)	(0.014)
Cash Flows for Investment	0.034**	0.025**
	(0.002)	(0.001)
Ordinary Dividends	0.049*	0.037*
	(0.007)	(0.005)
Long-term Investments	0.018*	0.016*
	(0.001)	(0.001)
Current Assets	0.014*	0.013*
	(0.001)	(0.001)
Shareholder's Equity	0.135*	0.139*
	(0.007)	(0.008)
Adjusted R-squared	0.222	0.254
Number of Observations	1131	1131

Note: 1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.
4. Stocks are computed using 15 percent annual depreciation rate.

Table 5.4. Computing the Impact of Knowledge Stocks

Ratios	Ratios evaluated at the:		
	Mean	***Median***	***Ratio of totals***
R&D	0.35	0.16	0.16
Patents	1.05	0.35	0.50
Citation-weighted Publication Stocks	7.95	6.33	7.46
Semi-elasticities*			
$\frac{\partial \log MKTVAL}{\partial (R\&D/A)}$	0.709^{***} (0.206)	0.876^{***} (0.037)	0.842^{***} (0.036)
$\frac{\partial \log MKTVAL}{\partial (PAT/R\&D)}$	0.016^{***} (0.004)	0.019^{***} (0.005)	0.019^{***} (0.004)
$\frac{\partial \log MKTVAL}{\partial (PAPERS\_CITES/R\&D)}$	0.0013^{***} (0.000)	0.0033^{***} (0.002)	0.0032^{***} (0.002)

*Computed using the estimated coefficients in column (2) of Table 3.

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.

Table 5.5. The Impact of Highly Cited Papers on Market Value

Dependent Variable: Market Value	Column 1 R&D, Patent and Publication Stocks	Column 2 Impact of Highly Cited Papers
R&D Stock	1.362***	0.926***
	(0.068)	(0.053)
Patent Stock	0.030***	0.025***
	(0.007)	(0.004)
Citation-weighted Publication Stocks	0.052***	
	(0.004)	
Dummies for # of scientific citations*		
[0-4]	-	-
5-6	-	0.006***
		(0.013)
7-10	-	0.097***
		(0.018)
11-20	-	0.353***
		(0.023)
>20	-	0.542***
		(0.029)
Adjusted R-squared	0.254	0.255
Number of Observations	1131	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.
4. Stocks are computed using 15 percent annual depreciation rate.
5. [ ] refers to omitted reference category.

Table 5.6. The Impact of Citations by Pharmaceutical Firms on Market Value

Dependent variable: Market Value	Column 1 Without splitting stock into Pharma/Non-pharma	Column 2 Splitting stock into Pharma/Non-pharma
R&D Stock	1.362***	1.261***
	(0.068)	(0.063)
Patent Stock	0.030***	0.028***
	(0.007)	(0.007)
Citation-weighted Publication Stock	0.052***	-
	(0.004)	
Citation-weighted Publication Stock: Only Pharma Citations	-	0.431***
		(0.017)
Citation-weighted Publication Stock: Only Non-Pharma Citations	-	0.046***
		(0.006)
Adjusted R-squared	0.254	0.260
Number of Obs.	1131	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.
4. Stocks are computed using 15 percent annual depreciation rate.

Table 5.7. Interaction Terms

Dependent Variable: Market Value	Column 1 R&D, Patent and Publication Stocks	Column 2 Including Interaction Terms
R&D Stock	1.362***	0.926***
	(0.068)	(0.053)
Patent Stock	0.030***	0.025***
	(0.007)	(0.004)
Citation-weighted Publication Stocks	0.052***	
	(0.004)	
(R&D Stock*Patent Stocks)	-	0.695***
		(0.112)
(R&D Stock*Citation-weighted Publication Stocks)	-	0.474**
		(0.138)
(Patent Stock*Citation-weighted Publication Stocks)	-	0.015**
		(0.004)
Adjusted R-squared	0.254	0.312
Number of Observations	1131	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.
4. Stocks are computed using 15 percent annual depreciation rate.

Table 5.8. Do Papers Explain Profits?

Dependent Variable: Profit Before Tax (PBT)	**Column 1 R&D, Patent and Publication Stocks**
R&D	0.452***
	(0.114)
Citation-weighted Patents	0.112*
	(0.018)
Citation-weighted Publication Stocks	0.212***
	(0.079)
Adjusted R-Squared	0.225
Number of Observations	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.

Table 5.9. Do Papers Explain Patents?

Dependent Variable: Patent Counts (PAT)	**Column 1 R&D and Publication Stocks**
R&D	0.532***
	(0.102)
Citation-weighted Publication Stocks	0.115*
	(0.008)
Adjusted R-squared	0.210
Number of Observations	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.

Table 5.10. Impact of Journal Quality on Market Value

Dependent Variable: Market Value	Column 1 R&D, Patent and Publication Stocks
R&D	1.452***
	(0.044)
Patents	0.033***
	(0.004)
Citation-weighted Publication Stocks	0.066***
	(0.007)
Adjusted R-squared	0.259
Number of Observations	1131

Note:
1. ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level. Heteroskedastic-consistent standard errors in parentheses.
2. Estimation method: Ordinary Least Squares.
3. Firm dummies, time dummies and control variables included in all estimations.
4. Stocks are computed using 15 percent annual depreciation rate.

Table 5.11. Cross Correlation of Key Explanatory Variables

	R&D/ Assets	Patents/ Assets	Publications /Assets	Citation-weighted Publications/ Assets	Profit Before Tax
R&D	1.000***				
Patents	0.484***	1.000***			
Publications	0.692***	0.438***	1.000***		
Citation-weighted Publications/ Assets	0.313*	0.293**	0.712**	1.000***	
Profit Before Tax	0.279*	0.242*	0.312*	0.438***	1.000***

Note: ***significant at 1-percent level; ** significant at 5-percent level; * significant 10-percent level.

Appendix 5.A. Control Variables

Below is a detailed description of the construction and use of the control variables used in my analysis.

Profit Before Tax. One of the classic accounting measures of firm performance is profit before tax. Unfortunately, given the venture capital led growth of biotechnology firms, most firms in the data set did not record profits for most of their lifetimes and instead relied upon interest payments and venture capital. Despite this paucity of profit figures, exceptions to the rule exist particularly as the younger firms enter maturity and the profit before tax is recorded and denoted by PBT. This variable captures the effect of appealing effect of actual profits on the market value of biotechnology firms'. Closely allied with profits are ordinary dividends paid out by firms to their shareholders. According to the dividend discount model (Myers, 2000), a firm's historic dividends are another important signal to its future performance, and are recorded as ORDIV.

Long Term Investments. In the absence of profits, the market looks for other cues to a firm's future financial prospects. Particularly, shareholders seek an indirect measure of cash flow. In this context, a firm's long-term investments and receivables (recorded as LTM) provide a valuable signal of its current financial condition.

Equity. Models of corporate finance (Myers 2003) demonstrate that equity, as an element of a firm's capital structure may embody useful valuation information abut its relative tax

breaks, relative agency costs (of equity and debt), and risk of financial distress and growth prospects. This variable, measured by shareholders' funds, is recorded as EQUITY.

Current Assets. Finally, a firm's current assets (recorded as CURA) can provide an indication of a firm's short-term management of cash flow and thus proxies factors such as risk of financial distress, and trade-off between liquidity and returns.

Aside from the impact of financial and accounting variables, the influence of unobserved factors is controlled by using firm-specific and time-specific[45] dummies. The time-specific dummies are used to control for inflation, since the data are given in nominal terms and are not readily convertible into real terms. The inclusion of firm dummies controls for the presence of unobserved differences between firms not captured by the observed variables in the dataset (i.e., such as the presence of superior management).

[45] The use of a time-based variable (such as day, month or year) instead of time dummies is prone to some ambiguity. For example, the date at which a scientific paper is made public is noisy – it may be presented at a conference, posted electronically as a working paper, and finally published formally as a journal article. For this reason, I chose to use time dummies. I also run regressions using a year variable instead of time dummies and find no significant difference in my results (available upon request from the author).

Chapter 6

Biotechnology Patents as Credence Goods: Implications for the Patent System and Firm Behavior

6.1. Introduction

Consumers of organically produced vegetables, car mechanic services, and biotechnology patents all have something in common: even after they purchase or make use of their goods, it is often not possible to comment accurately on the quality of what they just paid for. Such goods are often labelled "credence goods" (Leibi, 2002). This chapter is concerned with the unobservable quality of biotechnology patents, or their "credence good" nature. There are varying gradations of information asymmetry associated with credence goods and I argue in this chapter that the asymmetry in information at the far end of the spectrum is an accurate model for biotechnology patents. In particular, I argue that this asymmetry in biotechnology patents is borne out of three kinds of uncertainity associated with patents – "intrinsic" uncertainty that arises out of legal complexity of the patent document, "extrinsic" uncertainty regarding the underlying technology the patent is meant to protect and lastly variable quality in issuing patents at the patent office. I examine each of these later in the paper (See Box 6.1).

Credence goods by virtue of their uncertainty require external validation in order to ascertain their quality. Previously in this thesis, I have found that biotechnology firms publish scientific articles in peer-reviewed journals in collaboration with university

scientists and that by doing so they gain financially. This behavior seems irrational in classical models of patents and technology which assume that firms reveal as little information about their R&D activities as possible (see Chapter 4, and 5). In this chapter, I argue that conceptualizing biotechnology patents as credence goods interprets this "open" behavior of biotechnology firms as rational and necessary. Firms are able to signal the quality of their knowledge stock by publishing scientific articles in a peer reviewed journals which acts as an independent source of verification (or a "credence filter").

The rest of the chapter is organized as follows: In Section 6.2, I discuss what it means for a good or service to carry credence attributes and indicate how biotechnology patents may fit this description. In Section 6.3, I examine the inadequacy of the current literature in law and economics that views patents merely as a sharp proprietary right. In particular, I analyze the manner in which uncertainty and heterogeneity in quality associated with patents affects the proprietary value of the right. I also look at some of the ways in which patents are valued and the difficult nature of this process which accounts for the virtual unobservability of the quality of the patent right. In Section 6.4, I discuss how credence mechanisms – such as peer reviewed publications and collaboration with academia (see chapters 4 and 5) -- might operate in the biotechnology marketplace. In Section 6.5, I discuss implications ~~of~~ that the credence view of patents has for firms and the patent system. Finally in Section 6.6, I conclude with suggestions for future work.

6.2. Credence Attributes and Patents

The problem of credence goods is a problem of unobservable quality, and typically occurs in medical, legal, and financial advice services as well as a wide variety of repair professions. In most of these cases the seller provides the goods or services and also acts as the expert who determines what and how much of the goods and services are required. Searching for information to verify the expert's opinion is usually expensive (prohibitively so in certain cases), and creates opportunity for fraud on the part of the expert[46].

This problem of information asymmetry between buyers and sellers has received considerable attention in economics following on from George Stigler's classic paper on the economics of information[47] (Stigler, 1961). A transaction involves an asymmetry of information when one party to the exchange has more information (on quality of goods or relative price) than the other, leading to opportunities for fraudulent behavior. Although considerable attention in the literature has been paid to the problem of asymmetric information between buyers and sellers, the theoretical literature on fraudulent experts is "fairly small" (Emons, 1997).

[46]Asymmetry of information gives sellers several opportunities to exploit consumers – a problem that is dubbed "demand inducement" in the health economics literature. Winand Emons cites several examples where fraud was covered up. In Switzerland, for example, patients with the minimum level of schooling are twice as likely to have their womb or gall stone removed than patients with a university degree. Further, children whose parents are doctors are five times less likely to have their tonsils removed as regular children (Emons, 1997).

[47] Information is a valuable resource, yet "it occupies a slum dwelling in the town of economics". So starts the classic paper written by Stigler in 1961 that precipitated an explosion of theoretical research on the economics of information (Stigler, 1961).

Stigler's (1961) classic paper dealt with the problem of ascertaining "market price". In his paper, Stigler analyzed search costs that arise when a buyer (or seller) seeking the most favorable price canvasses various sellers (or buyers). According to Stigler, a consumer searches for information until the marginal benefit of additional information is equal to the marginal costs of obtaining the additional information. As a result, there is willingness to pay for information and there is a marginal cost of information. His central contribution lay in the conclusion that important aspects of economic organization take on a fresh meaning from the viewpoint of the search for information.

Following Stigler's lead, Nelson showed that the problem of determining quality of goods and services is even more intractable than the problem of determining price (Nelson, 1970). Based on the quality level of goods and services, Nelson distinguished between – 'search goods' and 'experience goods'. One can determine the quality of 'search goods' by searching; the quality of 'experience goods' can be determined by experiencing taste, durability or maintenance needs. Also for any brand, search qualities can be determined prior to purchase and experience qualities only after the event.

Typically, it is more expensive to get information about quality than it is to get information about the price. For some low cost goods, purchasing the product may be the best way of experiencing its quality – cans of tuna for example. If the purchase price is low enough, the consumer may prefer to get his information by way of 'experience'.

However if the costs of these procedures rises sufficiently high, the consumer will try to get the information in other ways.

Darby and Karni in a classic paper expand the categories to include 'credence goods' (Darby and Karni, 1973). They discuss how reputation, market conditions, and technological factors affect the amount of fraud. Credence goods constitute a category for which the non expert cannot verify the quality attributes of the good. For these goods one must rely on a third party to provide truthful information to the consumer about quality.

This discussion often provides theoretical backing for third party certification or introduction of government regulations. For example, this is often cited as the underlying economic theory for the eco-labeling of foods[48]. Research on consumer demand for credence products can be further unpacked as consumer demand for certain attributes of the product. Kevin Lancaster's framework in this context contends that the utility of the good does not lie in the good itself but in the characteristics or attributes that create utility (Lancaster, 1971).

This theoretical background that casts the market of goods as a search for information (for quality and price) has a unique resonance for biotechnology patents. The traditional view of patents has always regarded them as classic proprietary rights over information (Stigler, 1961; Fisher, 2001). However the credence attribute of patents in biotechnology arises from uncertainty in the true extent of information enclosed within such proprietary

[48] As an illustration see Cathy Roheim 'Early Indications of Market Impacts from the Marine Stewardship Council's Ecolabeling of Food' (Dec 26, 2002), at p 13 available at http://www.wto.org/english/forums_e/ngo_e/ccc_msc_e.doc

rights. The nature of the problem at hand is made clearer with reference to two scenarios. One could have firm fencing around a few thousand acres of land without knowing whether they enclose prairies, woodlands, arable land, wetlands or mining fields. A second uncertain scenario is represented by knowing what kind of land is enclosed with informal fencing specified only with reference to certain landmarks. Both scenarios lead to uncertainty in the extent and nature of the property rights held. Although a certain amount of uncertainty in property rights is to be expected, the uncertainty associated with patents in new technology is fundamental and crucial to an accurate understanding of the role of such rights in the market.

Immature technologies have particular characteristics that make them vulnerable to the credence good problem. These include, but are not restricted to, uncertain scope of terminology used and uncertain attributes of 'the person skilled in the art' – the notional person through whom questions of fact and law are often interpreted in patent law. The problem of uncertainty of technological and commercial implications, which is a problem common to all patents, is thus greater in the case of newer technologies. The utility of a patent right is not the patent itself but the attributes and characteristics that create the utility. Therefore the credence nature of the patents allows us to focus on the processes by which the attributes are valued and priced by the market.

6.3. Transcending the exclusivity view of patents

6.3.1. The Theory of Patents

Patent law is home to a number of seemingly freestanding theories that support and rationalize the need to give property rights over inventions[49], and as William Fisher notes, 'in law reviews and journals of economics and philosophy articles deploying "theories" of intellectual property proliferate'[50]. Nonetheless or perhaps because of this, controversies over patentability and the scope of these property rights have been around a long time. Machlup and Penrose describe how 'seesaw battles' around the patent system have raged on and off since the 19th century. The chief opponents of the system down the ages have been staunch proponents of free enterprise; the debate for and against tends to follow classic lines of argument. Indeed as the authors note, 'little, if anything, has been said for or against the patent system in the twentieth century that was not said equally well in the nineteenth' (Machlup and Penrose, 1950).

Patents and the need for patents has also been subjected to all manner of economic analysis. But in the classic sense there are four types of arguments that support the granting of exclusive property rights over inventions as the most appropriate response to the need for and creation of inventions.

---

[49] These theories often give one the feeling that no amount of effort at unraveling and understanding will suffice. Although active contradictions are rare, a number of gaps and ambiguities make the theoretical understanding of intellectual property an elusive goal. With respect to economic analysis for example, Landes and Posner (2003) note that there is a tendency among economic analysts on intellectual property to reduce the entire problem of intellectual property rights to a trade-off between 'incentive' and 'access'. They argue that to reduce the problem of intellectual property to this trade-off is to oversimplify greatly; 'Not that the incentive-access trade-off is nonexistent or even unimportant; but there is much else to consider in an economic analysis of intellectual property law'. Samuel Oddi (1996) in the course of his economic analysis of patents comes to the conclusion that it is not possible to have one single unified theory of patents that can predict appropriate scope of protection. See Samuel Oddi 'Un-Unified Economic Theory of Patents' 71 Notre dame Law Review 267 (1996).

[50] William Fisher 'Theories of Intellectual Property' in Stephen Munzer, ed., *New Essays in the Legal and Political Theory of Property* (Cambridge University Press, 2001). Fisher himself canvasses many of these theories, evaluates them and considers the role they ought to play in law making.

Type 1 says that a man has natural property rights in his own ideas; appropriation by others must be condemned as stealing. Type 2 says that inventors receive rewards for their services to society. The most appropriate way for society to reward them is by way of exclusive property rights. Type 3 says industrial progress is desirable to society and without incentives, inventors and capitalists will not endeavor to make those inventions that are central to industrial progress available. Exclusive property rights are the most appropriate way to ensure that they make profitable ventures and thereby avoid the 'free riding problem'[51]. Type 4 argues that in the absence of exclusive property rights, inventions that are essential for industrial progress, will have no protection against imitation and hence inventors will continue to keep these secret. Hence it is in the interests of society to induce the inventor to disclose his secret and the most appropriate way to do this is through exclusive property rights by way of patent rights (Machlup and Penrose, 1950). Most of the strands within these broad arguments generally exhibit a striking polarity for or against exclusive property rights.

Economic analyses of the need for patents and indeed other intellectual property rights, often assumes as a starting point that property rights are necessitated by the nature of information[52]. The standard economic rationale of patents is expressed as an efficient method of enabling the benefits of research and development to be internalised, thus making innovation and technological progress possible. Landes and Posner in their

[51] Consumers of a public good who do not contribute to the costs of creation or maintenance get a 'free ride' when others pay for these costs. Classic illustrations of such public goods include the police.
[52] Kenneth Arrow for instance, resolves the 'information paradox' by assuming well defined property rights as his starting point (Arrow, 1962).

classic account of the economics of patent law focus on another useful way of thinking about patents, i.e., as a response to economic problems inherent in trade secrecy and the market structure (Landes and Posner, 2003).

One of the more interesting strands within the economic analysis of intellectual property rights is their desirability based on their function as 'information mechanisms' -- illustratively, in the case of trademarks - the function of reducing consumers 'search costs'. This is a specific approach within a broader utilitarian force shaping property rights in the legal field. Thus, the primary economic benefits of trademark, are (1) the reduction of consumers' search costs (because it is easier to pick a box of "Cheerios" off the grocery shelf than to read the list of ingredient in each container, and because consumers can rely on their prior experiences with various brands of cereal when deciding which box to buy in the future. The second part to this view is that trademarks create an incentive for businesses to produce consistently high-quality goods and services (because they know that their competitors cannot by imitating their distinctive mark, take a free ride on the consumer good that results from consistent quality)[53].

There are two distinct ideas about information within this description that are relevant to our analysis. Firstly the idea that the intellectual property right, in this case trade marks, are a proxy for the value (in terms of quality and reputation) of information. This view of intellectual property right is uncommon in the case of copyright and seldom applied with respect to patent rights. The reason being that as we move along the spectrum of intellectual property rights, from trade marks to patent rights, the right becomes less of a

[53] See Landes and Posner seminal work on the economics of trademark law (Landes and Posner, 1987).

proxy and is valued more for being a 'proprietary' right in itself that demarcates the subject matter for exclusive use. This is why patents are regarded as 'stronger' intellectual property rights than copyright or trademarks.

However, if patents are indeed credence goods, as I argue here, then what is true for trademarks is also true for patents.[54] Patent rights, as per our view, are but a proxy for the value of the information they contains. However where trademarks typically code for 'search' or 'experience', the costs of investigating the quality of the information contained within patents (in newer technology) is prohibitively expensive.

Secondly, there is the question of the 'free rider' problem. This is reflected in the analysis of all of the intellectual property rights due to the non-rivalrous nature of information. It is particularly acute in the case of patent rights, where it might preclude the commercialization of inventions. Kenneth Arrow described this as a paradox in the valuation of information that stymies the free flow of information between inventors and producers (Arrow, 1962). Typically the inventor has many ideas but few resources, and the producer has the resources but few ideas. The close relationship between the two is played out within research, development and manufacturing. The relationship is a tense one as, minus property rights, the inventor is unlikely to want to disclose his invention in full and the producer is unlikely to want to invest in an ill defined idea[55].

[54] Perhaps to a lesser or qualitatively different extent but present nontheless.
[55] Arrow states, 'The value of information for the purchaser is not known until he has the information, but then he has in effect acquired it without cost' (Arrow, 1962).

The key to resolving Arrow's paradox then is treating patents as well defined property rights regardless of whether the invention is comparatively simple or complex. And such a view has allowed economists to focus on complex relationships among patents, innovation, competition and the diffusion of technology. Treating patents in biotechnology as credence goods means that Arrow's paradox is not resolved merely by the granting of patents. Because of the inherent uncertainty with patents it then becomes necessary to investigate how value is attributed to the information contained within the patent. In the next section, I more fully account for the sources of uncertainty and poor quality in patents.

6.3.2. Accounting for Uncertainty and Variable Quality in Patents

Several scholars have turned their attention to analyzing patents as ill-defined rather than the sharply defined exclusionary rights they were once thought to be (Long, 2002; Lemley and Shapiro, 2004)). In the literature, there are essentially three separate sources of uncertainty associated with patents which I classify as intrinsic uncertainty, extrinsic uncertainty and the uncertainty related to issues of quality (see Box 1). I examine each of these in greater detail later in the section.

6.3.2.1 Intrinsic Uncertainty

When a patent is granted, an extensive public document is created with a wealth of technological and firm information The first source which I call intrinsic uncertainty arises from the very document of the patent. Patents are complex scientific and legal

documents. Understanding the content of patents is a highly skilled task the difficulties of which are usually exacerbated in the case of new or immature technologies due to unsettled technical terminology and evolving jargon.

Box 6.1. Sources of uncertainty in biotechnology patents

Instrinsic Uncertainty – A patent application is a dense legal and scientific document where terminology acquires meanings based on state of the art and practice among persons skilled in the art. Language often bears the burden of immature technology that has to be interpreted in a specific to the legal context. To illustrate a recent case in the UK House of Lords involved fixing the meaning of 'host' cell. The essential question was whether the term 'host cell' commonly understood to refer to a cell that 'hosts' endogamous genetic material may include a cell with activated endogenous genetic material instead within its meaning.

Extrinsic Uncertainty – It is notoriously hard to predict the technological and hence commercial worth of a patent. Factors external to the patent such as cost and choice of materials available, cost and availability of complementary technologies, and the present or potential demand or market for a particular product or technology are decisive. Because of these factors patent counts alone are an extremely volatile indicator of a firm's innovative success.

Variable Quality – The patent systems in both the US and Europe have found it difficult to deal with rising numbers of patent applications (Lemley and Shapiro, 2004). Patent examiners have often been criticized as being the last to know of technological developments (Merges, 1999). Recently it was reported that as per staff surveys examiners at the EPO are losing confidence in their ability to ensure the quality of the patents issued (Abbott, 2004). Institutional design and interest group politics ensure a steady expansion of subject matter eligible for patent protection increasing the pressure on patent offices. Anecdotal evidence of 'bad' or 'poor quality' patents have contributed to a general mistrust of the technological or commercial worth of the information enclosed in a patent. The problem of 'poor quality' can be particularly acute in the case of immature technologies as patent examiners may have to rethink examination techniques quickly.

Source: Authors original contribution.

Moreover, new technology creates a period of doctrinal uncertainty that can colour the way the industry regards such rights. In the case of interpreting biotechnology patents, an important clause is that of non-obviousness and this is determined by the "notional person skilled in the art" which is a central concept in the law of patents and has been notoriously difficult to fix in the case of biotechnology both in Europe and the US (Cornish and Llewellyn, 2003). This notional person (or PHOSITA – "person having ordinary skill in the art") determines important claims such as obviousness and sufficiency of disclosure when granting a patent. The level of the skill in the art and the judgments the court makes about ordinary skill in the industry profoundly affects the scope of the patents that issue. On a macro level it can take a few years for the uncertainty of fixing notions upon which decisions are based to get converted to a more 'closed' form of interpretation. On a micro level, this means patents are often of uncertain validity and scope.

6.3.2.2. Extrinsic Uncertainty

A second source of uncertainty, which I call *extrinsic uncertainty*, is caused by how patent value is perceived in capital and labor markets. Among other things, such markets have a compelling need for such information in order to value firms and the assets they hold, to employ 'productive scientific groups', and make investment decisions. Often a thorough investigation directed towards intellectual property is called for in business transactions involving biotechnology firms as they rely so heavily on proprietary information[56]. The basis of the credence theory of patents is that this information is

[56] The wide ranging nature of such investigations is noted in a two part article by Gogoris and Clarke (2002a, 2002b) A due diligence investigation with respect to IP is called for typically when a company is

extremely hard to obtain in a credible way. Patents contain information in varying amounts and qualities, a result of an attribute of knowledge that Clarisa Long refers to as 'lumpy' (Long, 2002) . Patents can differ enormously in the value of the information they contain and hence patent counts are not in themselves proxies for the value of underlying inventions. This is borne out by extensive work on the relationship between patents and market value. It is the extremely skewed nature of the value distribution of individual patents (that is some are very valuable, while many are worth almost nothing) that makes firm patent totals a very noisy indicator of the underlying economic value of the innovations (Hall, Jaffe and Trajtenberg, 200; Harhoff, Scherer and Vopel, 2003).

The biggest problem faced in valuing patents, which makes the intrinsic uncertainty described above qualitatively different from extrinsic uncertainty, is the persistent inability to quantify the effect of novelty, inventive step, disclosure and breadth on a patent's economic value. Often the literature centres around parameters such as the number of times a patent is cited (as an indicator of its quality), the length of its renewal (renewing a patent costs money and therefore patent renewals are a proxy for probable value), or the number of countries where it is taken (patent breadth) [57]. Further, the methods outlined above are used by industrial economists and cannot easily be turned into handy predictors of patent value for an individual case[58]. Thus, we see that patent value hinges upon characteristics such as novelty and market potential, which are hard to quantify *ex ante*, thus leading to a source of extrinsic uncertainty in valuing patents.

---

about merge with, acquire or invest in another company, business or technology. A diligence team composed of at least one lawyer, one financial expert, one technology expert and a marketing professional conducts a thorough review of assets and liabilities.

[57]For a more in-depth exploration, see Markus Reitzig (2003)

[58] See for example, www.PatentValuePredictor.com

6.3.2.3. Uncertain Quality in Issuing Patents

A third source of uncertainty in patent quality stems from actions taken in the patent office. Recent empirical work shows that patent office examinations are increasingly recognized to be meaningless guarantors of the quality of the underlying innovation[59]. Sometimes this is due to the lack of resources. More worryingly, it is also seen to reflect a change in objectives of patent offices in many countries. As Lemley reports in the US, the patent office 'reengineered' itself, declaring its mission to be 'to help our customers get patents' (Lemley, 2001). This, as he says

> 'is an indefensible position for a quasi judicial administrative agency that is trusted with representing the public interest in deciding whether to issue patents. While the job of the PTO is certainly to issue good patents it is also to reject bad ones. The idea that applicants, rather than the public at large, are the intended beneficiaries of the patent system, cannot help but contribute to the push to issue patents regardless of quality' (Lemley, 2001).

This situation is aggravated by evidence of the seemingly systematic failings of patent offices. The US patent office (USPTO) for example reported that a patent examiner in the US spends 18 hours on average reading the application, searching for and reading prior

[59]Patent offices are also not obliged to evaluate a patent for the kind of information that capital markets would find interesting or valuable. There is often a gap between assessing an invention for patentability (technological significance) and assessing an invention for value (commercial importance) in the capital or labor market. Thus there are many perfectly patentable inventions that are never commercialized. This is one reason why only 1.5 % of patents are ever litigated and only 0.1 % are ever litigated to trial (Lemley and Shapiro, 2004).

art, writing one or more provisional rejections, reviewing responses and amendments, often conducting an interview with the applicant's attorney and writing a notice of allowance[60]. Against this backdrop there are constant demands to increase productivity often ensuing from the patent office itself. The 2004 USPTO Annual report sets itself the goal of accelerated processing timing through 'more focussed examination'[61]. Patent quality problems have also been experienced in the European Patent Office. Recently it was reported that as per staff surveys examiners at the EPO are losing confidence in its ability to ensure the quality of the patents that it issues. It is a devastating indictment to have two thirds of the 1,300 patent examiners to state that productivity demands within the EPO did not allow them 'to enforce the quality standards set by the European Patent Convention'[62].

Clearly the effect of performance reports like these adds strength to the perception of 'poor quality' patent rights with considerable implication for the system as a whole as well as the way the market values these rights. This information has to be internalised by the market which will either turn away from patents as a means to 'protect' inventions, or look for third party verifiers of the quality of a patent. The growing volume of patent applications indicates that the former has not taken place leading us once again to a 'credence' model of patents.

---

[60] The requirement of substantial examination of a patent, is a relatively recent addition to a long history of granting exclusive privileges to exploit new inventions. The examination process tries to assess the quality of an invention and passes a verdict on the 'patentability' of an invention. This verdict can be further litigated and even revoked in a court of law that examines, among other things the same evidence from prior art that the patent examiner has used.

[61] http://www.uspto.gov/web/offices/com/annual/2004/0402_performance.html

[62] The survey also noted that 90% of the patent examiners did not have time to keep up to date with advances in their scientific field (Abbott, 2004).

6.4. How Credence Works in the Biotechnology Marketplace

In this chapter, I have built an argument for biotechnology patents as credence goods, by showing that biotechnology patents contain a strong degree of uncertainty associated with their economic value, and hence there exists a corresponding asymmetry of information between buyers and sellers in the biotechnology marketplace. However, patenting continues unabated. Therefore it is likely, that like in the case of other credence goods, the market has developed its own mechanisms to mitigate the information asymmetries that exist in the case of biotech.

Taking cues from the literature on credence goods, I feel it would be right to characterize most patent owners as credence good monopolists as no patent can be replaced by another patent[63]. Further most patent transactions are one-shot relationships where the threat of future reputational sanctions may not have the expected effect[64].

Winand Emons has presented a simple framework which allows one to identify conditions under which the market mechanism can bridge informational asymmetries in the case of credence goods (Emons, 2001). In most cases of credence good monopolists, according to his model, the market mechanism does a fairly good job of mitigating the

[63] It may be replaced on supplemented in some cases by lead mover advantage or trade secrecy – this will also depend on whether the invention involved is a product or a process for example.

[64] Clarisa Long recognizes the significant reputational threat a firm faces if the information it sends out is inaccurate. Long believes that a patent itself is an investment in reputation that the firm makes. With patents some degree of formal sanctions exists for is a patentee is found to have misrepresented information in his patent application the patent will be invalidated ex post. Given the imbalances of patent litigation it is not clear that the threat of invalidation and consequent reputational loss may not necessarily address the information asymmetry between buyers and sellers in a patent transaction (Long, 2002).

information asymmetry of goods and services of credence quality. The market does so by creating incentives for behavior in 'good faith' and separates the 'expert' function into 'statement' and 'verification' (Emons, 2001). Applying this model to biotechnology patents, the statement made by the patentee-expert in his patent application has to be verified by third parties for its value to be recognized by the market.

One obvious mechanism is the patent office itself. Its rejection rate is an indiocator of quality of published patents. One other (less obvious) mechanism of third party verification is provided by the scientific peer review system. Firms regularly publish the results of their scientific research in peer-reviewed journals. Publishing peer-reviewed articles allows firms to convince investors and potential collaborators of the worth of their ideas. In my previous research, I found that these gains are indeed realized by biotechnology firms and that on average publishing fourteen scientific papers has the same effect on market value as obtaining a single patent (see Chapter 5). Further, viewing biotechnology patents as credence goods explains why biotechnology firms need to collaborate with academic institutions. In Chapter 4, I found that scientific papers that biotechnology firms co-author with academic institutions are more highly cited, and appear in more prestigious journals. In particular, the greatest gains are to be had with collaborating with academics with established research reputations (or "research stars" in the vocabulary of Zucker and Darby (1995)). Collaborating with academics give biotechnology firms the crucial "credence value" necessary to convince investors and other interested third parties such as pharmaceutical firms that the knowledge that they have is robust and valuable.

In Emons' model of a credence good market, consumers attempt to infer the quality from capacity alone and/or prices (Emons, 2001). Given that we cannot effectively 'price' patents, capacity could become a big factor in influencing decisions to buy. One such mechanism of adding capacity is that of patent portfolios. A recent paper by Parchomovsky and Wagner throws considerable light on why single patents can derive value from being part of a group of patents that are commonly controlled (Parchomovsky and Wagner, 2005). The authors propose a 'new theory of patent value', and argue that firms will typically seek to obtain a large quantity of related patents, rather than evaluating the actual worth of individual patents. This is a more accurate description of the reality of patenting behavior[65] and the explanatory power of the theory is further increased when viewed via the credence good aspects emphasised here.

In particular, Parchmovsky and Wagner (2005) argue that 'by distributing the importance of the total portfolio across constituent individual patents, a patent portfolio allows holders to significantly hedge against aspects of risk and uncertainty that are endemic to innovation in the modern economy' (p 35, Parchmovsky and Wagner, 2005). Thus a large enough portfolio will address uncertainty related to future market conditions (not just technology but changing cost or availability of materials for example). A large patent portfolio also addresses uncertainty related to future competitors. This seem to square with some of the concerns related to extrinsic uncertainty. Furthermore, a healthy patent portfolio can also address uncertainty in patent law (intrinsic uncertainty). That is,

[65] See Donna M Praiss (2001) arguing that 'a strong patent position is not only an important goal for a successful biotechnology business, but also the primary asset by which a company will be valued during all stages of its development'.

because no single individual patent conclusively determines the value of a portfolio, any uncertainty in the law that could alter the value of individual patents will have less impact.[66].

Thus, we see that patent portfolios function as a 'third tier information mechanism' (Long, 2002) or what I call a 'credence verifying' mechanism. On the one hand the market, based on the evidence, distrusts the quality of patents being granted and may be unwilling (or incapable) of an opinion on the long term viability of any firm based on individual granted patents. On the other hand, credibility cumulates over a patent portfolio and adds to the reputation of the firm, much like the stock of publications adds to the market value of a firm's intellectual property (see Chapter 4).

6.5. Implications for Firms and the Patent System

For firms, my research shows that simply owning patents is not enough. The patent holder also needs to signal to investors that the information enclosed with the patent is of credible value. This is not easy, particularly in relatively new industries with immature technologies. A metaphor used earlier in the chapter is particularly apt in the case of biotechnology patents. Taking the case of an enclosed piece of land, we find that there are two kinds of uncertainty associated with its property right and value – the quality of the land enclosed and uncertainty associated with the boundary of the land. Both kinds of

[66] Among other works the authors cite to support this is a previous paper by Wagner and Petherbridge that shows that determination of claim construction issues is highly variable, and dependent upon the identity of the judge hearing the case (Wagner and Petherbridge, 1999).

uncertainty exist (quality of the patent itself and the property right that it supposedly represents) with biotechnology patents (see Box 1).

Fortunately, publishing papers in peer-reviewed journals is one way uncertainity can be mitigated. Publication can offer a credible signal of the quality of the patent because the peer-review system acts as an external verifier of the knowledge owned by a firm. Thus, firms not only have to patent but also publish and make public otherwise proprietary knowledge. Particularly, I found that firms are able to produce better quality research output (scientific papers) as measured by citation counts and journal prestige by collaborating with academics established in well regarded research institutions. Also, bundling a patent within a portfolio of patents ameliorates uncertainty associated with the technological and legal boundaries of a single patent by diversification and strategic patenting.

For policy makers, it is important to note that the market for biotechnology patents is supported by both public (peer review publications) and private mechanisms (patent portfolios). Thus, ultimately, we find that public and private incentives are deeply intertwined in biotechnology innovation. This is why fostering a strong scientific establishment with a clear and transparent system of peer review is an essential ingredient in creating and maintaining a successful biotechnology industry.

Also, poor patent quality continues to be a problem. While the market has evolved mechanisms to screen out low quality patents (by disciplining the fraudulent patentee),

poor quality patents generate wasteful transaction costs in the economy. Merges (1999) suggest greater administrative and patent office review as a means of identifying 'bad' patents. There seems to be increasing awareness of the need to maintain the 'quality' filter role that patents play. The US for example is poised to adopt a 'European' style post-grant opposition procedure. Three recently issued reports in the US from the USPTO (United States Patent Office), the Federal Trade Commission and the National Academy of Sciences have all called for a post grant review process to provide a forum for validity challenges. They have all recognized the damage that questionable patents can cause and the limited options available in the existing system (Apple, 2005). There are considerable doubts as to how effective this procedure has been within the European system, but a number of recent high profile oppositions filed against controversial patents seem to have increased general confidence in this system (OECD, 2002).

There are also clearly markets and market conditions under which *ex ante* information gathering does not solve the potential problem of fraud. Emons (2001) refers to cases where prices are set by a regulator rather than by the seller, 'insurers pay for the services, distorting consumers' incentives to gather and process the necessary information'. This seems to indicate that greater third party regulation of the 'quality' of patents or the mechanisms that identify the 'quality' of patents would lead to a reduced incentive on the part of 'buyers' to decrease the informational asymmetry. Hence measures such as introduction of a post grant review process in the US may fail fundamentally to decrease the information asymmetry and further distort the process of gathering information about the *quality* of patents. Thus, further research is needed to precisely quantify the

transaction costs generated by low quality patents on the economy and the effect greater investments by the patent office to increase quality would have on the biotechnology marketplace.

Despite these concerns, clearly my own conclusions echoe those of Emons (1997, 2001) who observes that a more thorough understanding of these markets will be helpful for public policy purposes. Clearly studying how the market ascertains the value of patents has profound implications to the theory and reality of patenting behavior; and the work done here presents a step in this direction.

6.6.Conclusion

In this chapter, I have shown that biotechnology patents can be treated as credence goods insofar as the market for biotechnology patents is characterized by an asymmetry of information between buyers and sellers of publicly issued shares. This informational asymmetry is a result of uncertainty associated with biotechnology patents. The chief causes of uncertainty reside with the legal substance of the patent document itself, the technological and commercial uncertainty associated with patent value and poor quality in screening new innovations at the patent office when granting patents (see Box 1).

Despite these limitations, firms continue to patent in increasing numbers. Thus, the market has evolved mechanisms to "discipline" the fraudulent patentee and ascertain more accurately "true" patent value. These mechanisms, that I label *credence verifiers*,

include collaborating with academic scientists (private-public collaboration seen in Chapter 4), publishing scientific papers in peer reviewed journals and the practice of clubbing patents in patent portfolios. My research also suggests that despite these market mechanisms to verify value, the patent office needs to invest in screening out low quality patents that generate wasteful transaction costs in the economy.

It is hoped that the analysis presented in this chapter will stimulate further research into the credence good nature of biotechnology patents. The framework used in this paper was intended to extend our understanding of the incentives that drive innovation and the creation of intellectual property. Still more work is needed and shall be pursued by this author in subsequent research.

Conclusion

Following the opening three chapters which survey the literature related to knowledge, economic growth and firm behavior, I focus on the biotechnology industry to understand how firms translate basic scientific ideas into profitable ventures. I found that this industry is characterized by two unique stylized facts: first, firms publish the results of their scientific research openly in peer reviewed journals, and two, they collaborate with universities quite intensively. I explored this private-public nature of biotechnology innovation in three separate papers.

In my first paper (Chapter 4), I found that collaborative research with academics improves research quality for biotechnology firms. My results indicated that biotechnology firms should seek alliances with high status academics with established research reputations. One of the major policy implications of my paper was the recommendation to strengthen support for public science and encourage collaboration across the private-public divide.

The results of my first paper posed another question – namely, why should firms publish the results of their research openly in the first place? I address this question in my second paper (Chapter 5), by developing an open-science framework of innovation which argues that while R&D expenditures reveal the commitment of a firm's resources to innovation and patents record the completion of R&D activity, a firm's stock of scientific papers signals the quality of its innovative efforts. In biotechnology, quality of research is a

valuable signal and publishing peer-reviewed articles allows firms to convince investors and potential collaborators of the worth of their ideas. This proposition is tested using unique data of U.K. biotechnology firms during the years 1988-2000. The findings indicate that research publications bring real financial gains to biotechnology firms and that, on average, publishing fourteen scientific papers in academic journals has approximately the same impact on a firm's market value as obtaining a single patent. Furthermore, papers which are highly cited, particularly by pharmaceutical firms, have a greater impact on market value.

In a third theoretical paper (Chapter 6), I show that biotechnology patents can be treated as credence goods insofar as the market for biotechnology patents is characterized by an asymmetry of information between buyers and sellers. This informational asymmetry is a result of uncertainty associated with biotechnology patents. The chief causes of uncertainty are the legal substance of the patent document itself, the technological and commercial uncertainty associated with patent value and variable quality in screening new innovations at the patent office when granting patents. Despite these limitations, firms continue to patent in increasing numbers. Thus, the market has evolved mechanisms to more accurately ascertain "true" patent value. These mechanisms, that I label, credence verifiers, include publishing scientific papers in peer reviewed journals and the practice of clubbing patents in patent portfolios. Studying how the market ascertains the value of patents has implications for the theory and reality of patenting behavior; and by conceptualizing biotechnology patents as credence goods, this paper

makes an interdisciplinary contribution (combining law and economics) to understanding the incentives that drive innovation.

In this thesis, I have made original contributions on three levels: methodological, empirical and theoretical. I list these below.

Methodological contribution

[1] In this thesis, I have constructed an original database of firm performance and an extensive array of variables that capture innovative capabilities of U.K. biotechnology firms between 1988-2001. This is the first time such an extensive dataset has been collected on innovation and firm performance of the U.K. biotechnology industry.

[2] The central methodological contribution of this thesis is to introduce bibliometric measures of scientific knowledge in the study of innovation. As noted earlier (Chapter 3 and Chapter 5), a study of traditional measures of innovation such as R&D and patents which have given us a great deal of knowledge about innovation have several inadequacies in usage. In particular both R&D and patents (as well as weighted measures of patent data) are not accurate measures of quality of scientific innovation. In this thesis, I have argued that bibliometric measures of innovation act as a signal of quality and thus provide a more complete picture of the innovation process.

[3] In Chapter 5, I use a firm's financial performance as a proxy of its success following a strong empirical tradition started by Griliches et al (see Chapter 5). I have extended this study of innovation and performance by including bibliometric measures (publications weighted by citations, publications weighted by journal prestige) in a market value equation for the first time.

Empirical Contribution

Chapter 4. Do firms produce better quality research with greater academic collaboration?

[1] While previous research has focused on the importance of private-public partnership for success in biotechnology innovation, my research is the first empirical research to study the effect of private firms collaborating with public institutions on research quality.

[2] I conclusively found that research quality of private-public research output (as measured by scientific publications weighted by citations or journal prestige) is improved by increased academic input. I find clear evidence for this in the data that is robust to the inclusion of control variables, firm fixed effects and different model specifications. To some extent, my results mirror studies in the US in which star scientists have been found to impact firm research productivity (Zucker and Darby, 1995). In particular I confirmed the following empirical facts that characterize the effect of private-public collaboration on research quality:

a. The greatest impact of collaboration is felt in "basic" research although the results hold across the "basic-applied" spectrum.

b. Further, the research indicates that these quality gains are accentuated by collaborating with more than one laboratory and across disciplines.

c. The strongest gains are to be had in collaborating with academics with established research reputations ("star scientists", Zucker and Darby (1995)).

Chapter 5. Why Do Biotechnology Firms Make Private Knowledge Public?

[3] The key original empirical contribution of this chapter is to include citation-weighted publications in a market value equation of biotechnology assets and to analyze the strategic import of this variable. To the extent of knowledge of this author, this is the first time such an analysis has been carried out.

[4] I found that, on average, publishing fourteen scientific papers in academic journals has approximately the same impact on market value as obtaining a patent.
Further,

a. Highly cited papers have a much greater impact than average on market value.
I find that firms that average more than the median number of citations per patent exhibit a very significant increase in market value: 10% higher if having 7-10 citations per patent, and 35% higher if having 2-3 times the median (11-20 citations per patent). The most dramatic effect are for those scientific papers that receive more than 20 cites per

patent: the market value of these firms is 54% higher than if the firm's paper output had generated 4 or fewer citations.

b. In terms of market value, a pharmaceutical citation is worth nearly ten times more than a regular citation.

Theoretical contribution

[1] The central contribution of this thesis is to establish that public-private incentives drive innovation in biotechnology research. While a growing body of evidence (see Chapter 3 and 4) point to the importance of the role played by non-private institutions in high technology innovation, this thesis presents a significant step in more fully integrating public as well as private incentives in a single innovation framework.

[2] In Chapter 5, I develop a single public-private framework that integrates citation-weighted publications with more traditional measures of innovation such as R&D and patents. I argue that while R&D expenditures reveal the commitment of a firm's resources to innovation and patents record the completion of R&D activity, firms' publication of scientific papers signals the quality of its innovative efforts. To the extent of knowledge of this author, this is the first time such a framework has been developed and tested.

[3] The question of why peer-reviewed publications acts as an accurate signal of quality is posed by my thesis. I argue that biotechnology patents can be treated as credence goods

insofar as the market for biotechnology patents is characterized by an asymmetry of information between buyers and sellers. Collaborating with academics (Chapter 4) and publishing scientific papers in peer-reviewed journals (Chapter 5) bridge this informational asymmetry because the scientific community and the peer-review process acts as a "credence verifier" in ascertaining value in the biotechnology marketplace. The credence good framework which I develop (Chapter 6) integrates perspectives from both law and economics is a new one and provides a strong theoretical basis for understanding the incentives that drive the behavior of biotechnology firms. The standard reference for the use of credence good models is Emons (1997) in which Emons demonstrates how the market for medical services is characterized by credence properties. There is an asymmetry of information between buyers and sellers in this market which require third parties, usually government accreditation to bridge the gap. While, models such as these have been applied in other markets (accountants, university education for example), this framework has not been applied in the context of knowledge signalling in science-based firms (such as biotechnology firms).

In my thesis, I show that the stock of scientific papers published by biotechnology firms' signal the worth of their knowledge to financial markets.

This has implications for both firm behavior and public policy. For firms, it validates the use of an "open-science" framework thereby encouraging firms to share the results of their research with a peer scientific community. From a policy perspective, it strengthens the need for a strong independent peer-review system separate from industry. The stock

of scientific papers act as credible signal of knowledge, insofar as the peer-review system remains an impartial and independent judge of content. Thus, to maintain this independence and autonomy is crucial for the smooth functioning of the biotechnology industry.

In summary, this thesis presents a view of biotechnology innovation which integrates both public as well as private incentives and develops the literature along this basis on both empirical and theoretical fronts. I intend to further this research on the incentives and outcomes of public-private innovation in my future work.

*

References

[1] Abbott, A., 2004. Pressured Staff 'lose faith' in Patent Quality. Nature, Volume 429, p 423.

[2] Abramowitz, M., 1956. Resource and Output Trends in the United States since 1870. American Economic Review, 46, 5-23.

[3] Acs, Z. and Audretsch, D., 1990. Innovation and small firms. Cambridge MIT Press

[4] Acs, Z. J., Audretsch, D.B., and Feldman, M.P., 1992. Real Effects of Academic Research: Comment. American Economic Review, 82: 363-367.

[5] Adams, J., 1990. Fundamental Stocks of Knowledge and Productivity Growth. Journal of Political Economy 98(4), 673-702.

[6] Aghion, P., Howitt, P., 1992. A Model of Growth through Creative Destruction. Econometrica, 60, 323-351.

[7] Alcacer, J., Gittelman, M., 2004. How do I Know what you Know? Patent Examiners and the Generation of Patent Citations. http://ssrn.com/abstract=548003.

[8] Almeida, P., Kogut, B., 1999. Localization of Knowledge and the Mobility of Engineers in Regional Networks. Management Science, 45(7), 905-917.

[9] Apple, T., 2005. The coming US Patent Opposition. Nature Biotechnology, Volume 23, pp 245-247.

[10] Arora, A., David, P.A., Gambardella, A., 1995. Increasing Returns to Knowledge Capital in Scientific Research. Paper presented at NBER Summer Workshop on Industrial Organization, Cambridge, Mass.

[11] Arrow, K., 1962. Economic Welfare and the Allocation of Resources for Invention. In Nelson, R. (ed.) The Rate and Direction of Inventive Activity. Princeton University Press, 609-619.

[12] Azariadis, C., Drazen, A., 1990. Threshold Externalities in Economic Development. Quarterly Journal of Economics, 55(2), 501-526.

[13] Bain, J.S., 1954. Economies of Scale, Concentration, and the Condition of Entry in Twenty Manufacturing Industries. American Economic Review, 44, 15-39.

[14] Bain, J.S., 1968. Industrial Organization. Wiley & Sons.

[15] Barney, J., 1991. Firm Resources and Sustained Competitive Advantage. Journal of Management, 17, 99-120.

[16] Barro, R.J., 1989. Economic Growth in a Cross Section of Countries. National Bureau of Economic Research.

[17] Bator, F.M., 1958. The Anatomy of Market Failure. The Quarterly Journal of Economics, 8, 351-379.

[18] Baumol, W., Blackman, S., Wolff, E., 1989. Productivity and American Leadership. MIT Press.

[19] Baumol, W., Panzar, J.C., Willig, R.D., 1982. Contestable Markets and The Theory of Industrial Structure. Harcourt Brace Jovanovich, Inc.

[20] Ben-Zion, U., 1984. The R&D Investment Decision and its Relationship to the Firm's Market Value: Some Preliminary Results. R&D Patents and Productivity (ed. Griliches, Z.) Chicago University Press, 299-312.

[21] Blumenthal, M.D., Causino, N., Campbell, E.G., Louis K.S., 1996. Participation of life-science faculty in research relationships with industry, New England Journal of Medicine, 335(23), 1734-1739.

[22] Blundell, R., Griffith, R., Van Reenen, J., 1999. Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms. Review of Economic Studies, 66, 529-544.

[23] Bosworth, D.L., Mahdian, H., 1999. Returns to Intellectual Property in the Pharmaceuticals Sector. Economic Appliquée, LII (2), 29-93.

[24] Bosworth, D.L., Rogers, M., 1998. Research and Development, Intangible Assets and the Performance of Large Australian Companies. Melbourne Institute Working Paper 2/98.

[25] Bresnahan, T.F., 1986. Measuring Spillovers from Technical Advance: Mainframe Computers in Financial Services. American Economic Review, 76(4), 742-755.

[26] Brown, J.S., Duguid, P., 2001. Knowledge and Organization: A Social-Practice Perspective. Organization Science, 12(2), 198-213.

[27] Campbell, R.S., Nieves, A.L., 1979. Technology indicators based on patent data: The case of catalytic converters: Phase 1 report: Design and demonstration. Pacific Northwest Laboratories, Batelle.

[28] Carayol, N., 2003. Objectives, agreements and matching in science-industry collaborations: reassembling the pieces of the puzzle, Research Policy, 32, 887-908.

[29] Caves, R.E., Porter, M.E., 1977. From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition. Quarterly Journal of Economics, 91, 241-262.

[30] Chan L. K. C., Lakonishok J., Sougiannis, T., 1999. The Stock Market Valuation of Research and Development Expenditures. National Bureau of Economic Research Working Paper 7223: Cambridge, MA

[31] Chamberlin, E.H., 1933. The Theory of Monopolistic Competition. Harvard University Press.

[32] Chatterjee, S., Wernerfelt, B., 1991. The link between resources and type of diversification: theory and evidence. Strategic Management Journal, 12, 33–48.

[33] Chauvin, K.W., Hirschey, M., 1993. Advertising, R&D expenditures and the market value of the firm. Financial Management, 22, 128-140.

[34] Chesbrough, H., 2003. Open Innovation. Harvard Business School Press, Boston.

[35] Coase, R., 1937. The Nature of the Firm, Economica, IV, 386-405.

[36] Cockburn, I., Griliches, Z., 1988. Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents. American Economic Review, 78(2), 419-23.

[37] Cockburn, I., Henderson, R.M., 1994. Racing to Invest?: The Dynamics of Competition in Ethical Drug Discovery, Journal of Economics and Management Strategy, 3(3), 481-519.

[38] Cockburn, I., Henderson, R.M., 1996. Public-private Interaction in Pharmaceutical Research. Proceedings of National Academy of Sciences, 93, 12725-12730.

[39] Cockburn, I., Henderson, R.M., 1998. Absorptive capacity, Coauthoring behavior, and the Organization of Research in Drug Discovery. The Journal of Industrial Economics, June.

[40] Cohen, W., Levinthal, D., 1990. Absorptive Capacity – A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35, 128-152.

[41] Combs, J.G., Kitchen, D.J., 1999. Explaining Interfirm Cooperation and Performance: Toward a Reconciliation of Predictions from the Resource-Based View and Organizational Economics. Strategic Management Journal, 20, 867-888.

[42] Comroe, J., Dripps, R., 1976. Scientific Basis for the Support of Biomedical Research. Science, 192, 105.

[43] Cornish and Llewellyn, 2003. Intellectual Property: Patents, Copyright, Trade Marks and Allied Rights. Sweet and Maxwell, 5th edition, pp 229-230.

[44] Crane, D., 1972. Invisible College: Diffusion of Knowledge in Scientific Communities. University of Chicago Press, Chicago.

[45] Cyert, R.M., March, J.G., 1963. A Behavioral Theory of the Firm. Prentice Hall, Englewood Cliffs.

[46] Darby, M., and Karni, E., 1973. Free Competition and the Optimal Amount of Fraud. Journal of Law and Economics, Volume 16, pp 67-88.

[47] Dasgupta, P., David., P.A., 1994. Towards a New Economics of Science. Research Policy, 23, 487-521.

[48] David, P., Mowery, D., Steinmueller, E., 1992. Analyzing the Economic Payoffs from Basic Research. Economics of Innovation and New Technologies, 2(4), 73-90.

[49] De Solla Price, D.J., 1965. Networks of Scientific Papers. Science, 149, 510-515.

[50] Decarolis, D.M., Leeds, D.L., 1999. The Impact of Stocks and Flows of Organizational Knowledge on Firm Performance: An Empirical Investigation of the Biotechnology Industry. Strategic Management Journal, 20, 953-968.

[51] Del Canto, J.G., Gonzalez, I.S., 1999. A resource-based analysis of the factors determining a firm's R&D activities. Research Policy, 28, 891-905.

[52] Demsetz, H., 1974. Industry Structure, Market Rivalry, and Public Policy. Journal of Law and Economics, 16, 1-10.

[53] Demsetz, H., 1982. Barriers to Entry. American Economic Review, 72(1), 47-57.

[54] Department of Trade and Industry (DTI), 2005. Comparative Statistics for the UK, European and US Biotechnology Sector – Analysis Year 2003.

[55] Dierickx, I., Cool, K., 1989. Asset Stock Accumulation and Sustainability of Competitive Advantage. Management Science, 35, 1504-1514.

[56] Dosi, G., Teece, D.J., Winter, S., 1992. Toward a Theory of Corporate Coherence: Preliminary Remarks. In Dosi, G., Giannetti, R., Toninelli, P.A., 1993. Technology and Entreprise in A Historical Perspective. The Clarendon Press of Oxford University, Oxford.

[57] Emons, W., 1997. Credence Goods and Fraudulent Experts. RAND Journal of Economics, Volume 28, No: 1, pp 107-119 at p 109.

[58] Emons, W., 2001. Credence Goods Monopolists. International Journal of Industrial Organization, Volume 19, 375-389 at p 375.

[59] E. von Hippel 1987. Cooperation between rivals: Informal know–how trading. *Research Policy,* volume 16, issue 6, pp.291–302.

[60] E. von Hippel, 2001. Innovation by users communities: Learning from open-source software. *MIT Sloan Management Review,* volume 42, issue 4, pp. 82–86.

[61] Fama, E.F., Jensen, M.C., 1983. Separation of ownership and control. Journal of Law and Economics, 26, 301-325.

[62] Financial Times, *Putting U.K. Biotech in Perspective*, 23/08/2005.

[63] Fisher, F., McGowan, J., 1983. On the Misuse of Accounting Rates of Return to Infer Monopoly Profits, American Economic Review, 73, 82-97.

[64] Fisher, W., 2001. Theories of Intellectual Property in Munzer, S. ed., New Essays in the Legal and Political Theory of Property, Cambridge University Press, 2001.

[65] Garud, R., Rappa, M.A., 1994. A socio-cognitive model of technology evolution: The case of cochlear implants. Organization Science, 5, 344-362.

[66] Geroski, P.A., 1994. Market Structure, Corporate Performance and Innovative Activity. Oxford University Press, Oxford.

[67] Geroski, P.A., Machin, S., Van Reenan, J., 1993. The profitability of innovating firms. Rand Journal of Economics, 24(2), 198-211.

[68] Gittelman, M., Kogut, B., 2003. Does Good Science Lead to Valuable Knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science, 49(4), 336.

[69] Gogoris, A. and Clarke, P., 2001. Patent Due Diligence in Biotechnology Transactions. Nature Biotechnology Volume 19, No: 2 pp 175-177 (February 2001) and Volume 19, No. 3, pp 279-281.

[70] Grant, J., & Lewison, G., 1997. Government funding of research and development. Science 278, 878-879.

[71] Greene, W.H., 1981. Sample Selection Bias as a Specification Error: Comment, Econometrica, 49, 795-798.

[72] Greene, W.H., 1997. Econometric Analysis. Prentice Hall, Upper Saddle River, NJ.

[73] Griliches Z., 1981. Market Value, R&D and Patents. Economic Letters, 7: 183-187.

[74] Griliches, Z., 1973. Research Expenditures and Growth Accounting. In Williams, B.R. (ed.) Science and Technology in Economic Growth. Macmillan.

[75] Griliches, Z., 1979. Issues in Assessing the Contribution of Research and Development to Productivity Growth. Bell Journal of Economics, 10(1), 92-116.

[76] Griliches, Z., 1981. Market Value, R&D, and Patents. Economic Letters 7, 183-187.

[77] Griliches, Z., 1990. Patent Statistics as Economic Indicators: A Survey. Journal of Economic Literature, 28, 1661-1707.

[78] Griliches, Z., 1994. Productivity, R&D and the Data Constraint. American Economic Review 84 (1), 1-23.

[79] Griliches, Z., Hall, B.H., Pakes, A., 1987. The Value of Patents as Indicators of Inventive Activity. In Dasgupta, P., Stoneman, P. (eds.) Economic Policy and Technological Performance. Cambridge University Press, 97-124.

[80] Griliches, Z., Hall, B.H., Pakes, A., 1991. R&D, Patents, and Market Value Revisited: Is There a second (Technological Opportunity) Factor?. Economics of Innovation and New Technologies, 1, 183-201.

[81] Griliches, Z., Pakes, A., 1984. Patents and R&D at the Firm Level: A First Look. in Griliches, Z. (ed.) R&D, Patents and Productivity. University of Chicago Press.

[82] Grossman, G., Helpman, E., 1991. Innovation and Growth in the Global Economy. MIT Press.

[83] Hall, B., 1988. Innovation and Market Value. Mimeo, Oxford University and Berkeley.

[84] Hall, B., 1993. The Stock Market Valuation of R&D Investment During the 1980's. American Economic Review 83(2), 259-264.

[85] Hall, B., Jaffe, A., Trajtenberg, M., 2000. Market Value and Patent Citations: A First Look. NBER Working Paper 7741.

[86] Hall, B., Oriani, R., 2005. Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany and Italy, National Bureau of Economic Research Working Paper 10408.

[87] Hall, B.H., Griliches, Z., Hausman, J.A., 1986. Patents and R&D: Is There A Lag? NBER Working Papers 1454, National Bureau of Economic Research, Inc.

[88] Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools. CEPR Discussion Papers 3094.

[89] Hall, H., 2000. Common knowledge: an exploration of social learning in distributed organisations. Poster presented at ISIC 2000 - Information seeking in context: the third international conference on information needs, seeking and use in different contexts. Gothenburg, August 16-18, 2000.

[90] Hansen, G.S., Wernerfelt, B., 1989. Determinants of firm performance: The relative importance of economic and organizational factors. Strategic Management Journal, 10, 399-411.

[91] Harhoff, D., Narin, F., Scherer, F.M., Vopel, K., 1999. Citation frequency and the value of patented inventions. Review of Economics and Statistics, 81(3), 511-515.

[92] Harhoff, D., Scherer, F. and Vopel, K., 2003. Exploring the Trail of Patent Value Distribution in Granstrand, O. Ed. Economics, Law and Intellectual Property: Seeking Strategies for Research and Teaching in a Developing Field, Kluwer, pp 279-310.

[93] Hayashi, F., Inoue, T., 1991. The Relation between Firm Growth and with Multiple Capital Goods: Theory and Evidence from Panel Data on Japanese Firms. Econometrica, 59(3), 731-753.

[94] Heckman, J. J., 1979. Sample selection bias as a specification error, Econometrica, 47, 153-61

[95] Henderson, R., Cockburn, I., 2003. Scale, scope and spillovers : the determinants of research productivity in ethical drug discovery. Working papers 3629-93, Massachusetts Institute of Technology (MIT), Sloan School of Management.

[96] Hicks, D.M., 2002. The Changing Science and Technology Environment. In Teich, A.H., Nelson, S.D., Lita, S.J. (eds.) AAS Science and Technology Yearbook, 255-365.

[97] Hicks, D.M., Katz, J.S., 1996. Science policy for a highly collaborative science system. Science and Public Policy, 23(1), 39-44.

[98] HM Treasury, 2004. Science and Innovation Investment Framework 2004-2014, HM Treasury, London.

[99] House of Commons, 2003. UK Biotechnology Industry: Government Response to the Committee's Twelfth Report of Session 2002-03. 27 November. House of Commons.

[100] Hu, A.G., Jaffe, A.B., 2001. Patent citations and international knowledge flow: The cases of Korea and Taiwan. NBER Working Paper 8528.

[101] Ingham, H., Thompson, S., 1995. Deregulation, Firm Capabilities and Diversifying Entry Decisions: The Case of Financial Services. The Review of Economics and Statistics, 77(1), 177-83.

[102] International Biotechnology Trust, 2005. Annual Report.

[103] Jaffe, A.B., 1986. Technological Opportunity and Spillovers from R&D: Evidence from Firms' Patents, Profits and Market Value. American Economic Review 76, 984-1001.

[104] Jaffe, A.B., Trajtenberg, M., 2002. Patents, Citations, and Innovation: A window on the Knowledge Economy". MIT Press, Cambridge, MA.

[105] Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics, 108(3), 577-98.

[106] Jensen, R.A., Thursby, M.C., 2001. Proofs and Prototypes for Sale: The Licensing of University Inventions. American Economic Review, 91(1), 240-259.

[107] Katz, S., Martin, B., 1997. What is research collaboration? Research Policy 26, 1-18.

[108] Kleinknecht, A. and Verspagen, B., 1989. R&D and Market Structure: the Impact of Measurement and Aggregation Problems. Small Business Economics, vol. 1, no. 4, pp. 297-301.

[109] Knorr-Cetina, K., 1999. Epistemic Cultures. How the Sciences Make Knowledge. Harvard University Press, Cambridge.

[110] Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science, 3(3), 383-398.

[111] Krugman, P., 1983. New Theories of Trade among Industrial Countries. American Economic Review, 73(2), 343-347.

[112] Kuhn, T., S., 1962. The Structure of Scientific Revolutions. The University of Chicago Press, Chicago.

[114] Laband, D.N., Piette, M.J., 1994. The Relative Impacts of Economics Journals: 1970-1990. Journal of Economic Literature, 640-666.

[116] Lancaster, K., 1971. A New Approach to Consumer Theory. The Journal of Political Economy, Volume 74, pp 132-157.

[117] Landes,W. and Posner, R., 1987. Trademark Law: An Economic Perspective. Journal of Law and Economics, Volume 30.

[118] Landes, W. and Posner. R., 2003. The Economic Structure of Intellectual Property Law. Harvard University Press, pp 295 –333.

[119] Lanjouw, J.O., Schankerman, M., 1999. The Quality of Ideas: Measuring Innovation with Multiple Indicators. NBER Working Paper 7345.

[120] Lazarsfeld, P.F., 1972. Insiders and Outsiders: A Chapter in the Sociology of Knowledge. American Journal of Sociology 78, 9-47.

[121] Leibi, T., 2002. Monitoring Eco-Labels: You can Have Too Much of a Good Thing. Discussion Paper, Department of Economics, University of Bern http://ssrn.com/abstract=318540

[122] Lemelin, A., 1982. Relatedness in the patterns of interindustry diversification. Review of Economics and Statistics, 64, 646-647.

[123] Lemley, M., 2001. Rational Ignorance at the Patent Office. Northwestern University Law Review, Volume 95.

[124] Lemley, M. and Shapiro, C., 2004. Probabilistic Patents. Stanford Law and Economics Olin Working Paper 288.

[125] Leung, S.F., Yu, S. 1996. On the Choice Between Sample Selection and Two Part Models. Journal of Econometrics, 72: 196-229.

[126] Lev, B., Sougiannis, T., 1996. The Capitalization, Amortization and Value-Relevance of R&D. Journal of Business, Accounting and Economics, 21, 107-138.

[127] Lewison, G., Cottrell, R., Lucocq, B., Dawson, G., 2003. Mapping the Landscape, National Biomedical Research Outputs, 1988-95. Wellcome Trust, Policy Report No. 9.

[128] Lewison, G., Devey, M. E. 1999. Bibliometric methods for the evaluation of arthritis research. Rheumatology, 38, 13-20

[129] Lewison, G., Paraje, G., 2004. The classification of biomedical journals by research level. Scientometrics, 60 (2), 145-157.

[130] Lichtman, D., Baker, S., Kraus, K. 2000. Strategic Disclosure in the Patent System. Vanderbilt Law Review, 53, 2175-2217.

[131] Lieberman, M.B., 1987. Strategies for Capacity Expansion. Sloan Management Review, 28(4), 19-27.

[132] Liebeskind, J.P., Oliver, A.L., Zucker, G.L., Brewer, M. 1996. Social Networks, Learning, and Flexibility: Sourcing Scientific Knowledge in New Biotechnology Firms. Organization Science, 7, July/August, 428–43

[133] Lim, K., 2004. The Relationship between Research and Innovation in the Semiconductor and Pharmaceutical Industries (1981-97). Research Policy, 33 (2): 287-321.

[134] Lippman, S.A., Rumelt, R.P., 1982. Uncertain imitability: An analysis of interfirm differences in efficiency under competition. Bell Journal of Economics, 13, 418-438.

[135] Long, C., 2002. Patent Signals. University of Chicago Law Review, Volume 625.

[136] Machlup, F. and Penrose. E., 1950. The Patent Controversy in the Nineteenth Century. Journal of Economic History, Volume 10, pp 1-29.

[137] Maddison, A., 1991. Dynamic Forces in Capitalist Development. A Long-Run Comparative View. Oxford and New York: Oxford University Press.

[138] Mahoney, J., Pandian, J.R., 1992. The resource-based view within the conversation of strategic management. Strategic Management Journal, 13, 363-380.

[139] Mansfield E., Rapoport, J., Romeo, A., Wagner,S., Beardsley, G., 1977. Social and Private Rates of Return from Industrial Innovation. Quarterly Journal of Economics, 91, 221–240

[140] Mansfield, E., 1989. Basic Research and Productivity Increase in Manufacturing. American Economic Review. 70, 863-873.

[141] McGee, J., Thomas, H., 1986. Strategic groups: theory, research and taxonomy. Strategic Management Journal, 7, 141-160.

[142] Meade, J.E., 1952. External economies and diseconomies in a competitive situation. Economic Journal, 62, 54–67.

[143] Merges, R., 1999. As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform. Berkeley Technology Law Journal, Volume 14.

[144] Merton, R.K., 1968. The Matthew effect in science. Science, 159, 56–63.

[145] Meyer, M., 2000. Does Science Push Technology? Patents Citing Scientific Literature. Research Policy, 29, 409-434.

[146] Meyer-Krahmer, F., Schmoch, U., 1998. Science-based technologies: university-industry interactions in four fields. Research Policy, 27 (8), 835-851.

[147] Montgomery, C.A., Hariharan, S., 1991. Diversified expansion by large established firms. Journal of Economic Behavior and Organization, 15, 71-89.

[148] Mora-Valentin, E.M., Montoro-Sanchez, A., Guerras-Martin, L.A., 2004. Determining factors in the success of R&D cooperative agreements between firms and research organizations. Research Policy, 33, 17-40.

[149] Mowery, D.C., Oxley, J.E., Silverman, B.S., 1998. Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. Research Policy, 27, 507-523.

[150] Mueller, D.C., 1986. Profits in the Long Run. Cambridge University Press, Cambridge.

[151] Murray, F., 2002. Innovation as Co-evolution of Scientific and Technological Networks: Exploring tissue engineering. Research Policy, Vol 31, 1389-1403.

[152] Murray, F., 2004. The role of academic inventors in entrepreneurial firms: sharing the laboratory life. Research Policy, 33, 643-659.

[153] Narin, F., Noma, E., Perry, R., 1987. Patents as indicators of corporate technological strength. Research Policy, 16(2-4), 26-48.

[154] Narin, F., Olivastro, D., 1992. Status Report – Linkage Between Technology and Science. Research Policy, 21(3), 237-249.

[155] Narin, F., Rozek R.P., 1988. Bibliometric analysis of U.S. pharmaceutical industry research performance. Research Policy, 17, 139-154.

[156] Nelson, R.R. (ed.), 1962. The Rate and Direction of Inventive Activity. Princeton University Press.

[157] Nelson, P., 1973. Information and Consumer Behavior. Journal of Political Economy, Volume 78, pp 311-29 at p 311.

[158] Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. Belknap Press/Harvard University Press, Cambridge.

[159] OECD, 2002. Benchmarking Science-Industry Relationships. OECD, Paris.

[160] OECD Report., 2002. Short Summary of the Workshop on Genetic Inventions, Intellectual Property Rights and Licensing Practices.

[161] Oddi, S., 1996. Un-Unified Economic Theory of Patents. Notre dame Law Review, Volume 71.

[162] Oriani, R., and Sobrero M., 2003. A Meta-Analytic Study of the Relationship between R&D Investments and Corporate Value. In Calderini M., Garrone P., Sobrero M., (Eds), Corporate Governance, Market Structure and Innovation, Edward Elgar: Cheltenham.

[163] Pakes, A., 1985. On Patents, R&D, and the Stock Market Rate of Return. Journal of Political Economy, 93(2), 390-409.

[164] Pakes, A., 1995. On Patents, R&D and the Stock Market rate of Return. Journal of Political Economy, 93, 390–409.

[165] Pakes, A., Simpson, M., 1989. Patent Renewal Data. NBER Reprints 1265, National Bureau of Economic Research, Inc.

[166] Parchomovsky, G., 2000. Publish or Perish. Michigan Law Review, 98, 926-952.

[167] Parchomovsky. G. and Wagner, P., 2005. Patent Portfolios. University of Pennsylvania, Institute for Law & Economics Research Paper 04-16.

[168] Patel, P., Pavitt, K., 1997. The Technological Competencies of the Worlds Largest Firms: Complex and Path Dependent, but not Much Variety. Research Policy, 26, 141-156.

[169] Paula, S., 1996. The Economics of Science. Journal of Economic Literature, 34 (3), 1199-1235.

[170] Penan, H., 1996. R&D Strategy in a Techno-economic network: Alzheimer's disease therapeutic strategies. Research Policy, 25: 337-358.

[171] Penrose, E.T., 1959. The Theory of the Growth of the Firm. John Wiley, New York.

[172] Peteraf, M., 1993. The cornerstones of competitive advantage: A resource-based view. Strategic Management Journal, 14, 179-191.

[173] Pisano, G.P., 1997. R&D Performance, Collaborative Arrangements, and the Market-for-Know-How: A Test of the "Lemons" Hypothesis in Biotechnology. Working Paper, Harvard Business School, Boston.

[174] Polanyi, M., 1962. Personal Knowledge. Routledge and Kegan Paul, London.

[175] Polanyi, M., 1967. The Tacit Dimension. Anchor Books, New York.

[176] Porter, M.E. (ed.), 1986. Competition in Global Industries. Harvard Business School Press, Boston.

[177] Praiss, D., 2001. Creating a Winning Patent Portfolio. Nature Biotechnology, Volume 19.

[178] Prahalad, C.K., Hamel, G., 1990. The Core Competence of the Corporation. Harvard Business Review, 68, 79-91.

[179] Quah, D., 1992. Empirical cross-section dynamics in economic growth. Discussion Paper/Institute for Empirical Macroeconomics 75, Federal Reserve Bank of Minneapolis.

[180] Reitzig, M., 2003. What Determines Patent Value? Insights from the Semiconductor Industry. Research Policy, Volume 32, pp. 13-26.

[181] Robinson, J., 1933. The Economics of Imperfect Competition. Macmillan.

[182] Roheim, C., 2002. Early Indications of Market Impacts from the Marine Stewardship Council's Ecolabeling of Food. Available at http://www.wto.org/english/forums_e/ngo_e/ccc_msc_e.doc

[183] Romer, P.M., 1989. What Determines the Rate of Growth and Technological Change? The World Bank Policy Research Working Papers, WPS 279.

[184] Romer, P.M., 1990. Endogenous Technological Change. Journal of Political Economy 98 (5), 71-102.

[185] Rosenberg, N., 1982. Inside the Black Box: Technology and Economics. Cambridge University Press, Cambridge.

[186] Rosenberg, N., 1990. Why do Firms do Basic Research (with their own money). Research Policy, 19(2), 165-74.

[187] Ross, S.A., Westerfield, R.W., Jaffe, J.F., 1996. Corporate Finance (4th edition). Richard D. Irwin, London.

[188] Rumelt, R.P., 1984. Towards a strategic theory of the firm. In Lamb, R.B. (ed.) Competitive strategic management. Prentice-Hall, Englewood Cliffs.

[189] Rumelt, R.P., 1987. Theory, strategy and entrepreneurship. In Teece, D.J. (ed.) The competitive challenge: Strategies for industrial innovation and renewal. Ballinger Publishing Company, Cambridge, MA.

[190] Rumelt, R.P., 1991. How much does industry matter? Strategic Management Journal, 12,167-185.

[191] Santarelli, E. and Sterlachinni, A., 1990. Innovation, Formal vs. Informal R&D, and Firm Size: Some Evidence from Italian Manufacturing Firms. Small Business

Economics, 2(2), 223-228.

[192] Schankerman, M., Pakes, A., 1985. Valeur et Obsolescence des brevets. Revue Economique, 36, 917-941.

[193] Schankerman, M., Pakes, A., 1986. Estimates of the Value of Patent Rights in European Countries during the Post-1950 Period. Economic Journal, 96, 1052–1077.

[194] Schankerman, M. and Pakes, A., 1998. How Valuable is Patent Protection? Estimates by Technology Field. Rand Journal of Economics, Volume 29, pp 77-107.

[195] Scherer, F. (1965): Firm Size, Market Structure, Opportunity and the Output of Patented Inventions. American Economic Review, Volume 55, pp 1097-1123.

[196] Scherer, F.M., 1965. Firm Size, Market Structure, Opportunity, and the Output of Patented Innovations. American Economic Review, 55, 1097-1123.

[197] Scherer, F.M., 1982. Demand-pull and Technological Opportunity: Schmookler Revisited. Journal of Industrial Economics, 30, 225-37.

[198] Schmalensee, R., 1992. Sunk Costs and Market Structure: A Review Article. Journal of Industrial Economics, 40(2), 125-34.

[199] Schmookler, I., 1966. Invention and Economic Growth. Harvard University Press.

[200] Schumpeter, J.A., 1942. Capitalism, Socialism, and Democracy. Harper and Brothers.

[201] Shelanski, H.A., Klein, P.G., 1995. Empirical Research in Transaction Cost Economics: A Review and Assessment. Journal of Law, Economics and Organization, 11(3), 335-361.

[202] Solow, R.M., 1957. Technical Change and the Aggregate Production Function. Review of Economics and Statistics, 39, 312-320.

[203] Sorenson, O., Fleming, L., 2004. Science and the diffusion of knowledge. Research Policy 33, 1615-1634.

[204] Spence, A.M., 1981. The Learning Curve and Competition. Bell Journal of Economics, The RAND Corporation, 12(1), 49-70.

[205] Spender, J.C., Grant, R.M., 1996. Knowledge and the firm: Overview. Strategic Management Journal, 17, 5-9.

[206] Stern, S., 1999. Do Scientists Pay to be Scientists? NBER Working Paper Series, 7410.

[207] Stigler, G., 1968. The Organization of Industry. Chicago University Press.

[208] Stigler, G., 1961. The Economics of Information. Journal of Political Economy, Volume 69, pp. 213-225 at p 213.

[209] Spence, M., 2002. Signaling in Retrospect and the Informational Structure of Markets, American Economic Review 92 (3): 434-459.

[210] Teece, D.J., 1980. Economies of scope and the scope of the enterprise. Journal of Economic Behavior and Organization, 1, 223-247.

[211] Teece, D.J., 1982. Towards an economic theory of the multiproduct firm. Journal of Economic Behavior and Organization, 3, 39-63.

[212] Teece, D.J., 1986. Profiting From technological innovation: Implications for integration, collaboration, licensing and public policy. Research Policy, 15, 785-805.

[213] Teece, D.J., 1990. Contributions and impediments of economic analysis to the study of strategic management. In Fredrickson, J.W. (ed.). Perspectives on Strategic Management. Harper & Row, New York.

[214] Tijssen, R.J.W., Buter, R.K., Van Leeuwen, T.N., 2000. Technological relevance of science: An assessment of citation linkages between patents and research papers. Scientometrics, 47, 389-412.

[215] Toivanen O., Stoneman P., Bosworth D., 2002. Innovation and Market value of UK Firms, 1989-1995. Oxford Bulletin of Economics and Statistics, 64: 39-61.

[216] Tollison, R.D., 1982. Rent Seeking: A Survey. Kyklos, 35, 575–602.

[217] Trajtenberg, M., 1990. A penny for your quotes: patent citations and the value of innovations. RAND Journal of Economics, 21, 172–187.

[218] Wagner, P. and Petherbridge, L., 1999. Is the Federal Circuit Succeeding? An Empirical Assessment of Judicial Performance. University of Pennsylvania Law Review, Volume 77.

[219] Wall Street Journal, *Focus on Biotech*, 18/10/05.

[220] Ward. M., Dranove, D., 1995. The Vertical Chain of R&D in the Pharmaceutical Industry. Economic Enquiry, 33, 1-18.

[221] Wildasin, D., 1984. The tobin-q theory of investment with many capital goods. American Economic Review, 74(1), 203–210.

[222] Williamson, O.E., 1975. Markets and Hierarchies, Free Press, New York.

[223] Williamson, O.E., 1985. The Economic Institutions of Capitalism. Free Press, New York.

[224] Winter, S., 1988. On Coase, Competence, and the Corporation. Journal of Law, Economics and Organization, 4, 163-180.

[225] Zucker, G.L., Darby, M.R., 1995. Virtuous Circles of Productivity: Star Bioscientists and the Institutional Transformation of Industry. NBER Working Paper No. 5342.

[226] Zucker, G.L., Darby, M.R., Brewer, M., 1998. Intellectual Capital and the Birth of U.S. Biotechnology Enterprises. American Economic Review, 88 (1), 290-306.