AN ABSTRACT OF THE DISSERTATION OF

Chenggang Wang for the degree of Doctor of Philosophy in Agricultural and Resource Economics, presented on March 16, 2007.

Title: Public Investment Policy and Industry Incentives in Life Science Research

Abstract approved:____

Steven T. Buccola Yin Xia

The biorevolution in the 1970's greatly stimulated investment in life-science research. The present dissertation is aimed at evaluating the impact of US public investment on industrial investment in life-science research. The focus is on three major life-science fields: biology, medicine, and agriculture. A dynamic model of industrial R&D investment is developed to identify the channels of influence which public R&D investment has on industrial R&D investment. The model takes into consideration four determinants of industrial R&D investment: market demand, technological opportunity, supply of scientific labor, and adjustment costs. The model is estimated using R&D expenditures and patent counts constructed at the individual life-science field level.

Results show that, as far as the life sciences are concerned, the R&D performed by public institutions has been the primary cause of the past two decades' surge in industrial R&D investment. Even after accounting for the negative wage effect, public-

institution R&D has been strongly complementary to industrial R&D, both in agriculture and in medicine. Public institutions' basic biological research has had a significant "infrastructure" effect on industry's agricultural and medical research. Although analysts typically have argued that market demand and technological opportunity are equally important determinants of the pace and direction of technological change, we find that, in the life sciences at least, the dominating stimulant to industrial R&D investment has been technology push, i.e., the creation of new technological opportunities when advances are made in public institutions' research.

Public Investment Policy and Industry Incentives in Life Science Research

by

Chenggang Wang

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented March 16, 2007

Commencement June 2007

Doctor of Philosophy dissertation of Chenggang Wang presented on March 16, 2007

Approved:

Co-Major Professor, representing Agricultural and Resource Economics

---

Co-Major Professor, representing Agricultural and Resource Economics

---

Head of the Department of Agricultural and Resource Economics

---

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University Libraries. My signature below authorizes release of my dissertation to any reader upon request.

Chenggang Wang, Author

ACKNOWLEDGEMENTS

I would like to extend my deep gratitude and appreciation to my co-major advisors, Dr. Steven Buccola and Dr. Yin Xia, without whom this dissertation would not have been possible. I thank them for providing me the opportunity to work on the NRI project for the past two years, and allowing the freedom to use my work in that project to develop this dissertation. Steve has fostered my personal and intellectual growth by his exceptional professionalism and impeccable clarity. I thank Yin for her patience, time and friendship. I could not imagine how I would finish data construction without her help and guidance. I sincerely appreciate the contributions of Dr. Robbin Shoemaker at the USDA and Dr. William Folk at the University of Missouri–Columbia with data construction.

A special thanks goes to Professors Rolf Färe and Clark Seavert. Our collaborative work has been instrumental in helping me to get my first academic job. Rolf, a valuable friend and “tough” mentor, has taught me so much about research through numerous joyful discussions. Addressing his challenges has proven to be the most effective way for self-improvement. I thank Clark for providing financial support in the early stage of my graduate study, and for educating me on the fruit growing business.

I deeply thank Dr. Munisamy Gopinath for his valuable advice in each developmental stage of this dissertation. I would also like to thank Dr. Robert Smythe for his supervision of my study in statistics. Thank you to Dr. James Thompson for serving on my committee.

I deeply appreciate Vijay Satyal's friendship and help in the most trying days when I first came to the United States. I would also like to thank my friend Nicholas Rada for hours of proof-reading and valuable comments.

I am grateful for those who have made it possible for me to get to where I am: my family and friends in China. I particularly thank my uncle Wang Xiaode for encouraging me to pursue higher education in the United States and his faith in my academic potential.

Finally, I would like to express my gratitude to my best friend and wife Zhang Yan and our son Anthony for being my true inspiration.

Financial support for this study has been provided by the Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture, under NRI Grant No. 2005-35400-15881.

TABLE OF CONTENTS

TABLES OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF APPENDIX TABLES

DEDICATION

To my grandparents

Public Investment Policy and Industry Incentives in Life Science Research

Chapter 1: Introduction

Life science has become one of the most promising forces of technological change and economic growth. The last two decades have seen a surge of investment in life-science research by both the public and private sector. According to my estimates, in 1980 U.S. public institutions and industrial firms listed in the U.S. stock markets together invested $21.5 billion (2001 constant dollars) in the three major life science fields: biology, medicine and agriculture. By 2003, they were spending $103.7 billion (2001 constant dollars), about five times the investment in 1980. In contrast, the total U.S.R&D investment in all fields in 2003 was only less than 2.5 times the investment in 1980 (National Science Foundation, 2003, table B-21).

On the other hand, the sectoral composition of life science R&D has undergone a dramatic change in this period. Of the $21 billion R&D in 1980, 60% was invested and performed by public institutions such as federal and state governments, universities and colleges, and other nonprofit institutions. The other 40% was invested and performed by industry. By 2003, the public sector only accounted for 29% of the $103.7 billion R&D.

The rising importance of industrial investment calls for a comprehensive evaluation of the role played by public institutions in life science research. The economic rationale of public institutions' direct involvement in research activity is that a free-market economy would underinvest in research because of the two peculiar properties of knowledge—nonrivalry and partial excludability—as well as the inherent uncertainty in knowledge production. Thus, a socially desirable rate of technological change requires public institutions that are not governed by the profit-and-loss criterion for financing research activity [Arrow (1962); Nelson (1959).] Given that so many resources have been devoted to life science research by industry, should public institutions continue their direct involvement in this enterprise? Does public investment have anything to do with the surging industrial investment in life-science activity? Does public investment in life science research encourage, or substitute for, industrial investment? These questions are especially pertinent to the making of science policy in the federal government, given that a large proportion of federal R&D funds is now being invested in life sciences. In 2003 life sciences received 53.7% of total federal funds for R&D, according to the recent report, "National R&D Trends," by the National Science Foundation (2003).

In order to trace the channels of influence which public R&D investment has on industrial R&D investment, a thorough understanding of the determinants of industrial R&D investment is needed. The determinants of R&D investment have been extensively studied in the literature. Market demand, technological opportunity, appropriability of research outcomes, and research input price are considered to be the major factors a firm must take into consideration in making R&D investment decisions. Moreover, even potentially the most profitable R&D project requires the firm be in a position to finance the project; and when rapid R&D investment incurs adjustment costs, the firm must slow down the pace of investment.

The purpose of this dissertation is to evaluate the impact of public investment on industrial investment in life science research through a structural examination of industry's dynamic investment decisions. I focus on the three major life science fields: biology, medicine, and agriculture. I develop a multi-field, two-sector model through which I simultaneously examine knowledge spillovers and industry's dynamic investment decisions. I estimate the model with a unique dataset on R&D expenditures and patent counts, constructed at the individual field level. Results shed light on the contribution of public R&D to the growth of industrial R&D in life science, and provide guidance for future public investment policy in life science research.

Chapter 2: Literature Review

2.1 Economic Rationale of Public Intervention in Research and Development

Technological change has been widely viewed as one of the most important driving forces of economic growth [Romer (1990); Solow (1957).] It is also well established in the literature that a free market economy would fail to sustain a socially desirable rate of technological change and that the public sector should intervene in this enterprise. Two schools of thoughts underlie this argument. On the one hand, from a Schumpeterian perspective, innovation creates market power and therefore generates welfare losses in the short run. On the other hand, innovation reduces production cost and brings a greater variety of commodities to consumers, improving social well-being in the long run. Moreover, it is the potential quasi-rent available to the innovator as a monopolist in the market that constitutes the incentive for continuing innovation (Schumpeter, 1950). To sustain long-term prosperity, therefore, government should allow the innovator to have monopolistic market power for a certain period of time. If the monopoly period is too long, the long-run welfare gain will not cover the short-run welfare sacrifice. The optimal solution therefore lies somewhere in-between. This theory constitutes the economic rationale for the patent system and has played an important role in the evolution of antitrust law.

From the point view of welfare economics, the social benefit of basic research is far higher than the private benefit appropriable to profit-seeking firms, implying that a free market economy would underinvest in basic research (Nelson, 1959). More generally, the production and trading processes of knowledge as a good have certain peculiarities: increasing returns in use, inappropriability, and uncertainty—the three classic reasons for the failure of a competitive economy to achieve optimal resource allocation—all hold in this case (Arrow, 1962). The central theme in Nelson and Arrow's studies is that a free market economy would underinvest in R&D and optimal allocation of resources in inventive activity calls for governments—or some other agencies not governed by the profit-and-loss criteria—to finance R&D. Having said that, it is also likely that competition may cause overinvestment in R&D under certain circumstances. Dasguputa and Stiglitz (1980) show that competition in a specialized market may induce replicated R&D programs, resulting in a waste of social resources. These theoretical studies are the underpinnings of many government R&D policies, such as direct provision of research by the public sector, R&D tax reduction, and R&D subsidies.

Evidence of the success of public involvement in research activity is abundant. A successful example is the U.S. agricultural research system. As documented in Huffman and Evenson (1993), the agricultural sector has on average experienced a 1.61 percent annual growth in real output and 1.62 percent growth in productivity during the 20th century. Public R&D investments in agriculture are found to be the major contributors to this remarkable growth. Academic basic research has also been found to be conducive to economic growth in the United States. Combining scientific publication data with productivity data, Adams (1990) finds that fundamental stocks of knowledge generated in academic science have been the major contributor of productivity growth in the United States.

2.2 Interactions of Public and Private R&D Investment

Public-sector R&D investment not only directly enhances the private sector's productivity, but has a profound influence on the private sector's R&D investment decisions. A large body of literature has attempted to test whether public-sector R&D encourages, or substitutes for, industry's R&D investment. An excellent survey by David et al. (2000) shows that no satisfactory, conclusive answer can be drawn from this literature. According to these authors, econometric models and data used tend to be non-comparable; more importantly, the entire literature lacks the guidance of an appropriate structural model, leaving results difficult to interpret. To fill this gap, David and Hall (2000) propose a two-sector model to analyze channels of influence that public investment policy may have on private-sector investment decisions. On the one hand, knowledge generated by public-sector research may directly increase the research productivity of the private sector. On the other hand, the public and private sector are competing for the same input factor for knowledge production: scientific labor. Consequently, an increase in public-sector demand for scientific labor (namely, R&D investment) will increase wages and reduce private-sector R&D investment, *ceteris paribus*.

This model provides a useful framework for considering the research questions raised in the present dissertation. A comprehensive assessment of the effects of the public sector's R&D investment on private sector investment requires a good understanding of both the technical and behavioral aspects of the problem. The technical aspect refers to knowledge spillovers across sectors and across fields, while the behavioral aspect bears on how the private sector's R&D investment changes in response to the public sector's investment policy.

The David and Hall model (henceforth D-H model) is restrictive in the sense that adjustment costs in R&D investment are ignored. Yet both theoretical reasoning and empirical evidence suggest adjustment costs are present in R&D investment (Himmelberg and Petersen, 1994). As a result, R&D investment tends to approach gradually, rather than jump instantaneously toward, the equilibrium. This dynamic disequilibrium phenomenon is missing in the D-H model. In addition, David and Hall focus on theoretical analysis but are silent about how to empirically implement their approach.

In Chapter 3 of this dissertation I develop a theoretical approach which extends the static D-H model into a dynamic one by taking into account adjustment costs in R&D investment. The model provides the theoretical foundation for my econometric analysis of public and private interactions in life-science R&D investment. The remainder of this Chapter reviews the literature on the technical and behavioral aspects of public and private funding interactions.

2.3 The Search for Knowledge Spillovers

2.3.1 Performance Response to Knowledge Spillovers

The non-rivalry and partial excludability properties of knowledge are pertinent to a wide range of economics fields, from public finance to industrial organization and to growth theory. Not surprisingly, a large body of literature has emerged searching for empirical evidence on knowledge spillovers. The most popular approach is to examine the effect of R&D spillovers on the performance of the spillover-receiving unit, measured by productivity growth in a primal or dual framework. The statistical observations can be business lines, firms, or industries.

The central issue in this literature is the construction of the R&D spillover stock. The R&D spillover stock usually is defined as a weighted sum of R&D investments by the potential "spilling" units, with the weights measuring the "distances" between the spillover-receiving and spillover-giving units. Three weighting schemes have been adopted in the literature. The first uses the industry input-output matrix to construct a matrix of weights, the second is based on a technology flows matrix constructed using R&D expenditure and patent data (Scherer, 1984), and the third adopts a matrix of technological proximity weights derived from the diversification of a firm's patents across various patent classes (Jaffe, 1986).

Several surveys on this literature conclude that there exist reliable evidence of the presence of R&D spillovers. In his earlier survey of this literature, Griliches (1992) points out a long list of flaws in the conventional measures of the R&D spillover stock. Despite these measurement problems, he concludes with convincing evidence that R&D spillovers are present, their magnitude may be quite large, and social rates of return remain significantly above private rates. Mohnen's survey (1996) concludes that on the whole the literature suggests borrowed R&D exerts a more significant marginal effect on productivity than own R&D. Yet, own R&D and borrowed R&D affect productivity growth differently for different countries. For example, the United States seems to benefit more from own R&D than from foreign R&D, whereas Japan appears to benefit more from foreign R&D than own R&D (Cincera and Bruno Van Pottelsberghe de la, 2001). Evidence on cross-firm knowledge spillovers has been found in many studies. Cincera and Van Pottelsberghe de la (2001) review 38 econometric estimation results on this issue, 14 of them being insignificant and 23 of them significantly positive.

Although evidence seems in favor of the presence of R&D spillovers, caution must be taken in interpreting the results. In many interindustry studies, the performance response approach fails to distinguish two types of R&D spillovers, making it difficult to interpret which is actually being measured.

2.3.2 Knowledge Spillovers and Rent Spillovers

Griliches (1979) shows that two major sources of externalities are associated with R&D: knowledge spillovers and rent spillovers. Knowledge spillovers arise from imperfect appropriability of knowledge: knowledge can transcend the boundaries between various economic institutions with little cost. Different from knowledge spillovers, rent spillovers are associated with economic transactions. An innovator cannot capture all the benefits from inventing a new product unless she can exercise first-degree price discrimination. Additionally, conventional measures of productivity improvement do not correctly capture the benefits of an innovation to either the supplying or the buying industry, unless the price deflator for the new product is a perfect hedonic price index.

Thus, the distinction needs to be made clearly between the two different types of R&D spillovers, especially when one is assessing R&D investment policy. Confusing them may result in overestimation of knowledge spillover effects and therefore flawed policy implications.

The performance response framework discussed in 2.3.1 fails to distinguish between these two types of spillovers. The spilling industry's R&D may influence the spillover-receiving industry's performance through either knowledge spillovers or rent spillover or both. Consequently, the parameter on the R&D spillover stock is a measure of mixed spillover effects.

2.3.3 Alternative Approaches to Searching for Knowledge Spillovers

Alternative approaches to searching for knowledge spillovers exist in the literature. One of them is to directly estimate the spillover weights, instead of imposing any *a priori* weights. Bernstein (1989) and Bernstein and Nadiri (1988;1989) do so by treating each industry as a distinctive spillover source. As a result, their estimates yield a network of knowledge spillovers, where each industry is both the spillover sender and receiver. The advantage of this approach is that the weights are estimated directly. The disadvantage lies in the potential multicollinearity in explanatory variables, and in the weak degrees of freedom when a large number of industries under study are observed in a short time period. Note that, for interindustry studies, this approach does not distinguish between knowledge and rent spillovers either.

In an intra-industry study such as Bernstein and Nadiri (1989), however, the distinction between the two types of spillovers is not an issue because, presumably, there exist no rent spillovers within an industry. The problem remains for studies on international knowledge spillovers unless one is working on firm-level data within the same industry, or the industries under study do not use one another's products as production inputs.

Several authors estimate R&D spillovers in a knowledge production function framework, where R&D stocks are inputs and patent counts are the output [e.g., Henderson and Cockburn (1993) and Jaffe (1986)]. Instead of looking at the effects of R&D spillovers on productivity growth in the production of physical goods, these authors examine the effects of R&D spillovers on productivity improvement in research activity. The distinction between knowledge spillovers and rent spillovers is not at issue here because the former are the only spillover effects that prevail.

In summary, the performance response approach has the inherent difficulty in distinguishing between knowledge spillovers and rent spillovers. When external R&D

can influence the performance of an economic unit through both knowledge and rent spillovers, a simple regression of performance on R&D spillover stock links one common end of the two channels directly to another and therefore will necessarily mingle the two distinct forces of influence together.

There are at least two ways to avoid this problem. One way is to carefully select a statistical sample for which borrowed R&D influences the performance of the economic unit only through the channel of interest. The intra-industry study, for instance, is of this kind. Another way to avoid the problem is to focus on intermediate steps of the causality chain to identify the channel of influence in question. The knowledge production function approach falls into such a category. In the present dissertation, I estimate knowledge spillovers in a knowledge production function framework. Hence, my analysis is not subject to the problem discussed above.

2.4 Determinants of R&D Investment

The literature in the 1960s and 1970s has seen a heated debate on whether market demand or technological opportunity is the primary determinant of the rate and direction of technological change. In addition to market demand and technological opportunity, determinants of R&D investment include other factors such as appropriability, factors that affect supply of R&D input, financial constraints facing the R&D investor, and adjustment costs in R&D investment. In this Section I review and synthesize the literature on the determinants of R&D investment.

2.4.1 Demand-Pull vs. Technology-Push

Since Schumpeter’s (1950) masterpiece was published, economists have devoted a great amount of interest and effort to opening the black box of inventive activity for a deeper understanding of technological change, business cycles, and economic growth. The literature emanated from a fundamental question: what are the underlying forces that determine the rate and direction of technological change?

Schmookler (1966) argues that technological change—rather than residing in a black box independent of economic activities—is determined entirely by economic forces, and predominantly by demand forces. In his study on the railroad industry, Schmookler finds that increases in the purchase of railroad equipment and components are followed by increases in inventive activity measured by patent counts. Schmookler also finds evidence from cross-industry data to support his demand-pull hypothesis.

As documented in Ruttan (1997), the debate about the primary role of demand side forces and supply side forces intensified in the late 1960s. On the one hand, studies supported by the Office of the Director of Defense Research and Engineering show that the significant research programs which contributed to the development of 20 major weapons systems were motivated primarily by military need. On the other hand, a series of studies initiated by the National Science Foundation finds that scientific breakthroughs have been of much greater importance as a source of technical change.

This debate ended in the late 1970s, mainly owing to Rosenberg’s critique (1974) on Schmookler’s work, as documented in Ruttan (1997). The main objection Rosenberg raised is Schmookler's neglect of the supply side forces. According to her analysis, Schmookler's conclusions are based on an inappropriate assumption that the supply curve of inventions is perfectly elastic. Using an extensive list of historical events of scientific and technological breakthroughs, she demonstrates how technological progress has been stagnant in certain disciplines because of the constraints of human knowledge; even when strong demands exist. Rosenberg argues that technological opportunities are different for different industries, and therefore inventions are not equally possible in all industries. Accordingly, a better understanding of the nature of inventive activity calls for an integrated investigation into both demand factors and technological opportunities.

Schmookler's empirical study has been challenged by later researchers as well. Scherer (1982) retests Schmookler's empirical study using a more comprehensive dataset, which includes all manufacturing industries in contrast to only 23 industries in Schmookler's dataset. Instead of focusing only on capital goods inventions, Scherer also studies industrial materials inventions. Although Scherer finds significant correlation between demand indices and patent counts in all cases, the correlation is much weaker for materials inventions and the R-squares are far lower than those obtained by Schmookler. As for the relative importance of demand side forces and technological opportunity, Schmookler's conclusion is again only supported by the sample of capital goods inventions, but not by that of material goods inventions. Scherer then concludes that Schmookler uses the data best suited for supporting his hypothesis, and argues that both demand factors and differential technological opportunities must be taken into account for an adequate explanation of the rate and direction of technological change.

2.4.2 Appropriability

The firm's R&D investment, like any kind of investment, depends on the expected present value of the benefits that can be derived from the investment. The private benefits of R&D are the quasi-rents an innovating firm can extract as a monopolist. Because of the partial excludability of knowledge, the firm may not be able to appropriate all the benefits derived from its R&D effort. Appropriability, which may depend on entrepreneur ability, industrial market structure, and the general institutional framework in which the firm operates, is another crucial factor in the determination of the firm's demand for R&D investment.

In the early 1980s, two ambitious studies by Levin and Reiss (1984) and Pakes and Schankerman (1984) attempted to combine appropriability with market demand and technological opportunity to explain the distribution of R&D intensity across firms and across industries. Pakes and Schankerman (1984) find that demand factors only explain 5% of the intraindustry variance in R&D intensity, whereas about 70% of the intraindustry variance is explained by technological opportunity and appropriability. In the interindustry sample, the story is quite different. Demand factors account for the majority of the interindustry variance in R&D intensity. According to these authors, that different results are obtained at the two different levels of aggregation is due to factors affecting the R&D intensities of all firms within the industry. This finding complements those obtained by Schmookler (1966) and Scherer (1982) by specifying that it is the industry-specific demand factors rather than firm-specific demand factors that influence

R&D investment decisions. One of the limitations of this study is that the measurement of technological opportunity and appropriability is rather crude: the two of them are jointly measured by dummy variables.

Levin and Reiss (1984) propose a model to test whether interindustry differences in technological opportunity, appropriability, and demand can satisfactorily explain differences in R&D intensity, concentration, and advertising. The hypothesis that technological opportunity and appropriability do not matter is decisively rejected. The study stands out from previous ones because of its comprehensiveness. It not only incorporates all the three determinants of R&D intensity mentioned above, but also is featured with joint determination of market concentration and R&D intensity. In addition, a government R&D equation is included. They find that government funded R&D stimulates industry R&D by enhancing industry's technological opportunity conditions. The comprehensiveness of this study, however, comes at a cost. The structural model is based on a restrictive assumption that all the firms are assumed to be homogeneous and symmetric. Another limitation of this study is the use of dummy variables to capture the intraindustry variance in technological opportunity.

2.4.3 Supply of Scientific Labor

David and Hall (2000) probably is the only study to date that emphasizes the importance of factors that affect supply of R&D input in the determination of R&D investment. They propose a two-sector (the public and private sector) R&D investment model to study the channels through which the public investment policy may influence the private sector's investment. The private sector's demand for R&D investment is simply determined by equating marginal product value of R&D investment with wage of scientific labor. The two sectors are competing for the same input factor of knowledge production: scientific labor. When the supply of scientific labor is inelastic, an increase in governmental R&D investment may boost the wage of scientists and engineers and thus inhibit industry's investment incentive.

2.4.4 Financing of R&D

Even the best investment opportunity may be left unexploited if no money can be found to support the investment project. The optimal level of R&D investment— as is determined by demand factors, technological opportunity, and supply conditions of research input—will never be achieved if the firm is not in a position to finance that much R&D. Financial constraints frequently force R&D managers to suspend or drop potentially profitable R&D projects. On the one hand, the need for R&D investment can seldom be completely satisfied by internal cash flow. On the other hand, external financing is oftentimes more costly than internal financing because of asymmetric information between the investor and innovator (Hall, 2005).

In this context, the asymmetric information problem refers to the fact that the investor is not as informed about the likelihood of success—and the risks—of the R&D project as is the innovator. The investor's failure to distinguish good projects from bad projects generates an extra “lemons” premium for the transaction to take place. In the extreme case where the asymmetric information problem is too serious and the variance in the profitability of R&D projects is too large, the marketplace for the financing of R&D will disappear completely. Still worse, the non-rival nature of knowledge makes it infeasible to alleviate the problem through the innovator providing more information about the project to the investor. The innovator is unwilling to reveal detailed information about the R&D project to be financed; because once knowing the information, the investor can exploit the value of the information with little cost (Hall, 2005).

Empirical evidence is abundant that investment (including both physical and R&D investment) tends to be highly correlated with measures of internal financial resources; the correlation is most significant for firms likely to face asymmetric information problems [see Hubbard (1998) for a survey on this topic.] In their study of a panel of 179 firms in high-tech industries, Himmelberg and Pertersen (1994) find a large and significant relationship between R&D investment and internal cash flow, suggesting that internal financing conditions are an important determinant of R&D intensity. Hall (1992) provides similar evidence by examining a broader panel of U.S. manufacturing firms,

Although the importance of financing conditions in determining R&D investment is widely recognized, no empirical effort has been made to simultaneously study this factor with other determinants of R&D investment. Future research should make this attempt, given the fairly convincing evidence of the importance of financing conditions in the determination of industrial R&D investment.

2.4.5 Adjustment Costs and Investment Theory

In a study concerned with the firm's R&D investment behavior, a background discussion on the theory of investment seems in order. The theory of investment has occupied a large number of research programs. Hence, I will not attempt to review the entire literature as it is impossible and unnecessary in the present dissertation (an excellent survey can be found on the History of Economic Thought website: http://cepa.newschool.edu/het/essays/capital/invest.htm). Rather, I will focus on the neoclassical theory of investment which provides the foundation for the theoretical model proposed in this dissertation.

2.4.5.1 The Neoclassical Theory of Investment

In their early attempt to derive a new theory of capital and investment from the neoclassical marginal productivity theory, economists were frustrated by the difficulty of dealing with the flow variable, investment. The frustration is manifest in a passage by Haavelmo:

> "What we should reject is the naive reasoning that there is a 'demand schedule' for investment which would be derived from a classical scheme of producers' behavior in maximizing profits. The demand for investment

> cannot simply be derived from the demand for capital. Demand for a finite addition to the stock of capital can lead to any rate of investment, from almost zero to infinity, depending on the additional hypothesis we introduce regarding the speed of reaction of capital-users."
>
> — Haavelmo (1960, pp: 216)

To tackle this problem, Jorgensen (1963) derives the optimal capital stock from the firm's intertemporal optimization problem, and assumes that ordered investment goods are delivered in separate periods of time, when desired capital stock transfers from one equilibrium to another in response to changes in the decision environment. It follows from the latter assumption that the demand for investment is equal to the sum of weighted capital stocks in the past periods.

Jorgenson's *ad hoc* treatment of investment demand, which has been highly controversial, is forced by the fact that the first order conditions of the firm's intertemporal optimization problem are independent of investment. While the demand equations for capital and other variable inputs can be uniquely determined, the demand for investment is equal to either zero or infinity: it will equal zero if the demand for capital stays the same, and infinity if extra capital (or disinvestment) is needed immediately. Delivery lags are introduced to prolong the disequilibrium phase so that investment occurs every period.

Eisner and Strotz (1963) propose a more constructive solution to this problem by introducing the notion of adjustment costs. Suppose that a new long-run equilibrium in response to changes in the business environment entails an increase in capital. The new equilibrium will not come about instantaneously. Rather, it will take a certain number of periods before the final adjustment can be made to reach the new equilibrium. Various states of this adjustment process constitute the short- and intermediate-run equilibria, and the length of this adjustment process depends on the fixity of capital. Adjustment costs are specified as an *increasing* and *convex* function of the investment rate, so that a continuous adjustment path can be determined by the profit optimization problem of the firm.

The argument that there are adjustment costs in investment is further extended by Lucas (1967) and Treadway (1971) to include *n* fixed factors. In their models, the firm's decision rule is governed by a system of differential equations, the solution of which provides the demand function of investment goods. Lucas (1967) also shows that the demand function so obtained coincides with, and therefore rationalize, the empirically robust flexible accelerator model.

The backbone of these models is obviously the assumption that adjustment costs increase with investment at an increasing rate. Why this is so, however, has been explained in different ways. Some authors argue that adjustment costs are internal to the firm. When new equipment is purchased, for example, it needs to be installed and workers need to be trained to work effectively on it before it can finally be used as a productive input. This "integration" process will distract resources from production, and the more rapid is this "integration," the more resources it will distract. Other authors argue that adjustment costs are external to the firm: a high investment rate will raise the prices of the investment goods and therefore increase investment cost.

Empirical evidence of the adjustment cost hypothesis is sparse. Most empirical studies estimate only the derived factor demand system [e.g., Bernstein and Nadiri (1989) and Epstein and Denny (1983).] As an exception, Lichtenberg (1988) directly estimates both a production function and a labor demand equation and finds that a dollar of expansion in investment causes a 35 cent reduction in current output, providing strong support for the internal adjustment cost hypothesis.

2.4.5.2 Adjustment Costs in R&D Investment

The general investment theory discussed above can be directly applied to R&D investment. Moreover, economists argue that R&D investment incurs relatively high adjustment costs in comparison with physical investment. The argument runs as follows. The majority of R&D investment goes to wage payment to engineers, scientists, and other research workers. On the one hand, newly hired research workers need to be trained before their skills can be productively applied to the firm's R&D programs. On the other hand, since a significant proportion of R&D capital accumulated by the firm is embodied in its research workers, firing these workers implies an immediate loss of R&D capital (Grabowski, 1968).

Empirical evidence on the existence of adjustment costs in R&D investment is even more scanty than in physical investment. Bernstein and Nadiri (1989) estimate the R&D and physical investment demand equations simultaneously in a dynamic factor demand system, which is derived from the firm's intertemporal cost minimization problem. In all the four industries studied, the authors find that R&D investment adjusted more slowly than physical investment. This, however, can not be interpreted as evidence of the existence of adjustment costs in R&D investment, because without the adjustment cost assumption, the demand system would be totally different from the one they estimate. Consequently, the dynamic factor demand system derived from the firm's intertemporal optimization problem does not permit a nested hypothesis testing for the adjustment cost assumption. The present dissertation estimates the knowledge production function along with the first-order condition of the firm's dynamic profit optimization problem. In the case of no adjustment costs in R&D investment, the production function and the first order condition of the dynamic profit optimization problem collapse exactly to their respective counterparts in a static model, in the spirit of David and Hall's. This approach thus permits a nested hypothesis test of the adjustment cost hypothesis.

To conclude, the neoclassical theory of investment provides a convenient framework to derive R&D investment demand. One may use this framework to construct a model that takes into account the important dynamic aspect of R&D investment, as well as other determinants of R&D investment such as market demand and technological opportunity. There exist some limitations in this approach, however. The adjustment cost model typically assumes perfect competition in both the input and output markets. Yet it is widely recognized in the industrial organization literature that the firm may use R&D investment as a strategic instrument to acquire market power. In this sense the value of research output should depend on the firm's investment decisions. I did not find any study that attempts to take into account this dynamic factor in an adjustment cost model of investment. The difficulty is that a departure from the price-taking competition assumption will not permit a closed-form reduction of the system of differential equations which governs the firm's decision rule.

Another restrictive assumption underlying the adjustment cost model is that the firm finances at a constant interest rate. This assumption apparently contradicts empirical evidence of the firm's dependence on internal cash flow to finance investment. As discussed previously, asymmetric information in capital markets generates an extra "lemon" premium for external financing. Consequently, the interest rate schedule should jump up at the point where investment exceeds the firm's internal financing capacity. I did not find any study incorporating finance conditions into an adjustment cost model. Considering the amassed evidence of the firm's dependence on internal financing resources, there is a strong need for future research on this problem.

Chapter 3: Theoretical Model

The model presented in this Chapter describes knowledge spillovers among three life-science fields–biology, medicine, and agriculture–and public-private interactions in the scientific labor market. The public sector conducts research in all three life-science fields, but the private sector is engaged only in downstream medical and agricultural research, not basic biological research[1]. The thrust of the model is to understand R&D investment decisions in the medical and agricultural industries, and in particular, how these decisions may be influenced by the public sector's investment policy in life-science research.

The model is a dynamic generalization of that introduced by David and Hall (2000). Their model incorporates two effects public-sector R&D investment policy may have on the private sector's R&D investment. The first is a positive effect: research in the public sector may improve the private sector's productivity in knowledge generation. The second is a negative effect: since the two sectors presumably compete for the same input factor, scientific labor, an increase in R&D investment by the public sector may boost the wage of scientific labor, reducing private sector investment *ceteris paribus*. Hence, whether public R&D investment "crowds in" or "crowds out" private R&D investment hinges on the relative magnitudes of the spillover effect and wage effect, which in turn are determined by the technology and supply of scientific labor.

---

[1] The assumption is driven by the failure to identify industrial investment in biological research (see Section 5.2 for construction of industrial R&D investment series.)

Here David and Hall's model is reconsidered in the case in which "internal adjustment costs" exist in R&D investment. When internal adjustment costs exist in R&D investment, a rapid change in R&D investment is penalized by a reduction of knowledge output. Hence, it pays to approach the optimal level of investment gradually (see Section 2.4.5 for more about adjustment costs in investment). With the introduction of adjustment costs, the private sector's demand for scientific labor, or R&D investment, is the solution to a dynamic profit maximization problem. In the present dissertation, the decision problem is solved under both the static and quasi-rational expectations hypotheses, while the equilibrium concept adopted is the "rational expectations equilibrium" introduced by Lucas and Prescott (1971). It can be shown that without adjustment costs this dynamic model reduces to David and Hall's static model, and that under the static expectations hypothesis, the long-run equilibrium coincides with the equilibrium in David and Hall.

The specification of quadratic technology and linear supply function allows one to analytically solve the dynamic optimization problem and derive investment demand and knowledge output elasticities. Such results provide the theoretical basis for my empirical examination of the complementarity-vs.-substitutability hypothesis and the demand-pull-vs.-technology-push debate. In addition, I derive the formula for the private rate of return in a public investment plan, which may be used to evaluate the worth to the industry of one dollar of public investment.

3.1 Statement of the Model

Assume that the public sector conducts research in all three life-science fields, while there are only two research industries in the private sector, one engaged in medical research and another agricultural research. Knowledge may spill over from one field to another and from one sector to another. The mechanisms of the two sub-models of the medical and agricultural research industries are symmetric: substituting the subscript med for ag and ag for med in the following equations gives us the sub-model of the agricultural research industry. Hence, only the sub-model of the medical research industry is presented here.

Consider a medical research industry in which a finite number of identical competitive firms use a single input, scientific labor, to produce a single output, knowledge. The public sector and private sector compete for the same input factor, scientific labor for medical sciences. Let $w_{med,t}$ be the wage of scientific labor and $h_{med,t}$ be the supply of scientific labor. The (inverse) supply function is defined to be linear:

$$w_{med,t} = \alpha_1 + \alpha_2 h_{med,t}, \tag{3.1}$$

where α_1 and α_2 are parameters of the supply function. As usual, assume that a higher wage induces more supply of labor, namely, $\alpha_2 > 0$.

Let $p_{med,t}$ be the price of knowledge output in medical sciences. Research firms are competitive in the input and output markets, i.e., they take the output price sequence $\{p_{med,t}\}$ and wage sequence $\{w_{med,t}\}$ as given. One may consider the research industry as a competitive one with a single representative firm. I assume a quadratic knowledge production function with internal adjustment costs:

$$\begin{aligned} y^{prv}_{med,t} &= f(h^{prv}_{med,t-1}, h^{prv}_{med,t}, h^{pub}_{med,t}, h^{pub}_{bio,t}, h^{pp}_{ag,t}) \\ &= \beta_1 + (\beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t}) h^{prv}_{med,t} \\ &\quad + 0.5\beta_3 (h^{prv}_{med,t})^2 + 0.5\beta_4 (h^{prv}_{med,t} - h^{prv}_{med,t-1})^2, \end{aligned} \tag{3.2}$$

where $h^j_{i,t}$ denotes the amount of scientific labor engaged in knowledge production in the field i and the sector j at time t. Subscript $i = med, bio, ag$ stands for the three life science fields: medicine, biology, and agriculture, respectively. Superscript $j = pub, prv, pp$ represents the public sector, the private sector, and the combination of the two, respectively. β_k, $k = 1, 2, \cdots 7$, are the parameters of the knowledge production function f. Assume that $\beta_3 < 0$ and $\beta_4 < 0$, which guarantees the function f is strictly concave in $(h^{prv}_{med,t,} h^{prv}_{med,t-1})'$. Assumption $\beta_4 < 0$ implies that rapid change in R&D investment (employment of scientific labor) incurs adjustment costs in the form of a reduction of knowledge output. When $\beta_5 > 0$ there exist positive knowledge spillovers from the public sector's medical research to the private sector's: knowledge generated by the public sector helps improve the private sector's productivity. The interpretations of β_6 and β_7 are analogous.

The representative firm's R&D investment cost normalized by output price at time t is $\frac{w_{med,t}}{p_{med,t}} h^{prv}_{med,t}$. The return to the investment at time t thus equals to knowledge output (3.2) minus the normalized investment cost. The representative firm chooses contingency plans for $h^{prv}_{med,t}$ to maximize the expected present value of returns:

$$v^{prv}_{med,0} = E_0 \sum_{t=0}^{+\infty} r^t [\beta_1 + (\beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t} - \frac{w_{med,t}}{p_{med,t}}) h^{prv}_{med,t} \\ + 0.5\beta_3 (h^{prv}_{med,t})^2 + 0.5\beta_4 (h^{prv}_{med,t} - h^{prv}_{med,t-1})^2], \tag{3.3}$$

where r is the discount factor obeying $0 < r < 1$. The operator E_t is defined by $E_t x = Ex \mid \Omega_t$, where E is the mathematical expectation operator, and Ω_t is the information set available to the representative firm at time t. Assume that Ω_t includes at least the history of the variables, $h^{pub}_{med}, h^{pub}_{bio}, h^{pp}_{ag}, p_{med}$, and w_{med}, up to time t, and the history of h^{priv}_{med} up to time $t-1$. Note that the representative firm behaves as if the wage is independent of its own decision about how much scientific labor to employ despite the fact that its own demand for scientific labor affects the wage.

The solution to the problem (3.3) gives the representative firm's demand for scientific labor, or R&D investment. The demand-supply equilibrium in the scientific labor market must be defined to proceed. Following Lucas and Prescott (1971), the equilibrium pair of sequences $\{\overline{w}_{med,t}\}$ and $\{\overline{h}^{prv}_{med,t}\}$ satisfy the market clearing condition:

$$\overline{w}_{med,t} = \alpha_1 + \alpha_2 (\overline{h}^{prv}_{med,t} + h^{pub}_{med,t}), \tag{3.4}$$

and the condition that the sequence $\{\overline{h}^{prv}_{med,t}\}$ maximize (3.3).

3.2 Equilibrium Solution: Intuition

Before proceeding to solve for the equilibrium representation of the R&D investment sequence, it seems useful to provide an intuitive description of the equilibrium solution. For illustrative purposes, we consider the case where the initial value of investment is zero and the decision environment remains unchanged throughout. But the intuition gained in this thought experiment can be readily adapted to help understand the equilibrium solution when exogenous variables are nonconstant or even stochastic.

In problem (3.3) the expression in the bracket is the return to the representative firm at period t. Denote this return function by $R(h_{med,t}^{prv}, h_{med,t-1}^{prv}, \bullet)$. Note that without adjustment costs, i.e., if $\beta_4 = 0$, the return function is still quadratic. Let this function be $\hat{R}(h_{med,t}^{prv}, \bullet)$. Assume that all the other exogenous variables are constant and that a steady state, or long run equilibrium, of the decision variable $h_{med,t}^{prv}$ exists. By definition we have $h_{med,t}^{prv} = h_{med,t-1}^{prv} = \hat{h}_{med}^{prv}$ at the long-run equilibrium. Therefore, there will be no adjustment costs once the long-run equilibrium is achieved. It turns out that the long-run equilibrium of R&D investment $\hat{h}_{med}^{prv}$ must maximize the quadratic function $\hat{R}(h_{med,t}^{prv}, \bullet)$. Now the problem (3.3) may be interpreted as finding the optimal path for $h_{med,t}^{prv}$ to reach the long-run target $\hat{h}_{med}^{prv}$ — or the maximizer of the function $\hat{R}(h_{med,t}^{prv}, \bullet)$ — over an infinite horizon. Figure 3.1 graphically depicts this approaching process. In this figure, the solid

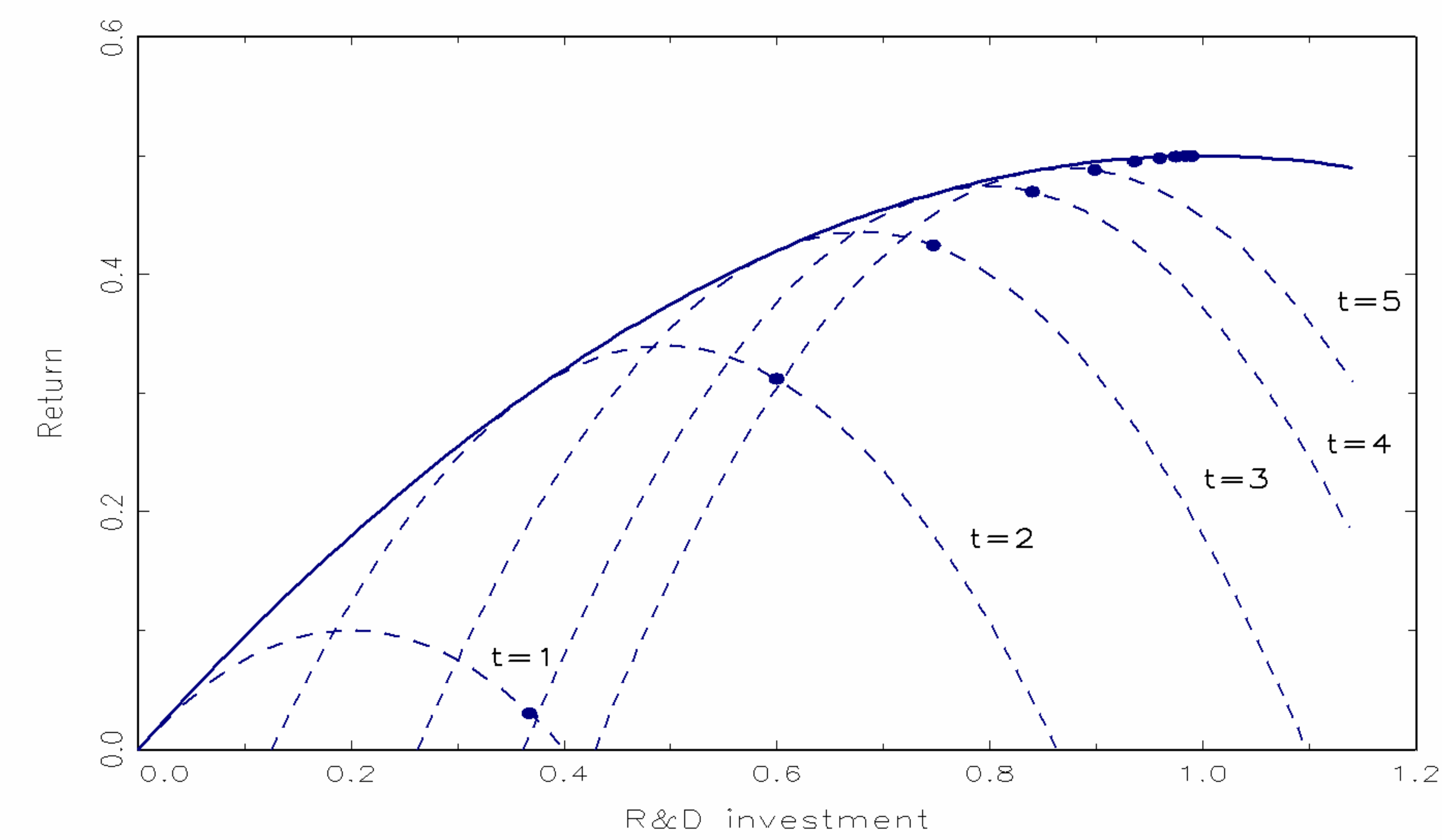

Fig. 3.1 Optimal path of R&D investment reaching the long-run equilibrium

curve represents the function $\hat{R}(h^{prv}_{med,t},\bullet)$ and the dashed curves are the return functions in each of the first five periods, assuming the initial value of R&D investment is zero. The solid circles are the optimal loci of R&D investment and return. Evidently, the optimal path of R&D investment is to approach the long-run target gradually. Moreover, at each period the optimal R&D investment is above the level that maximizes that period's return. Because of adjustment costs, the return function at period $t+1$ depends on R&D investment at period t. Specifically, the return at period $t+1$ is a positive function of R&D investment at period t. Although a higher level of R&D investment will incur more adjustment costs at the current period, it will reduce the adjustment costs at the nest period. It follows that the optimal R&D investment at period t must be such that no extra gain can be achieved by allocating money from period t to $t+1$ and vice versa. This is the essence of the first order condition of problem (3.3) as we will see soon.

In the case where there are no adjustment costs in R&D investment, we have $\hat{R}(h^{prv}_{med,t},\bullet)=R(h^{prv}_{med,t},h^{prv}_{med,t-1},\bullet)$. The optimal path for R&D investment is to jump to the equilibrium target that maximizes function $\hat{R}(h^{prv}_{med,t},\bullet)$, and to stay there until the decision environment changes. In that case, the dynamic model developed here will reduce to David and Hall's static model.

3.3 Equilibrium Solution under the Quasi-Rational Expectations Hypothesis

In this Section, I solve for the equilibrium representation of the R&D investment sequence $\{h_{med,t}^{prv}\}$, using the method used by Sargent (1987, pp. 399-404). The stochastic Euler equation for problem (3.3) is:

$$\begin{aligned} &\beta_2 + \beta_5 h_{med,t}^{pub} + \beta_6 h_{bio,t}^{pub} + \beta_7 h_{ag,t}^{pp} - \frac{\overline{w}_{med,t}}{p_{med,t}} + \beta_3 \overline{h}_{med,t}^{prv} \\ &+ \beta_4 (\overline{h}_{med,t}^{prv} - \overline{h}_{med,t-1}^{prv}) = r\beta_4 (E_t \overline{h}_{med,t+1}^{prv} - \overline{h}_{med,t}^{prv}), \\ &t = 0,1,2\cdots \end{aligned} \tag{3.5}$$

and the transversality condition is:

$$\begin{aligned} \lim_{t\to\infty} r^t E_0 \{ [\beta_2 + \beta_5 h_{med,t}^{pub} + \beta_6 h_{bio,t}^{pub} + \beta_7 h_{ag,t}^{pp} - \frac{\overline{w}_{med,t}}{p_{med,t}} \\ + \beta_3 \overline{h}_{med,t}^{prv} + \beta_4 (\overline{h}_{med,t}^{prv} - \overline{h}_{med,t-1}^{prv})] \overline{h}_{med,t}^{prv} \} = 0 \end{aligned} \tag{3.6}$$

On the left hand side of the Euler equation is the marginal cost (benefit) of R&D investment (disinvestment) at period t, while on the right hand side is the marginal benefit (cost) of R&D investment (disinvestment) at period $t+1$. The transversality condition can be interpreted as the requirement that the present discounted value of R&D investment at period t tend to zero as t tends to infinity.

As is well known, the quadratic dynamic optimization problem can be solved explicitly because the first order condition and transversality condition are linear in exogenous variables. In the current case, however, while the Euler equations (3.5) and transversality conditions (3.6) are linear in exogenous variables $h_{med,t}^{pub}$, $h_{bio,t}^{pub}$, and $h_{ag,t}^{pp}$, they are nonlinear in $p_{med,t}$. The unusual presence of nonlinearity in a linear-quadratic optimization problem is due to the fact that wage $w_{med,t}$ is treated as an endogenous variable in the present model, whereas conventionally real wage $\frac{w_{med,t}}{p_{med,t}}$ is treated as an exogenous variable (Sargent, 1978). To get around this difficulty, one may consider an approximation situation in which the representative firm has static expectations for output price $p_{med,t}$, or the representative firm believes that future prices $p_{med,t+j}$ have identical degenerate distribution functions localized at $p_{med,t}$, i.e., $p_{med,t+j} = p_{med,t}$ with probability 1 for all $t > 0$ and $j > 0$.

Substitute from (3.1) into (3.5) to obtain

$$\begin{aligned} &\beta_2 - \frac{\alpha_1}{p_{med,t}} + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t} + (\beta_3 - \frac{\alpha_2}{p_{med,t}})\bar{h}^{prv}_{med,t} \\ &+\beta_4(\bar{h}^{prv}_{med,t} - \bar{h}^{prv}_{med,t-1}) = r\beta_4(E_t\bar{h}^{prv}_{med,t+1} - \bar{h}^{prv}_{med,t}), \qquad t = 0,1,2\cdots \end{aligned} \tag{3.7}$$

Rearrange to obtain

$$E_t\bar{h}^{prv}_{med,t+1} - \theta\bar{h}^{prv}_{med,t} + \frac{1}{r}\bar{h}^{prv}_{med,t-1} = z_t, \qquad t = 0,1,2\cdots \tag{3.8}$$

where

$$\theta = 1 + \frac{1}{r} + \frac{\beta_3}{r\beta_4} - \frac{\alpha_2}{r\beta_4 p_{med,t}},$$

$$z_t = \frac{1}{r\beta_4}[\beta_2 - \frac{\alpha_1}{p_{med,t}} + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t}].$$

Note that (3.8) is a system of second-order stochastic difference equations. Following Sargent (1987, pp. 391-396), solve it forward using the unstable root to impose the transversality condition (see appendix A.1 for details), obtaining:

$$\bar{h}^{prv}_{med,t} = \lambda_1 \bar{h}^{prv}_{med,t-1} - \frac{1}{\lambda_2}\sum_{i=0}^{\infty}(\frac{1}{\lambda_2})^i E_t z_{t+i}, \quad t = 0,1,2\cdots \tag{3.9}$$

where $\bar{h}^{prv}_{med,-1} = h^{prv}_{med,-1}$, $\lambda_1 = 0.5(\theta - \sqrt{\theta^2 - 4/r})$, and $\lambda_2 = 1/(\lambda_1 r)$. And λ_1 and λ_2 obey $0 < \lambda_1 < 1 < \frac{1}{r} < \lambda_2$.

Expression (3.9) describes the representative firm’s R&D investment at time t as a function of the investment at time $t-1$ and the expected values of the exogenous state variables conditional on the information set at time t. If the probability distribution functions of these exogenous variables are known to the representative firm, the remaining task is to take the conditional expectations of the exogenous variables and substitute them into expression (3.9). If the representative firm does not know the probability distribution functions of the exogenous variables, one may assume that it forecasts the future values of the exogenous variables in an “optimal” way. As is often assumed in the literature, the forecasting criterion is to minimize the mean squared error. Suppose, say, that the exogenous variables evolve as markov processes:

$$\begin{aligned} h^{pub}_{med,t} &= c_1 + \rho_1 h^{pub}_{med,t-1} + \varepsilon_{1t}, \\ h^{pub}_{bio,t} &= c_2 + \rho_2 h^{pub}_{bio,t-1} + \varepsilon_{2t}, \\ h^{pp}_{ag,t} &= c_3 + \rho_3 h^{pp}_{ag,t-1} + \varepsilon_{3t}, \end{aligned} \tag{3.10}$$

where ε_{it}, $i = 1,2,3$ are forecast innovations. Applying the Wiener-Kolmogorov formula (Sargent, 1987, pp. 292) to (3.10) gives

$$P_t h^{pub}_{med,t+i} = \frac{1-\rho_1{}^i}{1-\rho_1} c_1 + \rho_1{}^i h^{pub}_{med,t}, \tag{3.11}$$

$$P_t h^{pub}_{bio,t+i} = \frac{1-\rho_2{}^i}{1-\rho_2} c_2 + \rho_2{}^i h^{pub}_{bio,t},$$
$$P_t h^{pp}_{ag,t+i} = \frac{1-\rho_3{}^i}{1-\rho_3} c_3 + \rho_3{}^i h^{pp}_{ag,t},$$

where $P_t x_{t+i}$ is the linear least square projection of x_{t+i} on the space spanned by $\{x_t, x_{t-1}, x_{t-2} \cdots\}$. Substitute the linear least square forecasts (3.11) for the conditional expectations in (3.9) to obtain

$$\begin{aligned}\bar{h}^{prv}_{med,t} = \lambda_1 \bar{h}^{prv}_{med,t-1} - \frac{1}{r\beta_4}[\frac{\beta_2 - \alpha_1/p_{med,t}}{\lambda_2 - 1} \\ + \frac{(\beta_5 - \alpha_2/p_{med,t})(c_1/(\lambda_2 - 1) + h^{pub}_{med,t})}{\lambda_2 - \rho_1} \\ + \frac{\beta_6(c_2/(\lambda_2 - 1) + h^{pub}_{bio,t})}{\lambda_2 - \rho_2} + \frac{\beta_7(c_3/(\lambda_2 - 1) + h^{pp}_{ag,t})}{\lambda_2 - \rho_3}],\end{aligned} \tag{3.12}$$

for $t = 0,1,2\cdots$ Expression (3.12) is the equilibrium solution to the representative's decision problem in its *feedback* form: the representative firm's equilibrium R&D investment depends only on the current realizations of the state variables, $\bar{h}^{prv}_{med,t-1}$, $h^{pub}_{med,t}$, $h^{pub}_{bio,t}$, $h^{pp}_{ag,t}$, and $p_{med,t}$, but not on the previous history of these variables. Note also that (3.12) demonstrates the essence of Lucas' critique: the parameters on the exogenous state variables will change if the policy interventions change the mode in which these variables evolve. Therefore, correct policy simulations must be based on a "structural" model. In the present context, the appropriate structural model should include both technology and supply of scientific labor, as well as a model describing the evolution of the exogenous variables.

Feedback solution (3.12) permits convenient derivation of the elasticities of demand for scientific labor[2]. The demand elasticity for private medical R&D investment with respect to a shock in public-sector medical R&D investment is

$$e_{m,m}^{d} = -\frac{(\beta_5 - \alpha_2 / p_{med,t}) h_{med,t}^{pub}}{r\beta_4(\lambda_2 - \rho_1)\bar{h}_{med,t}^{prv}}. \tag{3.13}$$

Note that, by assumption, $0 < r < 1$, $\lambda_2 > 1 > \rho_1$, and $\beta_4 < 0$. It follows that if $\beta_5 > \alpha_2 / p_{med,t}$ ($\beta_5 < \alpha_2 / p_{med,t}$), public-sector R&D investment will "crowd in (out)" private-sector R&D investment since the negative "wage effect" is dominated by (dominates) the positive "spillover effect."

The demand elasticity with respect to a shock in public investment in biological research $h_{bio,t}^{pub}$ is:

$$e_{m,b}^{d} = -\frac{\beta_6 h_{bio,t}^{pub}}{r\beta_4(\lambda_2 - \rho_2)\bar{h}_{med,t}^{prv}}. \tag{3.14}$$

In contrast with the previous case, whether public investment in biological research crowds in or crowds out private investment in medical research is determined solely by technology, and specifically by the sign on β_6. This is due simply to the distinction we have made between biological and medical scientific labor, which implies that public-sector demand for biological scientific labor has no wage effect on private-sector demand for medical scientific labor.

---

[2] All results of comparative statics analysis in this Chapter and the appendices are based on the assumption that no knowledge spillovers exist between medical and agricultural research, i.e., $\beta_7 = 0$ in both the medical and agricultural models. If $\beta_7 > 0$ in both industries, R&D investment in one industry becomes endogenous to another industry, and the equilibrium has to be redefined by integrating the two sub-models.

The elasticity of private-sector R&D expenditure with respect to knowledge output price $p_{med,t}$, denoted by $e^d_{m,p}$, is algebraically much more complicated than the previous two because the investment demand (3.12) is highly nonlinear in $p_{med,t}$ (note that both λ_1 and λ_2 are nonlinear functions of $p_{med,t}$.) The expression for $e^d_{m,p}$ therefore is removed to Appendix A.2. Note that comparing the demand elasticity with respect to output price and those with respect to public R&D investments provides a simple way to examine the demand-pull vs. technology-push debate (see Section 2.4.1 for a review of this literature) in the context of life science research. A comparison of these elasticities can tell us the relative importance to industrial innovations of technological opportunities created by public-sponsored research, and the market incentives derived from consumers' increasing demand for medical products.

The knowledge output supply function $y(h^{prv}_{med,t-1}, h^{pub}_{med,t}, h^{prv}_{bio,t}, h^{pp}_{ag,t}, p_{med,t})$ can be obtained by substituting the equilibrium R&D investment (3.12) into the knowledge production function (3.2). The elasticities of knowledge output supply with respect to shocks in the exogenous variables then can be computed readily (see appendix A.2.)

3.4 Equilibrium Solution under the Static Expectations Hypothesis

In the previous Section, the model has been solved under the quasi-rational expectations hypothesis (QREH), under which the subjective expectation coincides with the mathematical expectations over all exogenous variables except the knowledge output price, which is believed to stay constant in all future periods. In the present Section, I solve the model instead under the static expectations hypothesis, i.e., the representative firm naively believes that the current values of all exogenous variables will stay constant in all future periods. The static expectations hypothesis (SEH) evidently makes it much easier to solve for the equilibrium solution, and the approach has its own merit for analyzing policy questions that are interesting in the present context. In addition, a useful result is that the long-run equilibrium under the SEH reduces to the equilibrium in David and Hall's model.

A convenient way to solve for the equilibrium under the SEH is to replace $E_t z_{t+i}$ by z_t in (3.10). The intuition is simple: under the SEH, the representative firm believes that all the exogenous variables have degenerate distribution functions localized at their current values, meaning $E_t z_{t+i} = z_t$. Now we have:

$$\begin{aligned}\tilde{h}^{prv}_{med,t} = \lambda_1 \tilde{h}^{prv}_{med,t-1} - \frac{1}{r\beta_4(\lambda_2 - 1)}[\beta_2 - \frac{\alpha_1}{p_{med,t}} + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h^{pub}_{med,t} \\ + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t}],\end{aligned} \tag{3.15}$$

$t = 0,1,2\cdots$

At the steady state or long-run equilibrium, where $\bar{h}^{prv}_{med,t+j} = \bar{h}^{prv}_{med,t+j-1} = \hat{h}^{prv}_{med,t}$, we have

$$\hat{h}^{prv}_{med,t} = \frac{1}{\alpha_2 / p_{med,t} - \beta_3}[\beta_2 - \frac{\alpha_1}{p_{med,t}} + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t}], \tag{3.16}$$

$t = 0,1,2\cdots$

Now consider a one-field model like that of David and Hall (2000), by letting $\beta_6 = 0$ and $\beta_7 = 0$. In that case, the long-run equilibrium solution of R&D investment in the scientific labor market (3.16) is essentially that in the D-H model when technology is quadratic (but the validity of this claim is not restricted to quadratic technology; for a more general model, see Wang, et. al., unpublished manuscript.)

Now we can formally reinterpret the dynamic decision problem as the chasing-target problem described in Section 3.2. Rewrite equation (3.15) as

$$\tilde{h}^{prv}_{med,t} = \lambda_1 \tilde{h}^{prv}_{med,t-1} + (1-\lambda_1)\hat{h}^{prv}_{med,t}, \quad t = 0,1,2\cdots \tag{3.17}$$

Because rapid change in R&D investment incurs adjustment costs, it pays to approach the optimal level of investment gradually. The optimal level of investment at each period is a weighted sum of the level of investment in the last period and the target level of investment. The adjustment parameter $(1-\lambda_1)$ simply tells how fast investment approaches the long-run target. To see this more clearly, rewrite (3.17) as

$$\tilde{h}^{prv}_{med,t} - \tilde{h}^{prv}_{med,t-1} = (1-\lambda_1)(\hat{h}^{prv}_{med,t} - \tilde{h}^{prv}_{med,t-1}), \quad t = 0,1,2\cdots \tag{3.18}$$

The adjustment parameter represents the proportion of adjustment made within a single year in the process of approaching the long-run target. If adjustment costs are high (low) relative to investment benefits, the adjustment process will be comparatively slow (fast). In the extreme case of no adjustment costs, we have $\tilde{h}^{prv}_{med,t} = \hat{h}^{prv}_{med}$; investment reaches the steady state immediately, the situation described by the D-H model.

The short-run and long-run elasticities can be readily obtained by respectively taking derivatives of (3.15) and (3.16) with respect to the exogenous variables in question. And the equilibrium wage and output supply can be obtained by substituting (3.15) and (3.16) into the input supply function (3.4) and knowledge production function (3.2). The short-run and long-run elasticities of knowledge supply and equilibrium wage then can be derived (see appendix A.3 for these results.)

Finally, the equilibrium solution under the SEH may help to answer the following policy question. Assume that government plans to increase R&D investment in medical sciences by $1 million at period 0 and credibly maintain the new level of investment permanently. The policy is announced to the public before being put into effect. What is the shadow value of this investment plan to the medical research industry? In this scenario the equilibrium solution solved under the QREH reduces to the equilibrium solution solved under the SEH because the representative firm believes that the exogenous variables will stay constant in the future. The shadow value of public investment in medical sciences is the first-order derivative of the value function of problem (3.3) with respect to that investment variable. The marginal cost of the investment plan is simply $\$1/(1-r)$ million. The shadow value divided by the marginal cost gives a measure of the "private rate of return" on public investment in medical sciences (see Appendix A.4 for details.)

Chapter 4: Estimation Method and Econometric Model

This Chapter considers alternative approaches to estimating the model of R&D investment developed in Chapter 3. The model has been motivated to help us understand the influences which public R&D investment policy have on private sector R&D investment in the life sciences, so that the effectiveness of current policy can be evaluated. To this end, one should estimate a "structural" model in order to avoid the Lucas critique. In the present context, an appropriate structural model permitting meaningful policy evaluations should include the demand and supply of scientific labor (R&D investment), and technology. Two popular approaches have appeared in the literature for estimating rational expectations models. I begin by reviewing these two approaches.

4.1 MLE versus GMM

The first estimation strategy is to solve the model explicitly and apply the Maximum Likelihood Estimation (MLE) procedure. By imposing strict assumptions on the model and on the stochastic properties of exogenous variables, one may solve for an equilibrium representation of the endogenous variables as in Section 3.3. The objective function usually is assumed to have the quadratic form, which leads to first-order conditions linear in the exogenous variables. Without such a stringent assumption, dynamic rational expectations models typically do not have closed-from solutions. The solution to the dynamic rational expectations model usually involves the expectations of future values of exogenous variables, conditional on the current information set.

Consequently, stochastic properties of exogenous variables need to be specified in order to estimate the expectation terms. The resulting system to be estimated consists of the economic model, the model that describes the evolution of the exogenous variables, and a set of cross-equation restrictions implied by the rational expectations hypothesis. A full information estimation procedure such as MLE then can be applied to estimate this system [e.g., (Hansen and Sargent, 1980) and (Sargent, 1978).]

Alternatively, one may focus on estimating the stochastic Euler equation using the Generalized Method of Moments (GMM) procedure introduced by Hansen (1982) [for applications of this method see, e.g., (Hansen and Singleton, 1982) and (Pindyck and Rotemberg, 1983)]. That approach does not require an explicit representation of the rational expectations equilibrium. Hence, one may assume flexible functional forms for technology or preferences and impose fairly weak assumptions on the stochastic exogenous variables, yet still be able to fully identify the structural model parameters. The price of the simplicity, however, is that some restrictions implied by the theoretical model–e.g., the transversality condition–are ignored in estimation. Estimation efficiency loss may result.

In the specific case of investment models with adjustment costs, an important advantage of GMM over MLE is that the former approach permits simple *nested* hypothesis testing of the adjustment cost hypothesis. If in the GMM approach one imposes on Euler equation (3.5) the assumption of no adjustment costs, the technology and Euler equation in my dynamic model will reduce respectively to the technology and first-order condition of a static investment model without adjustment costs. However, one may not in the MLE approach impose the adjustment costs assumption on the final solution of the dynamic decision problem to obtain the solution of a corresponding static problem. This is because in solving the Euler equation as a set of difference equations, one implicitly imposes the adjustment cost assumption, and such a solution is valid only if the adjustment costs assumption holds [to see this, notice that β_4 cannot be equal to zero in expression (3.8).] Hence, by taking the MLE approach one loses the opportunity to test for the adjustment cost hypothesis in a simple way.

As shown in Section 3.3, the Euler equation in my model is nonlinear in output price despite an assumed quadratic technology. This is because, as explained above, wage is an endogenous variable in my model. To obtain an explicit expression of the equilibrium solution, I was forced in Chapter 3 to impose the "quasi-rational expectations" assumption. Altogether it seems that, in the current context, the GMM approach is preferable to the MLE. Hence, I adopt the GMM for econometric estimation.

4.2 GMM: A Review

Method of Moments estimation is based on the analogy principle that matches population moments with the appropriate sample moments. For example, let the population moment condition be $E(f(x_t,\theta^0))=0$, where x_t is the realization of a time series vector sequence at time t, θ^0 is the vector of parameters to be estimated, and $f(x_t,\theta^0)$ is a vector function of x_t and θ^0. The sample mean of $f(x_t,\theta^0)$ is

$$f_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} f(x_t, \theta), \tag{4.1}$$

a natural estimator of the population mean $E(f(x_t, \theta^0))$. Accordingly, the sample moment conditions are $f_T(\theta) = 0$. The estimator $\hat{\theta}_T$ which solves the moment condition provides an estimate of the true value θ^0. When the number of moment conditions is the same as the number of parameters to be estimated, one usually can find a unique solution for $\hat{\theta}_T$. When the number of moment conditions exceeds the number of parameters to be estimated, however, the system of equations are overidentified and there exists an infinite number of solutions for $\hat{\theta}_T$. In that case, the Generalized Method of Moments estimation introduced by Hansen (1982) is useful. The basic idea of GMM estimation is to minimize a quadratic form of the sample means,

$$J_T(\theta) = f_T(\theta)' W_T f_T(\theta), \tag{4.2}$$

with respect to θ, where W_T is a positive semidefinite weighting matrix. The GMM estimator of the true parameter θ^0 is

$$\tilde{\theta}_T = \arg\min_{\theta} J_T(\theta). \tag{4.3}$$

Since different weighting matrices may result in different GMM estimators, one needs to discover the optimal weighting matrix that leads to an efficient GMM estimator. Usually, a two-stage procedure is adopted. In the first stage an identity matrix, for example, is used as the weighting matrix. The estimation results of the first stage then are used to construct the optimal weighting matrix for the second stage.

The minimized value of the quadratic criterion (5.2), $J_T(\tilde{\theta})$, can be used to test the overidentifying restrictions when there are more moment conditions than parameters to be estimated. Hansen (1982) shows that sample size T times $J_T(\tilde{\theta})$ has an asymptotically chi-square distribution in which the degrees of freedom are the number of overidentifiying restrictions. This test is often called the Hansen J test, and $TJ_T(\tilde{\theta})$ accordingly the Hansen *J* statistic.

Specification tests also can be constructed with the aid of the Hansen *J* statistic. To test the validity of restrictions imposed on the parameters, a statistic can be constructed by subtracting the *J* statistic of the unrestricted model from that of the restricted model. The resulting statistic is distributed as chi-square in which the degrees of freedom equal the number of restrictions [see Ogaki (1993) for more detail].

4.3 Econometric Model and Moment Conditions

In order to fully identify all the parameters in the model developed in Chapter 3, I need to simultaneously estimate the knowledge production function, the Euler equation, and the inverse supply function of scientific labor. The econometric model for the medical research industry[3] can be obtained by defining the disturbances of the three equations as follows

$$\mu_{med,1t} = w_{med,t} - [\alpha_1 + \alpha_2 (h^{prv}_{med,t} + h^{pub}_{med,t})], \tag{4.4}$$

[3] To obtain the econometric model for the agricultural research industry, substitute the subscript "med" for "ag" and "ag" for "med" in each equation in this Section.

$$\mu_{med,2t} = y^{prv}_{med,t} - [\beta_1 + (\beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t}) h^{prv}_{med,t} + 0.5\beta_3 (h^{prv}_{med,t})^2 + 0.5\beta_4 (h^{prv}_{med,t} - h^{prv}_{med,t-1})^2], \tag{4.5}$$

where $y^{prv}_{med,t}$ denotes the medical industry's knowledge output; and

$$\mu_{med,3t} = \beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t} - \frac{w_{med,t}}{p_{med,t}} + \beta_3 h^{prv}_{med,t} + \beta_4 (h^{prv}_{med,t} - h^{prv}_{med,t-1}) - r\beta_4 (h^{prv}_{med,t+1} - h^{prv}_{med,t}). \tag{4.6}$$

Next, I need to define moment conditions for each of the three equations. The moment conditions crucially depend on the interpretations of the disturbances. We begin by equation (4.6). It follows immediately from stochastic Euler equation (3.5) that

$$E_t \mu_{med,3t} = 0. \tag{4.7}$$

Therefore, disturbance $\mu_{med,3t}$ is simply an expectation error that is independent of any variable in the representative firm's information set at period t, $\Omega_{med,t}$. In fact, expression (4.7) embodies the essence of the rational expectation thesis: whatever expectational errors made by economic agents are random errors uncorrelated with the information available to them. Let $z_{med,3t} = (h^{pub}_{med,t}, h^{pub}_{bio,t}, h^{pp}_{ag,t}, h^{prv}_{med,t-1}, h^{prv}_{med,t}, w_{med,t}, p_{med,t})'$. Since each element of vector $z_{med,3t}$ belongs to information set $\Omega_{med,t}$, by the law of iterated expectation we have the following moment conditions for equation (4.6):

$$E(z_{med,3t} \otimes \mu_{med,3t}) = 0. \tag{4.8}$$

The disturbance term in supply equation (4.4) can be interpreted as a labor supply shock. In order to identify equation (4.4) as a supply function rather than demand function, I need to find instruments that potentially are demand shifters yet not supply shifters. The candidates available include knowledge output price $p_{med,t}$, public-sector biological research investment $h^{pub}_{bio,t}$, and total agricultural R&D investment $h^{pp}_{ag,t}$. While knowledge output price is unequivocally a demand shifter but not a supply shifter, it is possible that R&D investments in biology and agriculture affect the supply of scientific labor in medical sciences. The reason is that scientific labor in biological and agricultural sciences may be, to some extent, substitutable for that in medical sciences. Hence, I define two sets of instruments for equation (4.4): $z^1_{med,1t} = (1, p_{med,t})'$ and $z^2_{med,1t} = (1, p_{med,t}, h^{pub}_{bio,t}, h^{pp}_{ag,t},)'$. The moment conditions for equation (4.4) therefore are one of the following two equations:

$$E(z^k_{med,1t} \otimes \mu_{med,1t}) = 0, \quad k = 1,2. \tag{4.9}$$

Finally, the disturbance term in production function (4.5) can be interpreted as a technological shock in knowledge production. Because, in the present study, knowledge output is measured by patent numbers, the disturbance may also be interpreted as a change in the frequency with which a patent application is awarded. We may call this an "institutional shock." The independent variables in production function (4.5) are chosen to be the instrumental variables:

$$z^1_{med,2t} = (1, h^{pub}_{med,t} h^{prv}_{med,t}, h^{pub}_{bio,t} h^{prv}_{med,t}, h^{pp}_{ag,t} h^{prv}_{med,t}, 0.5(h^{prv}_{med,t})^2, 0.5(h^{prv}_{med,t} - h^{prv}_{med,t-1})^2)'.$$

Note that because of the quadratic specification, the disturbance term can shift the production function up or down but can not affect its curvature. Consequently, the R&D

investment demand obtained from the first-order condition is independent of the disturbance term. This justifies the use of explanatory variables as instruments.

Additionally, one may assume that the disturbance term is independent of output price $p_{med,t}$ and wage $w_{med,t}$, leading to the alternative set of instruments:

$$z^2_{med,2t} = (1, h^{pub}_{med,t} h^{prv}_{med,t}, h^{pub}_{bio,t} h^{prv}_{med,t}, h^{pp}_{ag,t} h^{prv}_{med,t}, 0.5(h^{prv}_{med,t})^2, 0.5(h^{prv}_{med,t} - h^{prv}_{med,t-1})^2, p_{med,t}, w_{med,t})'.$$

Moment conditions for the production function (4.5) consequently would be one of the following two equations:

$$E(z^k_{med,2t} \otimes \mu_{med,2t}) = 0, \quad k = 1, 2. \tag{4.10}$$

Chapter 5: Data Construction

Table 5.1 contains all the data used in this dissertation.[4] R&D investments are measured by deflated R&D expenditures. Price indices used to deflate nominal R&D expenditure also serve as scientific labor wages, so that real expenditure is equivalent to labor quantity. Patent counts are used to measure knowledge output.

The National Science Foundation (NSF) Survey of Federal Funds for Research and Development formed the basis of the public-sector R&D expenditure data. Private-sector R&D expenditure data were constructed based on Standard & Poor's Compustat® database. The patent data were purchased from Thomson™. The Biomedical Research and Development Price Index, developed by the National Institutes of Health (NIH), was used to deflate medical and biological R&D expenditure, and the agricultural R&D deflator was developed by the U.S. Department of Agriculture, Economic Research Service (USDA/ERS) was adopted to deflate agricultural R&D expenditure. Annual data on knowledge output, that is patent, prices were developed by linking stock market data drawn from Compustat® to patent data obtained from Thomson™. Our data sample spans 1980 to 2004. Source data for private-sector R&D expenditure are unavailable before 1980. NSF has not updated its R&D expenditure data since 2003.

---

[4] Data construction described in this Chapter is part of a project funded by USDA/CSREES under NRI grant No. 2005-35400-15881. A national team of government researchers and university professors contributed to this project. As a research assistant of our research team, I participated in carrying out the original data collection plan, revising it as new problems arose, and designing new data procedures not covered in the original plan.

Table 5.1 Data

Year	*Private-sector R&D investment*		*Wage of Scientific Labor*		*Private-sector knowledge output*		*Knowledge output price*		*Public-sector R&D investment*		
	Millions of 2001 USD		Wage in 2001 = 100		# of patents		Millions of 2001 USD per patent		Millions of 2001 USD		
	Medicine	Agriculture	Medicine	Agriculture	Medicine	Agriculture	Medicine	Agriculture	Medicine	Agriculture	Biology
1980	5,019	3,382	37.48	38.72	3,265	3,574	3.77	0.32	4,889	2,453	5,799
1981	5,287	3,720	41.38	42.35	3,351	3,650	2.80	0.08	4,620	2,585	5,810
1982	6,008	3,737	44.93	45.59	2,961	2,973	3.72	0.98	4,781	2,607	5,320
1983	6,749	3,755	47.72	48.18	2,609	2,573	4.17	1.73	4,901	2,604	5,567
1984	6,953	4,147	50.54	50.84	3,223	3,045	3.45	2.28	5,423	2,416	5,823
1985	8,184	4,430	53.37	53.56	3,199	3,354	6.19	2.97	5,811	2,526	6,343
1986	10,471	4,742	55.61	56.44	3,203	3,138	10.32	6.96	5,786	2,530	6,357
1987	11,763	4,809	58.57	59.85	3,712	3,482	9.42	5.99	6,352	2,359	7,087
1988	13,694	5,193	61.51	62.17	3,628	3,375	9.73	7.72	6,779	2,336	7,302
1989	14,994	5,737	64.70	65.55	4,910	4,223	12.42	13.97	7,049	2,471	7,491
1990	18,177	5,566	68.21	68.95	4,765	3,904	14.56	14.32	7,164	2,441	7,494
1991	19,553	5,631	71.51	71.77	4,957	3,865	20.19	15.84	7,636	2,451	8,206
1992	22,204	5,889	74.66	74.00	5,426	3,993	13.79	13.83	7,398	2,412	7,423
1993	27,851	5,989	77.21	77.34	5,837	3,653	12.66	11.68	7,906	2,347	7,737
1994	29,157	6,422	80.20	79.94	5,157	3,611	12.28	9.96	8,279	2,360	7,388
1995	32,472	6,902	82.98	82.32	5,689	3,547	20.13	14.86	8,192	2,363	7,392
1996	36,893	7,525	85.10	84.67	6,494	3,696	24.88	17.31	8,193	2,314	7,804
1997	39,325	7,704	87.47	87.22	7,894	3,557	30.71	20.56	8,816	2,352	7,893
1998	45,013	8,859	90.43	89.90	9,809	4,153	39.52	18.47	9,196	2,462	8,135
1999	44,724	8,124	93.13	92.92	10,248	4,651	34.71	14.04	10,039	2,498	9,005
2000	49,622	6,836	96.57	95.96	9,693	4,662	42.11	15.02	7,789	2,579	12,297
2001	59,012	6,233	100.00	100.00	13,032	5,366	31.55	12.48	9,861	2,697	13,719
2002	62,554	6,066	103.19	103.59	23,164	7,791	15.91	8.86	9,654	2,689	15,069
2003	67,447	6,218	107.35	107.31	31,083	8,210	13.98	9.95	9,341	2,620	18,057
2004	69,770	5,955	111.21	111.17	30,323	8,717	11.42	8.93	.		.

5.1 Public-Sector R&D Expenditures

Data on public-sector R&D expenditure in the three life-science fields were constructed on the basis of the National Science Foundation Survey of Federal Funds for Research and Development (NSFSFFRD). Federal agencies and their subdivisions submit data in response to that survey each year. Reported R&D expenditures include R&D obligations and outlays "incurred or expected to be incurred in the reporting year, regardless of when the funds were appropriated." Obligations and outlays include those resulting from *research, development,* and *R&D plant,* which are defined as:

> a. Research is defined as systematic study directly toward fuller scientific knowledge or understanding of the subject studied. Research is classified as either basic or applied according to the objectives of the sponsoring agency.
>
> Basic research is defined as systematic study directed toward fuller knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications towards processes or products in mind.
>
> Applied research is defined as systematic study to gain knowledge or understanding necessary to determine the means by which a recognized and specific need may be met.
>
> b. Development is defined as systematic application of knowledge or understanding, directed toward the production of useful materials, devices, and systems or methods, including design, development, and improvement of prototypes and new processes to meet specific requirements.
>
> c. R&D plant includes acquisition of, construction of, major repairs to, or alterations in structures, works, equipment, facilities, or land for use in R&D activities at Federal or non-Federal installations.
>
> — National Science Foundation, NSF 05-307 (2005)

I selected obligations and outlays in *research* and *development* as the basis for public-sector R&D expenditure data. Those in *R&D plant* were excluded in order to render data on public R&D expenditure comparable to data on private R&D expenditure: firm R&D expenses reported on Form 10-k do not allow for the types of R&D expenses included in *R&D plant*.

In responding to the NSFSFFRD, reporting agencies are required to classify their research funds by academic discipline. Academic disciplines are classified into eight broad field categories, each consisting of a number of detailed fields. The broad field of *life sciences*, which is defined as "the scientific study of living organisms and their systems," is comprised of five detailed fields: *biological sciences (excluding environmental biology)*, *environmental biology*, *agricultural sciences*, *medical sciences*, and *life sciences not elsewhere classified* (National Science Foundation, 2005).

Of these five detailed fields, I chose *biological sciences (excluding environmental biology)*, *medical sciences*, and *agricultural sciences* to be the three life-science fields in the present study. The *environmental biology* field was excluded because of the difficulty in defining a parallel class in patent and private R&D expenditure data. Combining it with *biological sciences (excluding environmental biology)* to form a distinct field did not seem plausible either: the combination would confound the spillovers from biological research with those from environmental biology, the former with stronger effects than the latter. The *life sciences not elsewhere classified* field includes, by definition, multidisciplinary projects among the other life-science fields and projects that can not be assigned to any of the other four fields. Inconsistencies of its series may therefore arise across agencies and time, and it also was excluded from the life-science fields in the present study.

A number of problems would immediately arise if I directly used the NSFSFFRD R&D expenditure data. First, although obligations in *basic research* and *applied research* were reported by each agency at the detailed field level, obligations in *development* were not: only the aggregate *development* obligations of all fields were reported by each agency. Thus, I needed to find a way to estimate *development* funds in each of the three life-science fields.

Since *development* is a natural extension of *applied research*, it is reasonable to assume that the distribution of *applied research* funds across individual life-science fields approximates the distribution of *development* funds across those same fields. With this assumption, I estimated the *development* funds in each life-science field by (a) dividing each agency's total *development* funds into fields in proportion to the distribution of the agency's *applied research* funds in those same fields; and (b) aggregating across all federal agencies the resulting *development* funds in each field.

The second problem can be seen in Table 5.2, which provides a 2002 snapshot of the flows of R&D funds from funding to performing sector. Table 5.2 shows that a significant proportion of federally financed R&D is performed by industry through contracts or grants (under Industry and Industry FFRDCs in Table 5.2). Its contract and grant support does have an impact on the private sector's R&D investment. However, because the mechanism of its spillover impacts is different from what the Chapter 3 model describes, the latter is not appropriate for assessing publicly financed, industry performed R&D. I therefore exclude industry-performed R&D from public-sector R&D expenditure.

Table 5.2 National R&D expenditures (in millions of 1996 USD) from funding sectors (in the first row) to performing sectors (in the first column): 2002

	Fed Govt.	*State Govt.*	*Industry*	*U&C*	*Non-profits*
Federal Govt.	19,493				
Industry	18,920		171,653		
Industry FFRDCs	2,050				
U&C	20,365	2,235	2,117	6,741	2,428
U&C FFRDCs	5,476				
Non-profits	4,994		1,051		4,177
Nonprofit FFRDCs	1,917				

Source: National patterns of R&D sources (NSF).
FFRDC: federally funded research and development center
U&C: universities and colleges

The difficulty of this exclusion is that although the NSFSFFRD has collected reporting agencies' obligations in *basic research*, *applied research*, and *development* by performer, such obligations refer only to the broad field level, not to the detailed field level. To solve this problem, I assumed that for every reporting agency, the ratio between industry- and non-industry-performed R&D in each detailed field is the same as the ratio in the broad field to which the detailed field belongs. With that assumption, I then: (a) divided each federal agency's total R&D funds into industry-performed and non-industry-performed components at the detailed field level; and (b) summed the R&D funds in each detailed field across all federal agencies.

The third difficulty also can be seen in Table 5.2. The public sector in the present study consists of all institutions operated under a not-for-profit criterion. Table 5.2 shows the federal government is not the only funding institution that can be placed in the public domain. State governments, universities and colleges, and not-for-profit institutions are in the public sector as well. Unfortunately, R&D data for these public institutions have not been as systematically collected as federal R&D data have. Fortunately, NSF also conducts a Survey of Research and Development Expenditures at Universities and Colleges. In its survey, universities and colleges are asked to report R&D funds by funding source and academic discipline. Table 5.2 indicates that the majority (73% in 2002) of R&D funded by state governments, universities and colleges, and not-for-profit institutions is performed by universities and colleges. Non-federally-financed academic R&D funds therefore provide a good proxy for non-federally-financed public-sector R&D expenditures.

Thus, I subtracted federally funded academic R&D from total academic R&D, then added the result to federal R&D expenditure in each life-science field. Note that in this way industry-funded academic R&D also was included in public-sector R&D. In other words, by solving one problem we bring in another. Nevertheless, industry-funded academic R&D assumes only a small proportion (16% in 2002 as shown in Table 5.2) of non-federally funded academic R&D. The measurement bias generated by including non-federally funded academic R&D in public-sector R&D is smaller than that generated by not doing so.

In summary, I proceeded as follows:

Step 1. Estimation of federal *development* expenditures in individual life-science fields:

a. Calculated the distribution of each federal agency's *applied research* funds in various detailed fields in each of the years from 1980 to 2003;

b. Allocated total *development* funds in each agency-year to various detailed fields in proportion to the corresponding distribution in (a);

c. Aggregated the estimated *development* funds in each of the three life-science fields across all agencies.

Step 2. Estimation of non-industry-performed federally financed R&D expenditures:

a. Calculated the share of each agency-year's *basic research* performed by institutions other than for-profit firms;

b. Multiplied each agency-year's total federally funded *basic research* in each detailed field by the corresponding share computed in (a);

c. Aggregated the result in (b) across all agencies in each detailed field and each year;

d. Repeated steps (a), (b), and (c) for *applied research* and *development* funds;

e. Added the annual series of *basic research*, *applied research*, and *development* funds in each of the three detailed fields: *biological sciences (excluding environmental biology)*, *agricultural sciences*, and *medical sciences*. This yielded annual series of non-industry-performed federally funded R&D expenditures in the three life-science fields.

Step 3. Estimation of academic R&D expenditures from nonfederal funding sources:

In each of the three life-science fields mentioned in step 2, subtracted the annual series of academic R&D funds received from the federal government from the annual series of total academic R&D funds.

Step 4. Estimation of public R&D expenditures in the three life-science fields: Added, in each of the three life-science fields, the resulting annual series obtained in steps 2 and 3.

5.2 Private-Sector R&D Expenditures

According to Hall and Long (1999), two sources of micro-data are available for constructing industry R&D expenditures. One source is Form RD-1, on which U.S.companies report their R&D expenditures to the Census Bureau. These data have been aggregated to the 2-digit industry level and published by the National Science Foundation each year. The RD-1 data, however, are confidential and not available to the public (Hall and Long, 1999). Another source is Form 10-K of publicly listed firms in the United States. All companies traded on U.S.stock markets are required to file their R&D expenditures with the Securities Exchange Commission (SEC) on Form 10-K. These data are readily accessible via Standard and Poor's Compustat© database. According to Ernst & Young, R&D expenses of publicly traded firms, as reported on Form 10, accounted for 80% of industry's total R&D expenditures in 2002.

Private-sector R&D expenditure data in the present study were constructed on the basis of Form 10-K data. The construction procedure involved two steps: breaking down firms' R&D expenditures into various scientific fields, then aggregating the R&D expenditures in each life-science field across all firms. Breaking down firms' R&D expenditures is necessary because many companies' R&D projects cut across more than one scientific field.

The strategy adopted was to use information at the business segment level to break down life-science firms' R&D into various fields. The Compustat Business Information–Segment Products file contains firms' financial data—including R&D expenses—at the business segment level. The idea was to define an approximate mapping between business segments and scientific fields. The difficulty in defining such a mapping is that, on Form 10-K, business segments are classified based on industry classification systems such as the Standard Industry Classification (SIC) and the North American Industry Classification System (NAICS). Hence, any mapping method is inherently imperfect because SIC and NAICS are product-oriented classification systems, while production lines may be supported by R&D from different scientific fields. Moreover, subjective variations arise across reporting firms regarding which business segments belong to which SIC/NAICS industries. Worse still, when the SIC/NAICS data are missing, the matching method is no longer helpful in the mapping process. Solving these problems was the most laborious part of the data construction project. Private-sector R&D expenditures were constructed as follows:

Step 1: Selection of life-science-related NAISC and SIC industries:

We selected industries whose R&D activities, if any, belong in the broad field of life sciences. Using two industry classification systems, NAICS and SIC, I obtained a set of NAICS industries and another of SIC industries, denoted respectively by LS-NAICS (Table 5.3) and LS-SIC (Table 5.4).

Table 5.3 Selected NAICS industry classes matched to life-science fields

NAICS	*Industry Class Name*	*Life-Science Field*
111	Crop Production	Agricultural Sciences
112	Animal Production	Agricultural Sciences
113	Forestry and Logging	Agricultural Sciences
114	Fishing, Hunting and Trapping	Agricultural Sciences
115	Support Activities for Agriculture and Forestry	Agricultural Sciences
311	Food Manufacturing	Agricultural Sciences
312	Beverage and Tobacco Product Manufacturing	Agricultural Sciences
3253	Pesticide, Fertilizer, and Other Agricultural Chemical Manufacturing	Agricultural Sciences
54132	Landscape Architectural Services	Agricultural Sciences
56173	Landscaping Services	Agricultural Sciences
81291	Pet Care (except Veterinary) Services	Agricultural Sciences
3254	Pharmaceutical and Medicine Manufacturing	Medical Sciences
54194	Veterinary Services	Medical Sciences
621	Ambulatory Health Care Services	Medical Sciences
622	Hospitals	Medical Sciences
54171	Research and Development in the Physical, Engineering, and Life Sciences	Agricultural, medical, biological sciences or others
54169	Other Scientific and Technical Consulting Services	Agricultural, medical, biological sciences or others
541380	Testing Laboratories	Agricultural, medical, biological sciences or others

Table 5.4 Selected SIC industry classes matched to life-science fields

SIC	*Industry Class Name*	*Life-Science Field*
100	Agricultural Production Crops	Agricultural Sciences
200	Agriculture production livestock and animal specialties	Agricultural Sciences
700	Agricultural Services	Agricultural Sciences
800	Forestry	Agricultural Sciences
900	Fishing, hunting, and trapping	Agricultural Sciences
2000	Food And Kindred Products	Agricultural Sciences
2100	Tobacco Products	Agricultural Sciences
2870	Agricultural Chemicals	Agricultural Sciences
2830	drugs	Medical Sciences
8731	Commercial Physical and Biological Research	Agricultural, medical, biological sciences or others
8734	Testing Laboratories	Agricultural, medical, biological sciences or others

Step 2: Selection of life-science firms:

We used the LS-NAICS and LS-SIC industry classes to identify life-science firms. A firm was selected if at least one of its business segments was classified in an industry belonging in either LS-NAICS or LS-SIC and in at least one year from 1980 to 2004. The resulting set, denoted by LS-FIRM, contains 3,006 firms (The Compustat universe consists of more than 10,000 active and 7,600 inactive firms.)

Step 3: Data retrieval:

Annual data from 1980 to 2004 on *R&D expenses*, *net sales*, *business segment names*, *NAICS*, and *SIC*, were drawn from the Compustat Business

Information-Segment Products file for LS-FIRM firms. Corresponding firm-level data in the same time period were retrieved from the Compustat Industrial Annual Data file. These two datasets then were merged by firm and by year.

Step 4: Refinement of the life-science firm set:

A firm was deleted from set LS-FIRM if its R&D expenses were never reported, or reported to be zero in every year from 1980 to 2004. After the deletions I was left with 1,884 firms in LS_FIRM.

Step 5: Mapping of business segments into R&D fields:

Each business segment was assigned to one of the five R&D fields—agricultural sciences (AG), medical sciences (MED), biological sciences (BIO), non-life-science fields (NLS), and no R&D (NORD)—using a procedure combining computer programming and manual assignment. The programmable part of the mapping was executed by running the following routine:

a. A business segment is assigned to R&D field NORD if its *primary* SIC code indicates that it belongs to industries such as 4000 (Transportation), 5000 (Wholesale and Retailer), and 6000 (Insurance and Finance).

b. A segment is assigned to R&D field NLS if it cannot be assigned to field NORD according to instruction (a), and none of its primary and secondary business segment SIC code belongs to industry set LS-SIC.

c. A segment is assigned to R&D field AG if its primary SIC code belongs to one of the following SIC industry classes: 100 (Agricultural Production Crops), 200, 700, 800, 900, 2000, 2100, or 2870 (see Table 5.4 for definitions).

d. A segment is assigned to R&D field MED if its primary SIC code belongs to SIC industry class 2830 (drugs).

e. A segment whose primary SIC code belongs to SIC industry class 8731(Commercial Physical and Biological Research) or 6794 (Patent Owners & Leasers) is assigned to R&D field AG if its secondary SIC code belongs to one of the agricultural industries listed in instruction (c), to field MED if its secondary SIC code belong to SIC industry class 2870 (drugs), to field NLS if its secondary SIC code does not appear in industry set NS-SIC, and to field BIO otherwise.

After executing the above routine, I still was left with some business segments to which no R&D field was assigned. These segments largely fell into two categories. One included those whose SIC codes are not available in certain years. Another included those whose primary SIC codes do not—but secondary SIC codes do—belong to the set of life-science-related industries LS-SIC. I solved this missing-data problem using the following method:

f. For the data-missing segment-year observations, missing SIC codes were filled up with the SIC codes of the same or similar business segments in data-non-missing years. Whether two business segments in different years are in fact the same business segment can, in most cases, be identified by the names and IDs of the business segments in question.

g. Each of these business segments then was assigned to an R&D field following instructions a through e.

After solving the missing-data problem, I was left with 155 firms whose primary SIC codes did not—but secondary SIC codes did—belong to the set of life-science-related industries LS-SIC. Dr. William Folks, Professor in Biochemistry at the University of Missouri-Columbia, was asked to help me with this problem. I read to Dr. Folks the names of these business segments and Compustat's textual description of the 155 firms' businesses, and he used his judgment in suggestion to which R&D field each indeterminate segment should be assigned[5].

Step 6: Classification of life-science firms by R&D field:

With the R&D field of each business segment identified, the 1,884 life-science firms were classified into the following five groups: agricultural R&D (UNIAG), medical R&D (UNIMED), biological research (UNIBIO), non-life-science R&D (UNINLS), and firms conducting R&D in more than one field (MULTIF). Of the 1,884 firms in LS-FIRM, 679 fell into group UNINLS and thus were deleted from set LS-FIRM. Of the remaining 1205 firms, 180

[5] It is worth noting that for a business segment whose SIC code belongs to industry classes like 286 (Industrial Organic Chemicals) and 2844 (Perfumes, Cosmetics, and Other Toilet Preparations Cosmetics), we assigned it into the R&D field MED as long as its company's main business is drug manufacturing.

were assigned to UNIAG, 816 to UNIMED, 12 to UNIBIO, and 197 to MULTIF.

Step 7: Estimation of R&D expenditures of MULTIF firms' individual fields: Recall that our objective is to obtain industrial R&D expenditure series in the three life-science fields. Hence, I needed to aggregate all firms' R&D expenditures in each individual field. For firms only conducting R&D in a single R&D field, such as those in UNIAG, UNIMED, and UNIBIO, I could directly use the firm-level R&D expenditure data. For firms in MULTIF, however, I needed to obtain R&D expenditure data at the business segment level. The problem was that although all firms remaining in LS-FIRM reported R&D expenses at the firm level, not all did report R&D expenses at the business-segment level. Business-segment R&D expenditures were estimated using the following method:

a. Twenty one of the 197 MULTIF firms were found to have reported segment R&D expenditures for all but a few years. Moreover, business segments in the data-missing years were found to exist in recent data-available years. I assumed the distribution of the firms' R&D expenditures across business segments is the same in the data-missing as in the data-available years. With this assumption, I estimated the missing R&D expenditure data by i) computing the distributions of R&D expenditures across business segments in the recent three years of the data-missing year; ii) computing the geometric mean of the distributions in these three years; and iii) dividing the firm's total R&D expenses in the missing year into various business segments according to the average distribution computed in ii).

b. For the remaining 168 firms, I adopted the following formula to estimate the missing R&D expenditures:

1) If in the data-missing year no business segment belongs to R&D field NORD, and R&D expenditure data are missing for all business segments, then a business segment's R&D expenditure is equal to the firm's total R&D expenditure multiplied by the said business segment's sales share in the firm;

2) If in the data-missing year one or more segments belong to R&D field NORD, the R&D expenditure of a segment that does not belong to NORD is equal to the firm's total R&D expenditure multiplied by the said business segment's sales share in all the firm's business segments except those belonging to NORD;

3) If, in the data-missing year, R&D expenditure data are available for some but not all segments, then total R&D expenditure and sales of all data-missing segments are computed by subtracting data-non-missing segments' total expenditure and sales from the firm's total expenditure and sales, respectively; and the data-missing business segments are treated as a firm with R&D data missing for all

business segments to apply formulae 1) and 2), depending on whether there exists a data-missing segment belonging to NORD.

Step 8: Aggregation of R&D expenditure data in individual life-science fields: The resulting data were integrated into a dataset of R&D expenditures at the business segment level. The dataset has three dimensions—time, firm, and R&D field. I sorted the entire dataset into three life-science R&D fields in each year, then aggregated the R&D expenditures in each individual field each year.

Construction of private-sector R&D expenditure data for the individual life-science fields was the most laborious part of the data processing, and is the best that could be done given the data sources available. Nevertheless, the procedure might have generated some measurement errors. Following is an assessment of the potential errors:

a. As described above, a business segment's R&D field was solely determined in most cases by its *primary* SIC industry code. Sometimes, however, the *secondary* SIC code indicated that R&D activities in the business segment might cut across two R&D fields. Unfortunately, no information was available on the relative proportion of R&D in two different fields within a given business segment. I therefore was unable to separate a business segment's R&D expenditure further into two R&D fields.

b. An assumption of the above data construction was that the “private sector” is composed of all companies traded in U.S. stock markets regardless of the home country of the firm. However, the universe of Compustat firms has evolved over time on account of firm entry, exit, and merger. Although the merger of two Compustat firms would not affect my aggregate data, entry into and exit from U.S. stock markets might bring some measurement error. The extent of this potential mismeasurement is unknown.

c. My data indicate an almost negligible volume of industrial R&D expenditure in biology. This may be interpreted to mean that industry performs little basic biological research. Or the little it does is conducted with specific application in mind. It may therefore be included in either medical or agricultural research, depending upon the specific application in mind. In general, however, I expect the majority of this type of research is performed in pharmaceutical firms. This was confirmed by reading the business segment names in the UNIBIO group. I thus merged the UNIBIO firms with the UNIMED firms. Accordingly, biological and medical patents awarded to industrial firms were merged into a single group of medical patents (see Section 5.4.)

5.3 R&D Investment and Wage of Scientific Labor

R&D expenditure data are nominal measures in that they do not take account of changes in the price of the research input, which is predominantly composed of scientific labor. Appropriate deflators or price indices were needed to convert nominal into real R&D expenditures. I used the Biomedical Research and Development Price Index (BRDPI) constructed by the National Institutes of Health (NIH) to deflate medical and biological R&D expenditures, and the agricultural R&D deflator constructed by the USDA's Economic Research Service (ERS) to deflate agricultural R&D expenditures. The agricultural R&D deflator was developed on the basis of Klotz et al. (1995)'s method. The BRDPI is defined as follows:

> "The BRDPI measures changes in the weighted-average of the prices of all the inputs (e.g., personnel services, various supplies, and equipment) purchased with the NIH budget to support research. The annual change in the BRDPI indicates how much the NIH budget would need to change to maintain purchasing power— to compensate for the average increase in prices and to maintain NIH-funded research activity at the previous year's level."
>
> — National Institute of Health (2006)

The R&D investment series shown in Table 5.1 are these deflated R&D expenditures, and scientific labor wage rates are the price indices used in the deflation.

5.4 Knowledge Output: Patent Counts

Patent counts long have been used to measure knowledge output. In order to take into account the value differences among patents, citation data may be used to adjust patent counts. I constructed knowledge output data using patent count data purchased from Thomson™. Unfortunately, the cost of citation-count-adjusted patent counts was prohibitive.

In collaboration with Thomson's patent experts and my colleagues, I selected a set of life-science patent classes from four DWPI (Derwent World Patent Index) patent Sections: B (Pharmaceuticals), C (Agricultural Chemicals), D (Food, Detergents, Water Treatment, and Biotechnology), and P (General). The selected patent classes and subclasses are listed in appendix C.1. I then worked with Thomson's patent experts to design a search procedure to retrieve annual patent counts for the three life-science fields—biology, medicine, and agriculture—and five agricultural subfields: plants, animals, natural resources, food, and agricultural chemicals. I designed a filter to distinguish between patents awarded to institutions in the public and in the private sector. The resulting measures included the public sector's and private sector's annual patent counts from 1980 to 2005 in three life-science fields and five agricultural subfields. (See Table 5.1 for the private sector's data.)

Patent data were retrieved according to the following re-classification of DWPI patent classes:

BIOP, the set of biological patents, including all patents in class D16, i.e,

$$\text{BIOP} = \text{D16};$$

PHARMP, the set of medical patents, including patents in classes B01-06, exclusive of all patents in classes D and P, i.e.,

$$\text{PHARMP} = (\text{B01} \cup \text{B02} \cup \text{B03} \cup \text{B04} \cup \text{B05} \cup \text{B06}) \setminus (\text{BIOP} \cup \text{D13} \cup \text{D15} \cup \text{P11} \cup \text{P12} \cup \text{P13} \cup \text{P14} \cup \text{P15});$$

AGP, the set of agricultural patents, including patents in C01-06, D13 and D15, and P11-15, while excluding those that belong to either BIOP or PHARMP, i.e.,

$$AGP = (C01 \cup C02 \cup C03 \cup C04 \cup C05 \cup C06 \cup P11 \cup P12 \cup P13 \cup P14 \cup P15 \cup D13 \cup D15) \setminus (PHARMP \cup BIOP).$$

AGP was further partitioned into five subsets: PLANTP, ANIMALP, NATRESP, FOODP, and AGCHEMP, respectively representing patents in the agricultural subfields of Plants, Animals, Natural Resources, Food, and Agricultural Chemicals. The agricultural subfield classification consists of:

PLANTP, including all AGP patents in DWPI class P11-13 and P15, i.e.,

$$PLANTP = (P11 \cup P12 \cup P13 \cup P15) \cap AGP;$$

ANIMALP, including all AGP patents in class P14, i.e., $ANIMALP = P14 \cap AGP$;

NATRESP, including all AGP patents in class P15, i.e., $NATRESP = D15 \cap AGP$;

FOODP, including all AGP patents in class D13, i.e., $FOODP = D13 \cap AGP$;

AGCHEMP, including all AGP patents in class C01-C06, i.e.,

$$AGCHEMP = (C01 \cup C02 \cup C03 \cup C04 \cup C05 \cup C06) \cap AGP.$$

Finally, annual series of patent counts in each field and subfield were divided into two time series, one for the public sector and another for the private sector. For the reason explained in Section 5.2, the private sector's medical and biological patent counts were combined to measure industry's knowledge output in medical research and development. Along with Thomson's patent experts and my colleagues, I designed a list of keywords for finding the names of the public-sector patent assignees. Drawing public-sector patents from our patent pool left the private-sector patents. Following are the keywords employed:

> "CENTER" or "cent" or "AGENCY" or "INST" or "INSTITUTION" or "INSTITUTE" or "inst" or "ASSOCIATION" or "assoc" or "FOUNDATION" or "fond" or "found" or "ORG" or "PUBLIC" or "GOVERNMENT" or "UNIVERSIT?" or "univ" or "COLLEGE" or "dept" or "US dept" or "us admin" or "us sec" or "NAVY" or "NASA" or "ARMY" or "AIRFORCE" or "sec" or "us".

5.5 Price of Knowledge Output: Average Market Value of Patents

A patent is the property right awarded to the innovator of an invention, giving exclusive commercial use of the invention. If a market exists for a patent, its market price should exactly equal the present value of cash flow that can be generated by completely exploiting the property right associated with that patent. Hence, patent price reflects profitability in the industry in which the invention is used to produce new goods or reduce production costs.

One might question whether patent prices reflect social value, given that the distribution of patent values is highly skewed: a small number of blockbuster patents is extremely valuable, while the majority have no licensing revenues whatever. However, a time series of average patent value remains informative about changes in profitability conditions in the patent-using industries (Griliches, 1990). These changes have an impact on industry R&D investment and hence should be taken into account to explain fluctuations in industrial R&D investment demand.

The difficulty in constructing annual series of average patent values is that no integrated market exists for patents. Thus, one has to resort to indirect information—say, the implicit value of a patent as perceived by economic agents—to estimate patent value. Pakes (1986) and Lanjouw, Pakes, and Putnam (1998) use patent application and renewal data to estimate such value. Their idea is that every patent holder must posit an "optimal stopping" decision each year about whether or not to renew the patent monopoly, keeping in mind the annual renewal fee to which every patent is subject. Hence, the renewal data contain information about the value of the patent as perceived by the patent holder.

Alternatively, researchers have estimated patent market value through stock market information [Griliches (1981); Hall, et. al. (2005)]. A firm's stock price reflects its profitability as perceived by the buyers and sellers of the firm's stocks. The value of a new invention will, if significant, change the price of the inventing firm's stock, other things equal. With this in mind, Griliches (1981) regresses the firm's Tobin Q on the firm's patent counts along other control variables such as R&D expenditure. The estimated marginal effect of patent counts on Tobin's Q provides an estimate of patent value. Griliches shows that a successful patent is worth about $ 200,000. (The base year of this figure is unclear; but from the 1967-1974 sample period and 1981 article publication date, I assume it is in the late 1970's.) A recent study by Hall, et al. (2005)

applies citation-weighted patent counts to the Griliches model and finds that the marginal value of a patent is approximately $ 370,000 million (1980 U.S. dollars.)

This method is data-demanding. Employing it to obtain a time series of average patent market values in a given field requires time-series, cross-section data at the firm level. To my knowledge, the only data with this potential is the NBER patent data constructed by Bronwyn Hall and her coworkers. Constructing such panel data requires merging firm data in Compustat with NBER patent data. The work of matching Compustat firms with patent assignees is still in its infancy. Professor Hall and her coworkers have developed such a matching on the basis of firm ownership information in 1989. They have matched 2,592 Compustat firms with patent assignees who were granted USPTO patents in or before 1999. Unfortunately, of the over 1000 life-science firms in my dataset, only 134 appeared in Hall's firms-matched-to-assignees sample. Still worse, few of the 134 firms appeared in my dataset through the entire 1980-2004 sample period.

Confronted with these difficulties, I turned to a relatively crude estimation method. I subtracted each firm's book value from its market value, aggregated the net market values of all firms in a field and year, then divided the resulting aggregate net market value by the total number of patents awarded in the field and year.

Unlike the previous two approaches, this requires data only at a high level of aggregation. Of course, it assumes that the profit potential of a firm's research entirely explains the firm's market value net of its physical assets. This is a strong assumption. Other intangible assets—such as consumer loyalty established through advertising, and management efficiency—are reflected in stock price as well. Nevertheless, the approach does capture the most important changes in the profitability conditions of patent-using industries.

In summary, following are the procedures used to construct time series of average patent values in each life-science field:

Step 1. Compute book and market values of the 1205 life-science firms in LS-FIRM (see appendix C.2 for formulae and definitions).

Step 2. Estimate book and market values of business segments in the 197 firms in MULTF, which were found to conduct R&D in more than one R&D field. Each firm's book and market values were divided into business segments according to sales shares.

Step 3. Aggregate book and market values across firms and business segments in each field and year.

Step 4. Calculate market-to-book ratios in each field and year.

Step 5. Multiply aggregate book values by one minus the corresponding market-to-book ratio to obtain the total net market value of all firms and segments in a given field.

Step 6. Normalize annual net market values obtained in step 5 into 2001 dollars using the Consumer Price Index (CPI).

Step 7. Construct patent stocks by applying the perpetual inventory formula: $K_t = R_{t-1} + K_{t-1}(1-\delta)$, $t = 1981, \cdots 2004$; where K is patent stock, R is annual patent counts, and δ is obsolescence rate (δ = 0.15). Initial knowledge stock was estimated by $K_{1980} = R_{1980}/\gamma$, where $\gamma = 0.23$.

Step 8. Divide the aggregate net market value in a research field and year by patent stock constructed in Step 7. This obtains the average patent value in the given field and year.

Chapter 6: Results

As discussed in Chapter 4, the R&D investment models developed in Chapter 3 were estimated using the GMM estimation technique. Estimation was executed in GAUSS with the Hansen/Heaton/Ogaki GMM package. Each model was estimated four times, successively using instrument sets $I_{j,a}$, $I_{j,b}$, $I_{j,c}$, and $I_{j,d}$, where $j = med$ for the medical research industry and $j = ag$ for the agricultural research industry. (See Appendix B for definitions of these instrument sets). The GMM estimator is a 3SLS estimator that corrects for estimated heteroskedasticity. Estimation results were fairly robust across the four instrument sets and across alternative discount factors (see Appendix B for details.) Here, I discuss the results obtained from instrument set I_b [6] under the assumption of a 0.95 discount factor.

6.1 Parameter Estimation and Specification Tests

6.1.1 Medical Research Industry

Tables 6.1 displays parameter estimates from the medical research industry. The *J* statistic in this table represents the Hansen *J* statistic, which has an asymptotic chi-

[6] I use a “minimum mean squared prediction error” criterion to compare results obtained from the four sets of instrument variables. Specifically, for each observation, I substitute the observed exogenous variables into the estimated model to compute the equilibrium values of the endogenous variables. I then compute the mean squared prediction error for each endogenous variable. In both the medical and agricultural research industry, the model estimated with I_b generates the smallest prediction error for all endogenous variables.

Table 6.1 Parameter estimates in the medical research industry

	Model 1	***Model 2***	***Model 3***	***Model 4***
α_1	0.3852 (0.0101)	0.3411 (0.0091)	0.3957 (0.0098)	0.3837 (0.0088)
α_2	0.0101 (0.0002)	0.0107 (0.0002)	0.0097 (0.0002)	0.0101 (0.0002)
β_1	5.7310 (0.8044)	-0.8324 (0.4379)	5.4541 (0.8013)	5.5571 (0.5765)
β_2	-0.3785 (0.1209)	-0.1469 (0.1185)	-0.7529 (0.0734)	-0.3532 (0.0892)
β_3	-0.0056 (0.0030)	-0.0112 (0.0030)	-0.0168 (0.0010)	-0.0052 (0.0028)
β_4	-0.0445 (0.0114)	0.0016 (0.0104)		-0.0457 (0.0107)
β_5	0.0820 (0.0136)	0.0797 (0.0135)	0.1178 (0.0099)	0.0815 (0.0135)
β_6	0.0251 (0.0045)	0.0310 (0.0045)	0.0408 (0.0021)	0.0244 (0.0038)
β_7	0.0024 (0.0077)	-0.0358 (0.0066)	0.0126 (0.0073)	
β_8	-0.9140 (0.0940)		-0.7617 (0.0855)	-0.8991 (0.0808)
J	15.1417	109.7443	30.3407	15.2377
P-Value	0.0871	0.0000	0.0008	0.1236
d.f.	9	10	10	10

Note: Models are estimated with instrument set I_b. Numbers in parentheses are asymptotic standard errors.

square distribution. The degrees of freedom of the chi-square distribution equal the number of moment conditions minus the number of parameters estimated. The J statistic therefore provides a natural test of the validity of the moment conditions. The statistic also can be used for specification tests. The J statistic of the restricted model minus that of the unrestricted model is distributed chi-square with degrees of freedom equal to the number of restrictions imposed. Model 2 in Table 6.1 is that described in Chapter 4.

Model 1 adds a trend term to the production function in Model 2. The parameter of the trend term is β_8. The *J* statistic in Model 2 (109.74) minus that in Model 1 (15.14) equals 84.60, which is far greater than 3.84, the 5% critical value of a chi-square distribution with 1 degree of freedom. Thus, Model 2 is rejected against Model 1. Model 3 imposes restriction $\beta_4 = 0$ on Model 1, implying no adjustment costs in medical R&D investment. The same test procedure leads me to reject the no-adjustment-cost hypothesis at the 5% level. Model 4 imposes restriction $\beta_7 = 0$, implying no knowledge spills over from agricultural research, be it public or private, to private-sector medical research. The chi-square statistic equals to 15.24 - 15.14 = 0.1, far less than the 5% critical value of a chi-square distribution with 1 degree of freedom. The null hypothesis thus is not rejected. However, the comparable hypotheses regarding knowledge spillovers from public-sector medical and biological research to private-sector medical research, i.e., $\beta_5 = 0$ and $\beta_6 = 0$ respectively, are rejected. The joint hypotheses that there have been no knowledge spillovers from pairs of spilling sources were also tested and successively rejected. Thus, Model 4 is the final one adopted.

In Model 4, assumptions $\alpha_2 > 0$, $\beta_3 < 0$, and $\beta_4 < 0$ in the theoretical model (Chapter 3) are satisfied. In medical research, that is, scientific labor supply is a linearly increasing function of wage and the representative firm's technology is quadratic with internal adjustment costs. These assumptions guarantee that the representative firm's dynamic decision problem (3.3) is well defined. The *J* statistic in Model 4 equals 15.24, which is less than 18.31, the 5% critical value of the chi-square distribution with 10 degrees of freedom. The overidentifying restrictions are therefore not rejected, supporting the use of the moment conditions.

The failure to reject the adjustment-cost hypothesis suggests a dynamic model is more appropriate for the medical research industry than is a static model. Recall that adjustment parameter $(1-\lambda_1)$ represents the proportion of the adjustment that occurs within a single year [see equation (3.18)]. As shown in expressions (3.8) and (3.9) λ_1 is a nonlinear function of interest rate r, knowledge output price $p_{med,t}$, and parameters β_3, β_4, and α_2. The average adjustment parameter over the 1980-2003 sample period is 0.29 (quasi-rational expectations are assumed in computing this parameter), implying that on average only 29% of the adjustment occurred within a single year. This result is consistent with Bernstein and Nadiri's (1989) finding that 36%, 32%, 26%, and 22% of R&D investment adjustment cost occurred within a year in, respectively, the chemical, petroleum, machinery, and instrument industries. The chemical industry (SIC28) in that study includes as a subclass the same medical industry (SIC283), from which I have drawn my medical research firm data. The result obtained here thus complements that obtained by Bernstein and Nadiri by suggesting medical R&D tends to incur higher adjustment costs than do other chemical industries in SIC28.

Model 4 implies knowledge has spilled over from public-sector medical and biological research to private-sector medical research. That is, public biological and medical research has enhanced medical firm research productivity by creating technological opportunities and providing knowledge infrastructure. Interestingly, the estimated parameter on public-sector medical research investment (0.082) is larger than that on public-sector biological research investment (0.024). Yet it remains unclear whether the former contributed more than the latter to the growth of industrial demand for medical R&D investment, since the former's negative wage effect has not been accounted for.

Although the sign on the parameter of agricultural R&D investment is positive in Model 1, the rejection of Model 1 against Model 4 indicates that knowledge spillovers from agricultural research, be it public or private, to private-sector medical research are not statistically significant. It is well known that, thanks to advances in biotechnology, many subfields of medical and agricultural research now are organized within the same discovery process. This is confirmed by my observation, when describing construction of private-sector R&D expenditures in the medical and agricultural sciences, that many large life-science firms have both agricultural and pharmaceutical business segments. A recent study by Xia (2002) finds evidence of within-firm knowledge spillovers between agricultural biotech research and pharmaceutical biotech research. The result obtained here, however, is not inconsistent with Xia's result. Agricultural and medical research are in the present study broadly defined to consist of both non-biotech-based and biotech-based (DNA-based) research programs. Knowledge spillovers from biotech-based agricultural research to biotech-based medical research likely are diluted in aggregation.

Finally, the estimated negativity of β_8 implies that, after controlling for knowledge spillover effects, productivity of private-sector medical R&D has been declining since 1980. This productivity regress may be due to technological exhaustion, to a diminishing rate of USPTO patent approvals, or to both.

6.1.2 Agricultural Research Industry

Table 6.2 provides the parameter estimates from the agricultural research industry. Models 1 and 2 in this table are analogous to those in Table 6.1 for medical research. Unlike in medical research**,** the hypothesis that there has been no trend in technological or institutional shocks is not rejected in the agricultural research industry. Model 2, that is, is not rejected against Model 1 because the J statistic in Model 2 (12.02) minus that in Model 1 (11.64) is 0.38, far smaller than 3.84, the 5% critical value of the chi-square distribution with 1 degree of freedom. Model 3 is used to test the no-adjustment-cost hypothesis in agricultural research by imposing restriction $\beta_4 = 0$ on Model 2. The no-adjustment-cost hypothesis is not rejected, even though β_4 assumes the correct sign in Model 2. Model 4 allows for adjustment costs while assuming no knowledge spillovers from medical research, public or private, to private-sector agricultural research, i.e., $\beta_7 = 0$. I fail to reject this hypothesis. Finally, using Model 5, I test hypotheses $\beta_4 = 0$ and $\beta_7 = 0$ jointly. The joint hypothesis is not rejected either. In the same fashion, assumptions $\beta_5 = 0$ and $\beta_6 = 0$ respectively are tested jointly with assumption $\beta_4 = 0$. The two joint tests are each rejected. The P-value of the J statistic in Model 5 is 0.37, strongly supporting the use of the moment conditions. Thus, I choose Model 5 as the final one for the agricultural research industry. This final model is a special case of David and Hall’s static model with linear labor supply and quadratic technology.

Table 6.2 Parameter estimates in the agricultural research industry

	Model 1	***Model 2***	***Model 3***	***Model 4***	***Model 5***
α_1	-0.1260	-0.1276	-0.1370	-0.1112	-0.1204
	(0.1290)	(0.1290)	(0.1285)	(0.1258)	(0.1252)
α_2	0.1047	0.1043	0.1062	0.1022	0.1041
	(0.0161)	(0.0161)	(0.0159)	(0.0157)	(0.0155)
β_1	0.3840	1.0433	0.9900	1.1108	1.0580
	(1.2837)	(0.7189)	(0.7157)	(0.7091)	(0.7059)
β_2	-0.0819	-0.4981	-0.4504	-0.5324	-0.4848
	(0.7295)	(0.2852)	(0.2788)	(0.2788)	(0.2723)
β_3	-0.1799	-0.1167	-0.1225	-0.1062	-0.1119
	(0.1127)	(0.0479)	(0.0473)	(0.0443)	(0.0437)
β_4	0.0175	-0.0681		-0.0685	
	(0.1626)	(0.0860)		(0.0860)	
β_5	0.3439	0.3889	0.3793	0.3879	0.3782
	(0.1232)	(0.0996)	(0.0988)	(0.0995)	(0.0988)
β_6	0.0446	0.0458	0.0464	0.0471	0.0477
	(0.0043)	(0.0038)	(0.0037)	(0.0030)	(0.0029)
β_7	0.0025	0.0005	0.0005		
	(0.0034)	(0.0008)	(0.0008)		
β_8	-0.0334				
	(0.0539)				
J	11.6386	12.0230	12.6497	12.3481	12.9833
P-value	0.2345	0.2835	0.3168	0.3381	0.3703
d.f.	9	10	11	11	12

Note: Models are estimated with instrument set I_b. Numbers in parentheses are asymptotic standard errors.

In Model 5 the positive estimate of α_2 implies that agricultural scientific labor linearly increases with wage. The estimate of α_2 in the agricultural model (0.10) far exceeds that in the medical model (0.01), implying agricultural scientific labor supply is much less elastic than is medical scientific labor supply. This in turn indicates that, in agricultural research, public investment has had a much stronger wage effect on private investment than it has in medical research.

On the other hand, the own-field knowledge spillover effect in agricultural research ($\hat{\beta}_5 = 0.38$) also is much stronger than in medical research ($\hat{\beta}_5 = 0.08$). As documented in Huffman and Everson (1993), research funded by public institutions such as the USDA and Agricultural Experiment Stations has been extraordinarily instrumental in driving technological change in the U.S. agricultural sector. The comparison here with medical research provides a new perspective on the contribution of public to private agricultural research. Yet the strong negative wage effect indicated by the inelastic agricultural scientific labor supply may well neutralize the knowledge spillover effect on private investment incentives. It remains unclear whether, in agriculture, public investment has been more complementary to private investment than it has been in medicine.

Public biological research also had a significant spillover effect on private agricultural research ($\hat{\beta}_6 = 0.48$), stronger than that on private medical research ($\hat{\beta}_6 = 0.24$). As in the medical model, however, public biological research has exerted a much weaker knowledge spillover effect than public own-field research has. In addition, no significant knowledge spillovers are found from either public or private medical research to private agricultural research. This may be interpreted in a way analogous to the medical model. That is, knowledge spillovers from biotech-based (DNA-based) medical research on biotech- based agricultural research is diluted, both in agriculture and medicine, in the aggregation of biotech-based and non-biotech-based research programs.

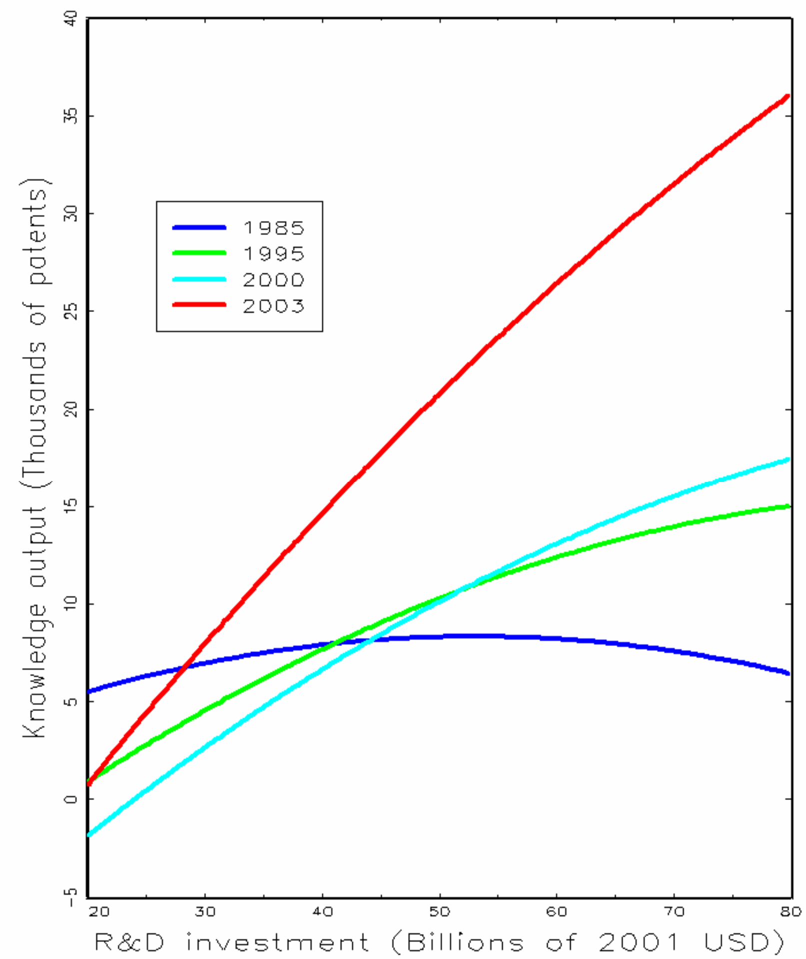

Fig. 6.1 Medical knowledge production function

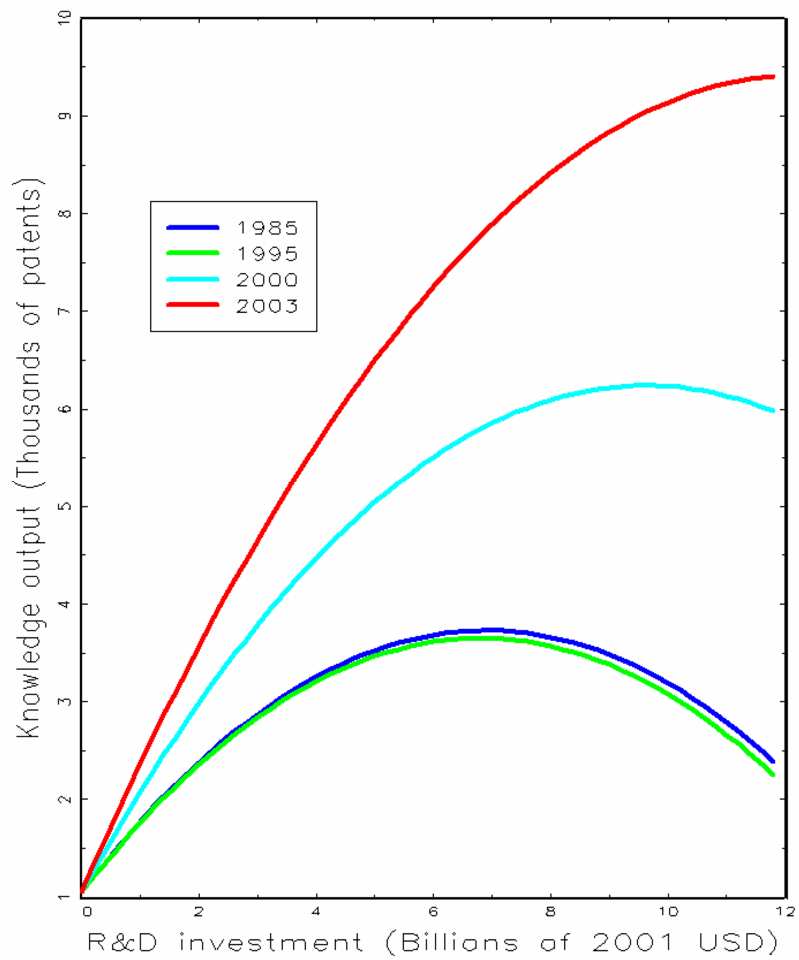

Fig. 6.2 Agricultural knowledge production function

Finally, a comparison of the estimate of the curvature parameter in the medical model ($\hat{\beta}_3 = -0.005$) and that in the medical model ($\hat{\beta}_3 = -0.112$) suggests agricultural research has a more concave production function than does medical research. In other words, the marginal productivity of private investment decreases at a faster speed in agricultural research than in medical research. This is evident in Figures 1 and 2, which show the evolutions of the two industries' knowledge production functions during the sample period. In common with both industries, however, the marginal productivity of private investment has increased dramatically over the sample period, thanks to public-sector investments in basic biological and own-field research.

In summary, results in this Section indicate that the medical research industry follows a model quite different from that in agricultural research. In the first place, the tests of the adjustment cost hypothesis suggest that a dynamic model is appropriate for the medical research industry, while a static one is appropriate for the agricultural research industry. In the second place, the technology and supply of scientific labor are quite different in the two industries. In common with both industries is the knowledge spillover pattern: public own-field and biological research improve private research (although the former did more strongly than the latter); and no significant knowledge spillovers are found between medical and agricultural research. The latter implies the two industry's R&D investments are exogenous to each other, dramatically simplifying the ensuing computation of elasticities and simulations (see footnote 1 on page 36.)   In the next Section, we study the elasticities of the representative firm's demand for scientific labor (R&D investment), providing insight into the complementarity vs. substitutability hypothesis and the demand-pull vs. technology-push debate.

6.2 Elasticities

Model 4 in Table 6.1 and Model 5 in Table 6.2 are used to compute input demand and output supply elasticities by applying the formulae derived in Chapter 3 and Appendix A. Since the specification tests suggest a dynamic model is appropriate in medical research, I compute elasticities under both quasi-rational and static expectations. Under the SEH, I compute both short- and long-run elasticities. Under the QREH, the long-run equilibrium is stochastic because of the stochastic nature of the exogenous variables. Thus, only short-run elasticities are computed under the QREH. Elasticity formulae include (3.13 -14) and those in Sections 2 and 3 of Appendix A. For the agricultural research industry, the model adopted is a static one in the spirit of David and Hall's because the adjustment cost hypothesis is rejected. In this case no distinction needs to be made between short- and long-run equilibrium. As shown in Chapter 3, equilibrium in such a static model is the same as the long-run dynamic equilibrium solved under static expectations. I therefore apply the long-run dynamic elasticity formulae, under the SEH, to the agricultural R&D model derived (see Appendix A.3.)

Table 6.3 contains elasticities of the medical research industry's demand for R&D investment with respect to public-sector investment in medical research (h_{med}^{pub}), public-sector investment in biological research (h_{bio}^{pub}), and knowledge output price (p_{med}).

Table 6.3 Elasticities of demand for R&D investment in the medical research industry

	Elasticity of private-sector demand for R&D investment					
	1982-1990			1991-2003		
	Short-run		Long-run	Short-run		Long-run
With respect to	SEH†	QREH‡	SEH	SEH	QREH	SEH
h_{med}^{pub}	1.55	0.97	3.05	0.67	0.51	1.27
h_{bio}^{pub}	0.52	0.38	1.03	0.22	0.20	0.42
p_{med}	0.31	0.23	0.88	0.04	0.04	0.13

Note: Numbers are average elasticities over each indicated time period.
†Elasticities are computed under the static expectations hypothesis.
‡Elasticities are computed under the quasi-rational expectations hypothesis, under which the representative firm holds static expectations on output price p_{med}, and rational expectations on the other two state variables, h_{med}^{pub} and h_{bio}^{pub}. The latter two variables are assumed to evolve in a first-order autoregressive mode. OLS estimation yields $E_{t-1}(h_{med,t}^{pub}) = 1.24 + 0.86 h_{med,t-1}^{pub}$ and $E_{t-1}(h_{bio,t}^{pub}) = 0.47 + 0.96 h_{bio,t-1}^{pub}$.

Elasticities are evaluated at all observations and averaged over two successive periods: 1982-1990 and 1991-2003. The private sector's investment demand elasticity with respect to public investment permits examination of the complementarity vs. substitutability hypothesis discussed in Chapters 2 and 3. The positive signs on the demand elasticities with respect to h_{med}^{pub} indicate that public-sector R&D investment is complementary to private-sector R&D investment in medical research, even after accounting for the negative wage effect. Evidently, the positive spillover effect dominates the negative wage effect, consistent with the finding in the previous Section that medical scientific labor supply is highly elastic. That the demand elasticities with respect to h_{bio}^{pub} are positive implies public-sector investment in biological research helps improve the medical research industry's productivity. Recall that in the previous Section the question whether private R&D has been more responsive to public biological or to own-field research was unsolved because at that point the equilibrium had not been solved and the magnitude of the wage effect was unknown. We can now conclude from Table 6.3 that private-sector medical R&D investment has been more responsive to public-sector medical R&D investment than it has been to public-sector biological R&D investment.

The medical research industry's investment demand elasticities with respect to knowledge output price (p_{med}) are all positive as theory predicts. The magnitudes of these elasticities, however, are smaller than those with respect to public-sector investments in medical and biological research (h_{med}^{pub} and h_{bio}^{pub} , respectively). More strikingly, investment demand after 1990 became almost perfectly inelastic to knowledge output price, in turn implying that the post-1990 growth of industrial R&D in medical sciences was driven almost solely by productivity improvements induced through knowledge spillovers from public institutions. That result has interesting implications for the demand-pull vs. technology-push debate of the 1960's and 1970's. As reviewed in Section 2.4.1, most of the 1970's literature concluded that both demand-pull and technology-push are the major driving forces of inventive activity in modern economies. I find instead that, at least since 1980, technological opportunity rather than market demand has been the dominant driver in private-sector medical research investment. Indeed, industry investment in medical R&D has since 1990 been almost exclusively determined by technological opportunities created through publicly-funded biological and

Table 6.4 Elasticities of knowledge output supply in the medical research industry

	Elasticity of private-sector supply of knowledge output					
	1982-1990			1991-2003		
	Short-run		Long-run	Short-run		Long-run
With respect to	SEH†	QREH‡	SEH	SEH	QREH	SEH
h_{med}^{pub}	1.79	2.78	3.13	3.14	4.32	3.49
h_{bio}^{pub}	0.57	0.66	1.03	1.01	1.18	1.15
p_{med}	-0.13	-0.41	0.30	-0.18	-0.24	0.05

Note: Numbers are average elasticities over each indicated time period.
†Elasticities are computed under the static expectations hypothesis.
‡Elasticities are computed under the quasi-rational expectations hypothesis (see Table 6.3 for a precise definition.)

medical research.

Observe that, especially after 1990, elasticities estimated under the QREH are quite similar in magnitude to those estimated under the SEH. This provides part of the justification for employing the SEH in the simulations discussed below in Section 6.4.

Table 6.4 provides the medical research industry's output elasticities. A comparison of the elasticities with respect to h_{med}^{pub}, h_{bio}^{pub}, and p_{med} further confirms the conclusion drawn above about the demand-pull vs. technology-push debate. Knowledge output, measured by numbers of patents, has been more responsive to technology push than to demand pull throughout the sample period, especially after 1990.

The negative short-run output supply elasticities in table 6.4 are worth noting. The negative signs seem to contradict the standard economic theory that output supply is monotonically increasing in output price. However, a moment's reflection leads us to reject this conjecture. More favorable investment environments lead the representative firm to recruit more scientific workers. When adjustment costs are high, integrating the new with the old research workers may reduce the productivity of ongoing research programs. The loss, however, is transient: when the integration is completed, productivity will resume. Thus, the long-run output supply elasticity with respect to output price must be positive, an hypothesis confirmed in Table 6.4. The adjustment cost model does not contradict the static output supply theory, it enriches that theory.

Table 6.5 provides the estimates of the input demand and output supply elasticities in the agricultural research industry. Because the agricultural model is static, no distinction needs to be made between static and quasi-rational expectations or between short- and long-run equilibrium. As in medical research, public-sector R&D investment in agricultural research has, during the 1980 – 2003 sample period, been complementary to private-sector investment, even after accounting for the negative wage effect. Elasticities in public-sector agricultural research are uniformly higher than the corresponding elasticities in public-sector biological research. As shown in Table 6.2, public-sector agricultural R&D investment has had a much stronger spillover effect than has public-sector biological research ($\hat{\beta}_5 = 0.38$ vs. $\hat{\beta}_6 = 0.05$). Even after being neutralized by the wage effect, the strong own-field spillover effect still has generated more incentives for private investment than have spillovers from public-sector biological research.

Table 6.5 Input demand and output supply elasticities in the agricultural research industry

With respect to	*Elasticity of private-sector demand for R&D investment*		*Elasticity of private-sector supply of knowledge output*	
	1982-1990	1991-2003	1982-1990	1991-2003
h_{ag}^{pub}	1.19	1.03	1.67	1.58
h_{bio}^{pub}	0.43	0.51	0.57	0.78
p_{ag}	0.28	0.08	0.09	0.01

Note: Numbers are average elasticities over the indicated time period.

Relative magnitudes of elasticities with respect to public R&D investment and output price exhibit the same patterns in the agricultural research industry as in the medical industry. In particular, public research has played a much more important role than market demand in driving industrial agricultural R&D in the past two decades. Indeed, private agricultural R&D investment and patent output have since 1990 become almost perfectly insensitive to demand pull. Productivity improvements induced by knowledge spillovers from public-sector research have been the main cause of the growth in industrial agricultural R&D.

To gain further perspective on the contrasting characteristics of agricultural and medical research, we next compare elasticities across these two industries. Note that elasticities in the agricultural research industry are comparable only to SEH long-run elasticities in the medical research industry. Broadly speaking, the comparison suggests that R&D investment and knowledge output supply have in agricultural research been less responsive to exogenous variables than in medical research. Our earlier comparison of the parameter estimates in these two models shows that, on the one hand, public-sector own-field and biological research have exerted a stronger spillover effect on private sector research in the agricultural research industry than in the medical research industry. On the other hand, medical scientific labor supply is much more elastic than agricultural scientific labor supply. An inelastic supply implies that induced incentives for R&D investment tend to be neutralized by the rising wage. Indeed, this is exactly what happens in the agricultural research model: strong spillover effects have been offset by the strong, counteracting wage effect. The net result is that industrial investment in agricultural research has become even less responsive to exogenous stimulants than has industrial investment in medical research.

6.3 Private Rate of Return to Public Investment

Considering the predominant role that public-sector R&D investment has played in private sector R&D investment decisions, it seems useful to characterize the industrial value of public-sector R&D investment. Let us consider a policy of permanently boosting public investment, for example in the medical sciences, by $1 million per year. What is the private rate of return on this investment plan to the medical research industry? By the private rate of return to public investment, I mean the worth to the industry of a dollar of public investment. In Section 3.4 and Appendix A.4, I have shown that such a return rate may be computed as the marginal value to the industry of the investment plan divided by the marginal cost of the investment plan incurred by the public investor. For the medical research industry, I apply the formulae derived in Appendix A.4 to Model 4 in Table 6.1. For the agricultural research industry, I apply the same formulae to Model 5 in Table 6.2.

Table 6.6 Private rates of return in the medical research industry to public-sector R&D investments

	To the medical research industry			
	1982-1990	1991-2003	1982-1990	1991-2003
Rate of return on	patents/$ million		$ million/$ million	
h_{med}^{pub}	2.47	7.32	25.77	183.33
h_{bio}^{pub}	0.75	2.20	7.79	55.06

Note: Numbers are average rates of return over the indicated time period.

Table 6.7 Private rates of return in the agricultural research industry to public-sector R&D investments

	To the agricultural research industry			
	1982-1990	1991-2003	1982-1990	1991-2003
Rate of return on	patents/$ million		$ million/$ million	
h_{ag}^{pub}	1.51	2.64	12.40	36.43
h_{bio}^{pub}	0.20	0.34	1.62	4.69

Note: Numbers are the average rates of return across the indicated time period.

As shown in Table 6.6, the private rate of return to the medical research industry of public medical R&D in the 1980's was 2.47 patents (worth of $25.77 million) per $1 million of investment. The same rate of return increased in the 1990's to 7.32 patents (worth of $183.33 million) per $1 million of investment. Rates of return to public biological R&D in biological research are evidently much smaller. A similar pattern can be observed in Table 6.7 for the agricultural research industry. In particular, \$1 million of public investment in agricultural research in the 1980's produced 1.51 patents, or \$12.40 million, to the private agricultural research industry, and in the 1990's it produced 2.64 patents or \$36.43 million.

Because of the partial equilibrium nature of the present study, these results do not necessarily imply that public research resources be reallocated from basic biological to downstream medical or agricultural research. The broader social value of science is extremely difficult to measure.

In summary, public R&D investments in the life sciences have generated impressively high rates of return to the private sector, and these rates have continually risen in the last two decades. The secular increase is due in part to rising private-sector R&D and its complementary relationship with public-sector R&D, and in part to increases in patent values. The suggestion is that public direct involvement in research activity remains an effective policy tool for creating incentives for private-sector innovation.

6.4 Simulations

With the models estimated in Section 1 of this Chapter, we may simulate the effects of exogenous on endogenous variables in selected years. I conduct the simulations under the static expectations hypothesis in a deterministic context because, as we have seen in Section 6.2, elasticities obtained under static expectations approximate those under quasi-rational expectations. Moreover, the SEH is more appropriate than the QREH for policy experiment (i) considered below.

We examine how the endogenous variables respond over time to the following separate events: (i) one of the three exogenous variables — public-sector own-field R&D investment, public-sector biological R&D investment, and knowledge output price — increases by 1% in the fifth year after the system has reached equilibrium, permanently remaining at this new level; (ii) one of the three exogenous variables unexpectedly increases by 1% in the fifth year after the system has reached equilibrium, dropping back to its original level in the following year.

Simulation results are shown graphically in Figures 6.3-6.8. In each figure, the three leftmost charts describe the evolutions of the exogenous variables. Each of the three charts is matched, to its right, with three additional charts describing the evolutions of the three endogenous variables — private-sector R&D investment, private-sector knowledge output, and scientific labor wage — in response to changes in the corresponding exogenous variable. 1985 and 2003 data are chosen for these simulations in order to observe interim changes in the manner in which the endogenous variables have responded to the exogenous variables.

Figures 6.3 and 6.4 simulate event (i) in the medical research industry with 1985 and 2003 data, respectively. In both years, the endogenous variables are most responsive to a 1% change in public-sector medical R&D and least responsive to a 1% change in the price of the knowledge output. Indeed, changes in knowledge output prices have negligible effects, consistent with our discussion above of the demand-pull vs. technology-pull debate.

Because of adjustment costs, the endogenous variables approach their new equilibria gradually. Adjustments appear to be quite long, especially in private-sector R&D investments and in wages, consistent with the fact that the average adjustment parameter in the medical research industry is 0.29. Finally, comparing Figures 6.3 and 6.4 indicates each endogenous variable becomes less responsive to a given change in the exogenous variables as time passes.

Figures 6.5 and 6.6 provide the counterparts of Figures 6.3 and 6.4 for the agricultural research industry. Because no significant adjustment costs are evident in agricultural R&D investment, the endogenous variables jump to their new equilibria immediately. The manner, however, in which they respond to changes in the exogenous variables, and how those responses vary over the years, are the same as we have observed in the medical research industry in Figures 6.3 and 6.4.

Figures 6.7 and 6.8 show, respectively for the medical and agricultural research industry, how the endogenous variables respond to an unanticipated 1% increase in each of the three exogenous variables. 2003 data are used for these simulations. Not surprisingly, demand-side shocks have little impact on private-sector R&D investment decisions. Because the exogenous variable is assumed to return to its original level immediately after the shock, endogenous variables in medical research do not have time to reach their new equilibrium. Adjustment costs, that is, act as a buffer for exogenous variable shocks.

In summary, the above simulations further confirm the contribution of public investment to the growth of industry demand for R&D investment. In particular, simulations with the 2003 data suggest industrial R&D investment has remained responsive to public R&D investment, but hardly to knowledge output price, in recent years.

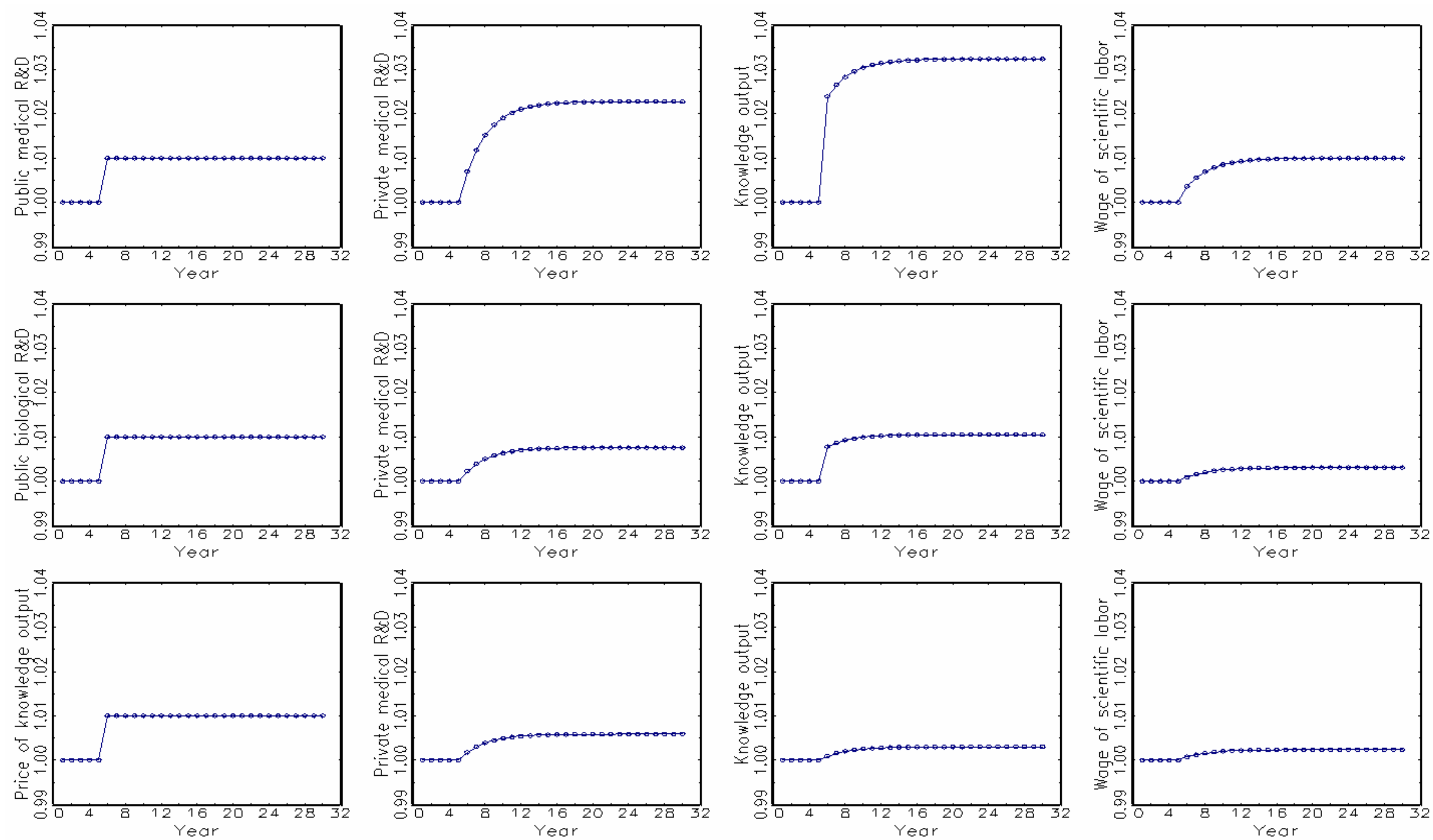

Fig. 6.3 Simulation of the effects on endogenous variables of a 1% permanent increase in public-sector medical R&D, public-sector biological R&D, and output price: medical research industry (based on data in year 1985)

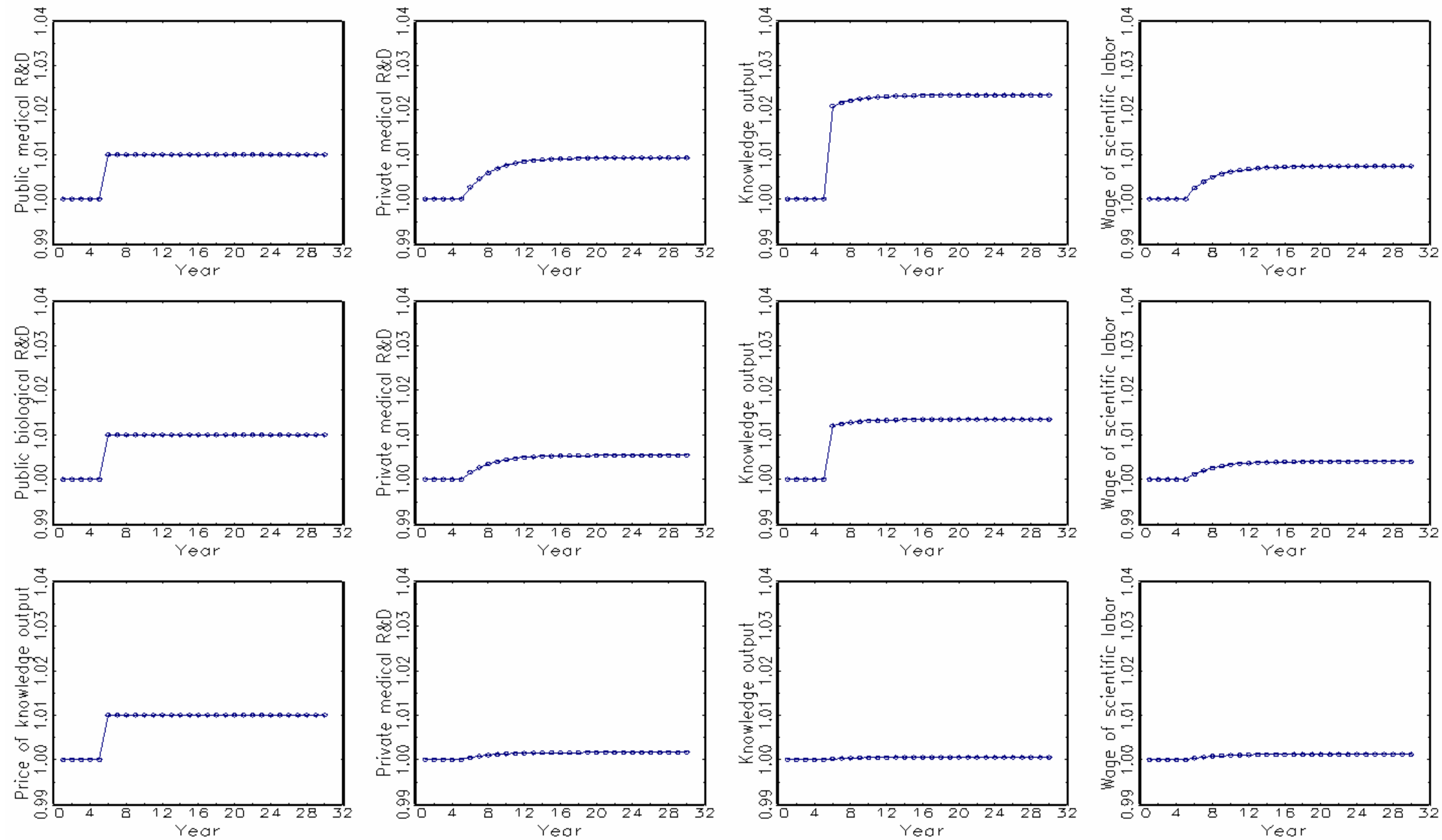

Fig. 6.4 Simulation of the effects on endogenous variables of a 1% permanent increase in public-sector medical R&D, public-sector biological R&D, and output price: medical research industry (based on data in year 2003)

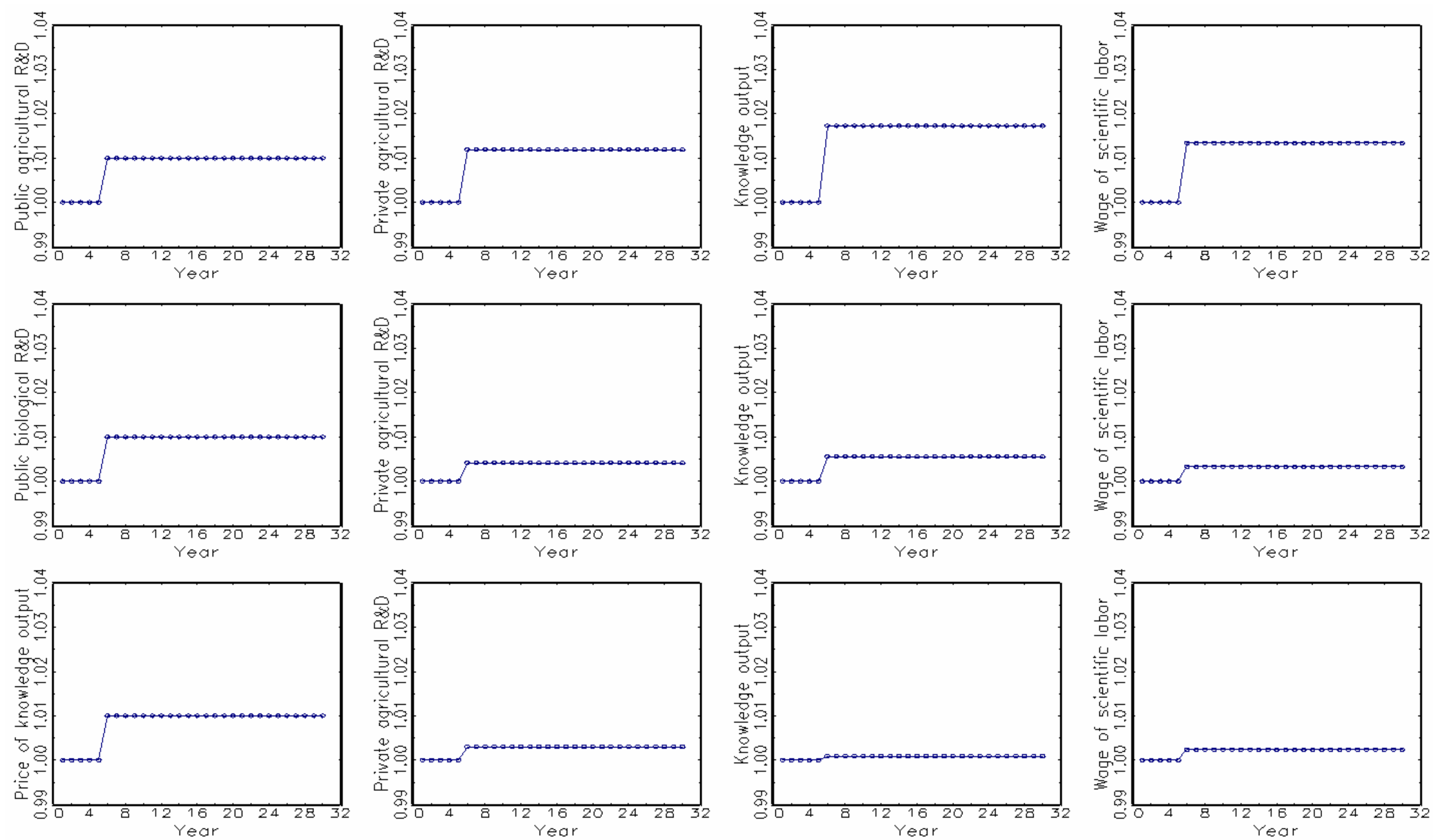

Fig. 6.5 Simulation of the effects on endogenous variables of a 1% permanent increase in public-sector agricultural R&D, public-sector biological R&D, and output price: agricultural research industry (based on data in year 1985)

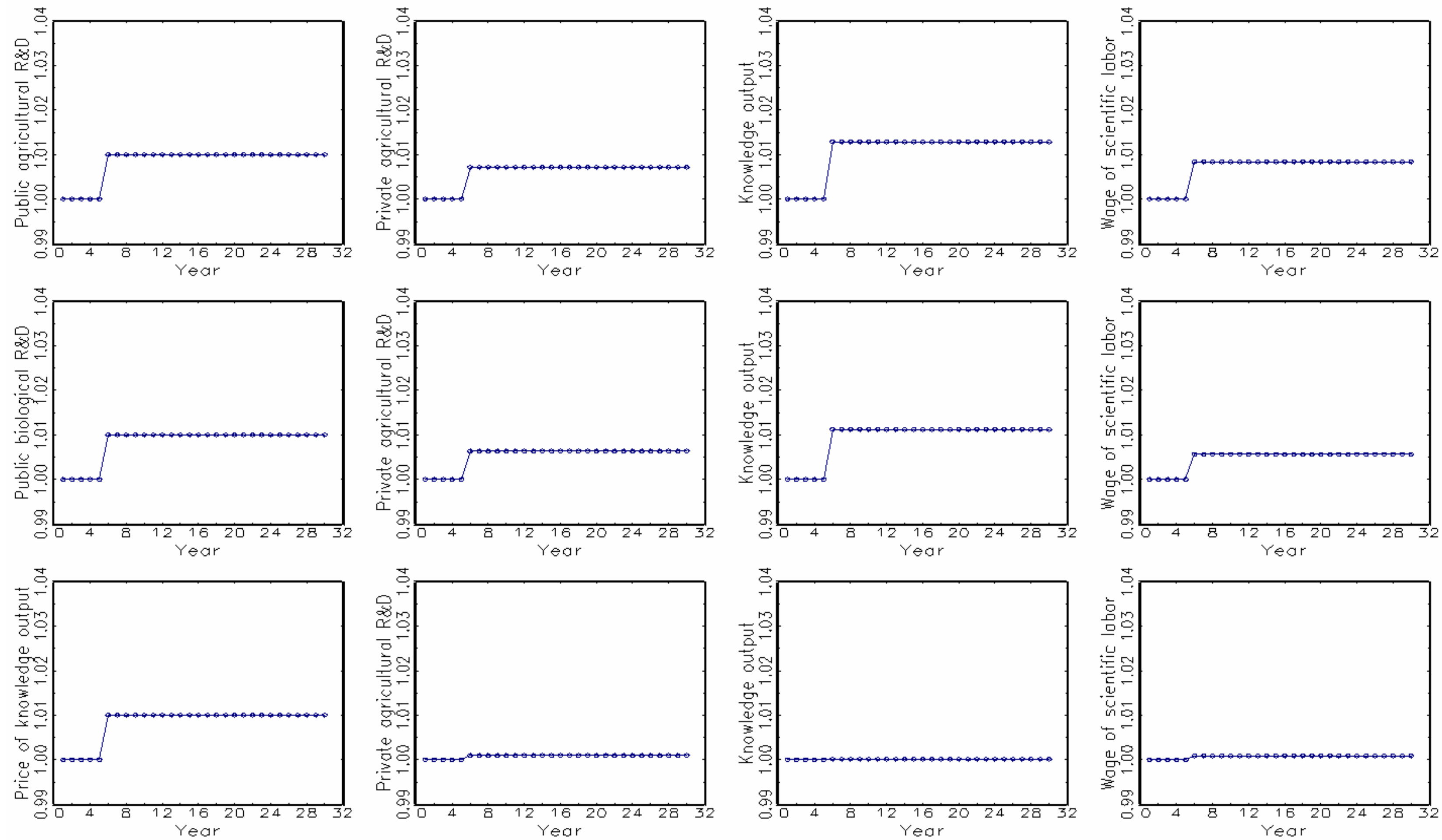

Fig. 6.6 Simulation of the effects on endogenous variables of a 1% permanent increase in public-sector agricultural R&D, public-sector biological R&D, and output price: agricultural research industry (based on data in year 2003)

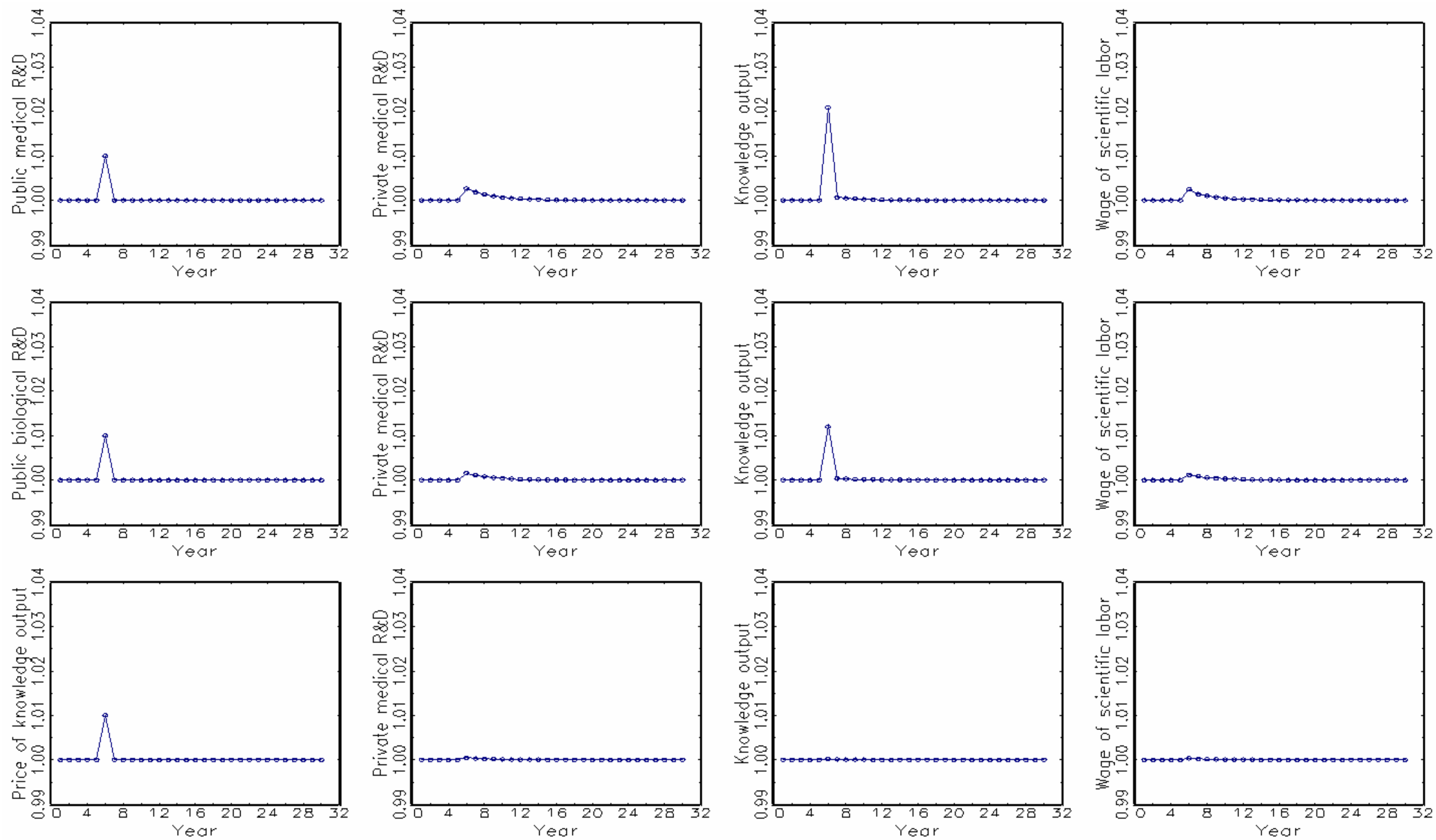

Fig. 6.7 Simulation of the effects on endogenous variables of a shock (1% increase) in public-sector medical R&D, public-sector biological R&D, and output price: medical research industry (based on data in year 2003)

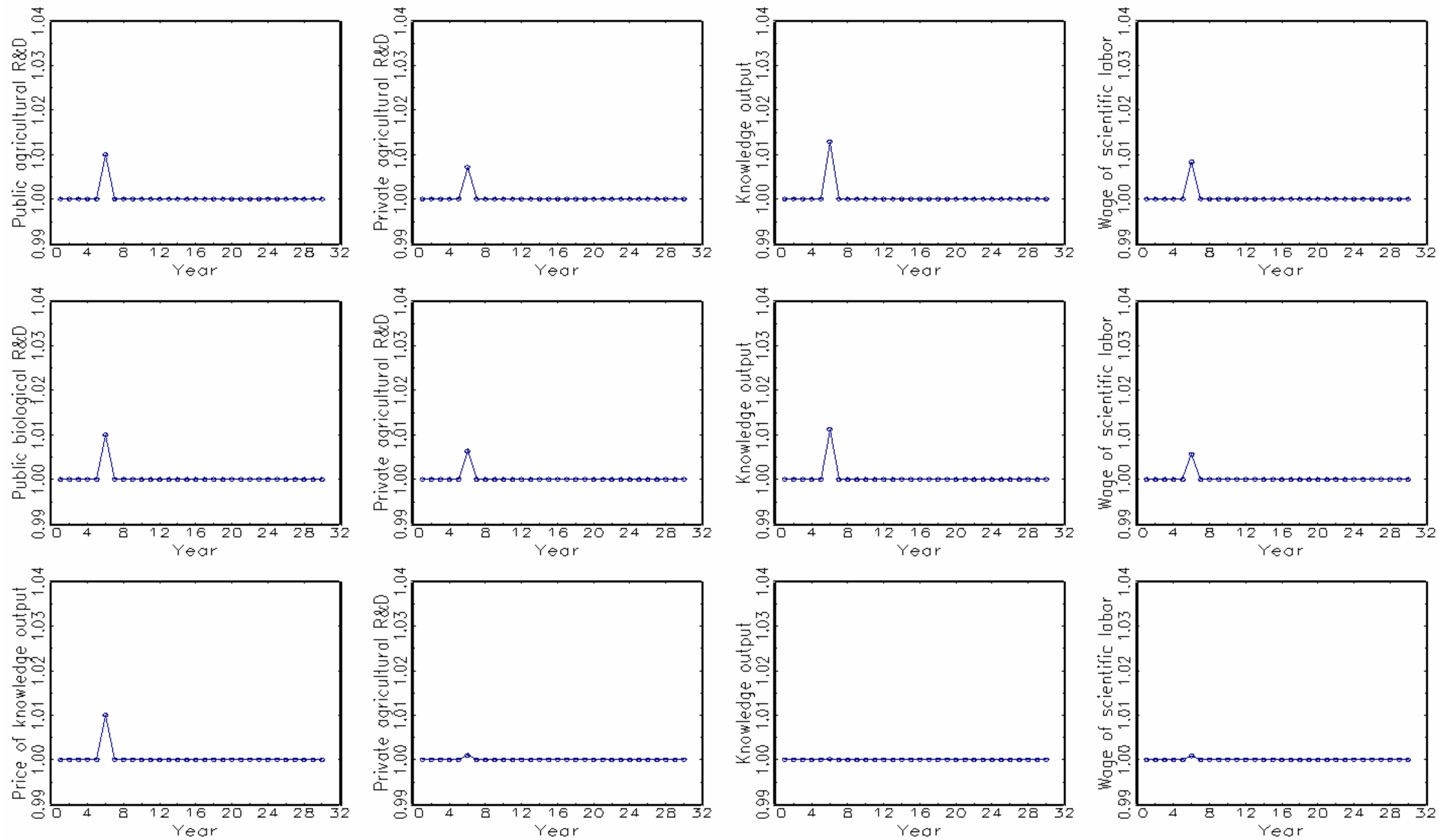

Fig. 6.8 Simulation of the effects on endogenous variables of a shock (1% increase) in public-sector agricultural R&D, public-sector biological R&D, and output price: agricultural research industry (based on data in year 2003)

Chapter 7: Conclusions

The two decades since the biorevolution of the 1970's have seen an investment surge in the life-science research of both public institutions and private firms. Over 50% of federal funds for research and development now go to the life sciences. Yet investment growth in industry has far outpaced that in public institutions, and the majority of life-science R&D now is being invested and performed in the private sector.

The objective of the present dissertation has been to evaluate the contribution of public research to industrial life-science R&D growth during the past two decades. To this end, I have developed a dynamic model of industrial R&D investment permitting a structural examination of cross-field and cross-sector knowledge spillovers, public and private interactions in the research input market, and the technological characteristics of knowledge production in the principal life-science fields. The model has been estimated with a unique dataset of R&D expenditures and patent counts constructed at the individual life-science field level.

Results show that the technology of private agricultural research differs markedly from that of medical research. The marginal productivity of R&D investment decreases at a quicker rate in agricultural than in medical research. High internal adjustment costs are evident in medical R&D investment, and estimates obtained here of medical R&D adjustment parameters are consistent with previous results. Nonsignificant adjustment costs, however, are found in agricultural R&D. The implication is that a dynamic model is appropriate for characterizing medical R&D investments, while a static model suffices for the agricultural research industry.

Estimates are provided in this dissertation of the knowledge spillovers between the three major life-science fields and between the public and private sector. Basic biological research performed in public institutions is found to have a significant “infrastructure” effect on research productivity in both agriculture and medicine. *Own-field* research performed in the public sector also has strong spillover effects on private agricultural and medical research. However, no significant knowledge spillovers are found from either public or private agricultural research to medical research, or from public or private medical research to agriculture research.

A comparison of medical and agricultural knowledge production functions shows that spillovers from public own-field and basic biological research to agricultural research are greater than they are to medical research. This result provides new perspectives on the important role of public institutions in the U.S. agricultural research system.

Public sector R&D investment may affect the private sector not only through knowledge spillovers but through competition for the same research input, scientific labor. In theory, knowledge spillovers encourage industrial investment, whereas the wage effect inhibits industry investment incentives. I find that, both in agriculture and medicine, the positive spillover effect dominated the negative wage effect during the 1980 - 2003 sample period. Altogether, public institutions’ R&D investments have been strongly complementary with those in private firms. And rates of return to private firms of public investments in the life sciences have been impressively high.

The supply of scientific labor in agricultural research is found to be much more inelastic than that in medical research, implying a stronger wage effect in agricultural than in medical research investment. In agricultural research, the strong wage effect neutralizes the strong spillover effect. As a consequence, private R&D investment in agriculture has been less responsive to exogenous factors than it has in medicine.

Technological opportunities created through public research have been the dominating force in private R&D investment, swamping any market-demand effects throughout the sample period. Indeed, after 1990, industry investment in both agriculture and medicine became almost perfectly inelastic to market demand, driven instead by technological opportunities created through public-sector research.

Evidence of public investment’s predominant role in creating incentives for industrial investment in life-science research has important implications for government science and technology policy. During the past two decades, the federal government has greatly expanded life-science research investment in response to rising public demands surrounding health and food issues. I have shown this investment strategy to be highly successful in creating technological opportunities for life-science firms. Simulation results indicate public investment in the life sciences will continue to be extremely valuable for the private sector. R&D tax credits may, of course, have exerted similar knowledge spillover effects, and represent alternative policy instruments for stimulating industry R&D investment.

This study has its limitations. It has not taken into explicit account any changes during the sample period in industry financing opportunities or in the strength of intellectual property rights in the life sciences. Those issues raise important difficulties in modeling and variable construction, which would have greatly complicated the present framework. On the other hand, although no rigorous study has yet simultaneously considered all four determinants of private R&D investment – market demand, technological opportunity, scientific labor supply, and adjustment costs – the present research has taken a significant step in that direction.

Bibliography

ADAMS, J. D. (1990): "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy*, 98, 673-702.

ARROW, K. J. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity*, ed. by R. Nelson.

BERNSTEIN, J. I. (1989): "The Structure of Canadian Inter-Industry R & D Spillovers, and the Rates of Return to R & D," *The Journal of Industrial Economics*, 37, 315-328.

BERNSTEIN, J. I., and M. I. NADIRI (1988): "Interindustry R&D Spillovers, Rates of Return, and Production in High-Tech Industries," *The American Economic Review*, 78, 429-434.

— (1989): "Research and Development and Intra-Industry Spillovers: An Empirical Application of Dynamic Duality," *The Review of Economic Studies*, 56, 249-267.

CINCERA, M., and P. BRUNO VAN POTTELSBERGHE DE LA (2001): "International R&D Spillovers: A Survey," *Brussels Economic Review/Cahiers Economiques de Bruxelles*, 0, 3-31.

DASGUPTA, P., and J. STIGLITZ (1980): "Industrial Structure and the Nature of Innovative Activity," *Economic Journal*, 90, 266-93.

DAVID, P. A., and B. H. HALL (2000): "Heart of Darkness: Modeling Public-Private Funding Interactions inside the R&D Black Box," *Research Policy*, 29, 1165-1183.

DAVID, P. A., B. H. HALL, and A. A. TOOLE (2000): "Is Public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence," *Research Policy*, 29, 497-529.

EISNER, R., and R. H. STROTZ (1963): "Determinants of Business Investment," in *Impacts of Monetary Policy*. Englewood Cliffs, N.J.: Prentice-Hall.

EPSTEIN, L. G., and M. G. S. DENNY (1983): "The Multivariate Flexible Accelerator Model: Its Empirical Restrictions and an Application to U.S. Manufacturing," *Econometrica: Journal of the Econometric Society*, 51, 647-674.

GRABOWSKI, H. G. (1968): "The Determinants of Industrial Research and Development: A Study of the Chemical, Drug, and Petroleum Industries," *The Journal of Political Economy*, 76, 292-306.

GRILICHES, Z. (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10, 92-116.

— (1981): "Market Value, R&D, and Patents," *Economics Letters*, 7, 183-187.

— (1990): "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 28, 1661-1707.

— (1992): "The Search for R&D Spillovers," *Scandinavian Journal of Economics*, 94, S29-47.

HAAVELMO, T. (1960): *A Study in the Theory of Investment*. Chicago: University of Chicago Press.

HALL, B. H. (1992): "Investment and Research and Development at the Firm Level: Does the Source of Financing Matter?," University of California at Berkeley.

— (2005): "The Financing of Innovation," in *Blackwell Handbook of Technology and Innovation Management*, ed. by S. Shane. Oxford: Blackwel Publishers Ltd.

HALL, B. H., A. JAFFE, and M. TRAJTENBERG (2005): "Market Value and Patent Citations," *RAND Journal of Economics*, 36, 16-38.

HALL, B. H., and W. F. LONG (1999): "Differences in Reported R&D Data on the Nsf/Census Rd-1 Form and the Sec 10-K Form: A Micro-Data Investigation."

HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029-1054.

HANSEN, L. P., and T. J. SARGENT (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2, 7-46.

HANSEN, L. P., and K. J. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269-86.

HENDERSON, R., and I. COCKBURN (1993): "Scale, Scope and Spillovers: The Determinants of Research Productivity in the Pharmaceutical Industry," *NBER working paper #4466*.

HIMMELBERG, C. P., and B. C. PETERSEN (1994): "R & D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries," *The Review of Economics and Statistics*, 76, 38-51.

HUFFMAN, W. E., and R. E. EVENSON (1993): *Science for Agriculture: A Long-Term Perspective*. Blackwell.

JAFFE, A. B. (1986): "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value," *American Economic Review*, 76, 984-1001.

JORGENSON, D. W. (1963): "Capital Theory and Investment Behavior," *The American Economic Review*, 53, 247-259.

KLOTZ, C., K. FUGLIE, and C. PRAY (1995): "Private-Sector Agricultural Research Expenditures in the United States, 1960-92," USDA, Economic Research Service.

LANJOUW, J. O., A. PAKES, and J. PUTNAM (1998): "How to Count Patents and Value Intellectual Property:The Uses of Patent Renewal and Application Data," *The Journal of Industrial Economics*, 46, 405-432.

LEVIN, R. C., and P. C. REISS (1984): "Tests of a Schumpeterian Model of R&D and Market Structure," in *R&D, Patents, and Productivity*, ed. by Z. Griliches. Chicago: University of Chicago Press, 175-204.

LICHTENBERG, F. R. (1988): "Estimation of the Internal Adjustment Costs Model Using Longitudinal Establishment Data," *The Review of Economics and Statistics*, 70, 421-430.

LUCAS, R. E., JR. (1967): "Optimal Investment Policy and the Flexible Accelerator," *International Economic Review*, 8, 78-85.

LUCAS, R. E., JR., and E. C. PRESCOTT (1971): "Investment under Uncertainty," *Econometrica*, 39, 659-681.

MOHNEN, P. (1996): "R&D Externalities and Productivity Growth," Paris: OECD, 39-66.

NATIONAL INSTITUTE OF HEALTH (2006): "Revision of Biomedical Research and Development Price Index: Revised Fy 2005 Update and Projections for Fy 2006-2016."

NATIONAL SCIENCE FOUNDATION (2003): "National Patterns of Research and Development Resources: 2003," Arlington, VA 2005: National Science Foundation, Division of Science Resources Statistics.

— (2005): "Federal Funds for Research and Development: Fiscal Years 2002, 2003, and 2004," Arlington, VA: National Science Foundation, Division of Science Resources Statistics.

NELSON, R. R. (1959): "The Simple Economics of Basic Scientific Research," *The Journal of Political Economy*, 67, 297-306.

OGAKI, M. (1993): "Generalized Method of Moments: Econometric Applications," in *Handbook of Statistics*, ed. by G. S. Maddala, C. R. Rao, and H. D. Vinod. Amsterdam: North-Holland, 455-488.

PAKES, A., and M. SCHANKERMAN (1984): "An Exploration into the Determinants of Research Intensity," in *R&D, Patents, and Productivity*, ed. by Z. Griliches. Chicago: University of Chicago Press, 209-259.

PAKES, A. S. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, 755-84.

PINDYCK, R. S., and J. J. ROTEMBERG (1983): "Dynamic Factor Demands and the Effects of Energy Price Shocks," *American Economic Review*, 73, 1066-79.

ROMER, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71-S102.

ROSENBERG, N. (1974): "Science, Invention and Economic Growth," *The Economic Journal*, 84, 90-108.

RUTTAN, V. W. (1997): "Induced Innovation, Evolutionary Theory and Path Dependence: Sources of Technical Change," *The Economic Journal*, 107, 1520-1529.

SARGENT, T. J. (1978): "Estimation of Dynamic Labor Demand Schedules under Rational Expectations," *Journal of Political Economy*, 86, 1009-44.

— (1987): *Macroeconomic Theory*. New York: Academic Press, Inc.

SCHERER, F. M. (1982): "Demand-Pull and Technological Invention: Schmookler Revisted," *The Journal of Industrial Economics*, 30, 225-237.

— (1984): "Using Linked Patent and R&D Data to Measure Technology Flows," in *R&D, Patents, and Productivity*, ed. by Z. Griliches. Chicago: University of Chicago Press, 417-464.

SCHMOOKLER, J. (1966): *Invention and Economic Growth*. Cambridge: Harvard University Press.

SCHUMPETER, J. (1950): *Captalism, Socialism, and Democracy*. New York: Harper and Row.

SOLOW, R. M. (1957): "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 39, 312-320.

TREADWAY, A. B. (1971): "The Rational Multivariate Flexible Accelerator," *Econometrica: Journal of the Econometric Society*, 39, 845-856.

WANG, C., Y. XIA, R. SHOEMAKER, and S. BUCCOLA (2006): "Public Investment Policy in Life-Science Research," *Unpublished Manuscript*.

XIA, Y. (2002): "Science and Technology in Cutting-Edge Agricultural Biotechnology Research," Corvallis: Oregon State University, PhD Dissertation.

APPENDICES

Appendix A: Mathematical Details for Chapter 3

A.1 Equilibrium Solution under the Quasi-Rational Expectations Hypothesis

Following Sargent (1987, pp. 391-396), we solve the stochastic Euler equations (3.8) forward using the unstable root in order to impose the transversality condition (3.6), and then solve it backward using the stable root to impose the initial condition.

Assume $\{h_{med,t}^{pub}\}$, $\{h_{bio,t}^{pub}\}$, and $\{h_{ag,t}^{pp}\}$ are exogenous stochastic processes of exponential order less than $1/\sqrt{r}$, i.e., for some $M > 0$ and $1 \le s \le 1/\sqrt{r}$,

$$\left|E_t h_{med,t+j}^{pub}\right| < Ms^{t+j}, \ \left|E_t h_{bio,t+j}^{pub}\right| < Ms^{t+j}, \ \left|E_t h_{ag,t+j}^{pp}\right| < Ms^{t+j}, \tag{A.1.1}$$

for all t and $j \ge t$. Rewrite equation (3.8) as

$$(B^{-2} - \theta B^{-1} + \frac{1}{r})E_t \overline{h}_{med,t-1}^{prv} = E_t z_t \,, \tag{A.1.2}$$

where the operator B is defined by $B^{-j} E_t x_t = E_t x_{t+j}$. Let $\lambda_1 + \lambda_2 = \theta$ and $\lambda_1 \lambda_2 = \frac{1}{r}$ so that (A.1.2) can be written as

$$(B^{-1} - \lambda_1)(B^{-1} - \lambda_2) E_t \overline{h}_{med,t-1}^{prv} = E_t z_t \,. \tag{A.1.3}$$

Without loss of generality, let $\lambda_1 \le \lambda_2$. Under the assumptions $\alpha_2 > 0$, $\beta_3 < 0$, and $\beta_4 < 0$, it is straightforward to show $0 < \lambda_1 < 1 < \frac{1}{r} < \lambda_2$. Dividing both sides of (A.1.3) by $B^{-1} - \lambda_2$ yields

$$\bar{h}^{prv}_{med,t} = \lambda_1 \bar{h}^{prv}_{med,t-1} + \frac{E_t z_t}{B^{-1} - \lambda_2} = \lambda_1 \bar{h}^{prv}_{med,t-1} - \frac{1}{\lambda_2} \frac{E_t z_t}{1 - B^{-1}/\lambda_2}$$
$$= \lambda_1 \bar{h}^{prv}_{med,t-1} - \frac{1}{\lambda_2} \sum_{i=0}^{\infty} (\frac{1}{\lambda_2})^i E_t z_{t+i} + c\lambda_2^{\ i}, \tag{A.1.4}$$

where c is constant. Since $\lambda_2 > 1$, we must have $c = 0$ in order for the transversality condition (3.6) to be satisfied. This gives expression (3.9) in Section 3 of Chapter 3.

After replacing least square predictions with expectations to get (3.13), the final solution to the problem can be obtained by solving (A.1.4) backward using the stable root λ_1. Write (3.13) as $\bar{h}^{prv}_{med,t} = \frac{1}{1-\lambda_1 L}\eta_t$, where

$$\eta_t = -\frac{1}{r\beta_4}[\frac{\beta_2 - \alpha_1/p_{med,t}}{\lambda_2 - 1} + \frac{(\beta_5 - \alpha_2/p_{med,t})(c_1/(\lambda_2 - 1) + h^{pub}_{med,t})}{\lambda_2 - \rho_1}$$
$$+ \frac{\beta_6(c_2/(\lambda_2 - 1) + h^{pub}_{bio,t})}{\lambda_2 - \rho_2} + \frac{\beta_7(c_3/(\lambda_2 - 1) + h^{pp}_{ag,t})}{\lambda_2 - \rho_3}].$$

With $h^{prv}_{med,-1}$ given, this can be expressed as

$$\bar{h}^{prv}_{med,t} = \lambda_1^{\ t+1} h^{prv}_{med,-1} + \sum_{i=0}^{t} \lambda_1^{\ t} \eta_{t-i}, \ t = 0,1,2\cdots \tag{A.1.5}$$

It is straightforward to confirm that (A.1.5) satisfies the transversality condition (3.6):

$$\lim_{t\to\infty} r^t E_0\{[\beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t} - \frac{\bar{w}_{med,t}}{p_{med,t}}$$
$$+ \beta_3 \bar{h}^{prv}_{med,t} + \beta_4(\bar{h}^{prv}_{med,t} - \bar{h}^{prv}_{med,t-1})]\bar{h}^{prv}_{med,t}\} = 0$$

To see this, notice that by (A.1.1) and (A.1.5), $\bar{h}^{prv}_{med,t}$ is a sequence of exponential order less than $1/\sqrt{r}$, i.e., $\left|\bar{h}^{prv}_{med,t}\right| < Ms^t$, where $M > 0$ and $1 \le s \le 1/\sqrt{r}$; and so is $\bar{w}_{med,t}$ by the market clear condition (3.4). Since by assumption $p_{med,t} = p_{med,0}$ for $t \ge 0$, the transversality condition is satisfied. Hence, (A.1.5) is the unique solution for the equilibrium sequence $\{\bar{h}^{prv}_{med,t}\}$. Substituting from (A.1.5) into the market clear condition (3.4) yields the unique solution for the equilibrium sequence $\{\bar{w}_{med,t}\}$.

A.2 Elasticities under the Quasi-Rational Expectations Hypothesis

As shown in Chapter 3, it is simple to derive the elasticities of the private sector's equilibrium R&D investment with respect to the public sector's R&D investment. The elasticity with respect to knowledge price $p_{med,t}$, however, is a bit more complicated than the other two to derive, because the equilibrium representation of R&D investment (3.12) is highly nonlinear in $p_{med,t}$. What follows is the derivation of this elasticity.

Notice first that $\lambda_1 = 0.5(\theta - \sqrt{\theta^2 - 4/r})$, $\lambda_2 = \frac{1}{\lambda_1 r}$, and $\theta = 1 + \frac{1}{r} + \frac{\beta_3}{r\beta_4} - \frac{\alpha_2}{r\beta_4 p_{med,t}}$,

so that λ_1 and λ_2 are nonlinear functions of $p_{med,t}$. The derivates of the former two variables with respect to the latter are

$$\begin{aligned}\frac{d\lambda_1}{dp_{med,t}} &= \frac{0.5\alpha_2(1-\theta/\sqrt{\theta^2-4/r})}{r\beta_4 p_{med,t}{}^2},\\ \frac{d\lambda_1}{dp_{med,t}} &= \frac{0.5\alpha_2(1+\theta/\sqrt{\theta^2-4/r})}{r\beta_4 p_{med,t}{}^2}.\end{aligned} \tag{A.2.1}$$

Differentiating $\bar{h}^{prv}_{med,t}$ with respect to $p_{med,t}$ in (3.12) gives

$$\frac{d\bar{h}_{med,t}^{prv}}{dp_{med,t}} = \bar{h}_{med,t-1}^{prv}\frac{d\lambda_1}{dp_{med,t}} - \frac{1}{r\beta_4}[-\frac{\beta_2 - \alpha_1/p_{med,t}}{(\lambda_2-1)^2}\frac{d\lambda_2}{dp_{med,t}} + \frac{\alpha_1/p_{med,t}{}^2}{\lambda_2-1}$$
$$-\frac{(\beta_5-\alpha_2/p_{med,t})(c_1/(\lambda_2-1)+h_{med,t}^{pub})}{(\lambda_2-\rho_1)^2}\frac{d\lambda_2}{dp_{med,t}}$$
$$+\frac{\frac{\alpha_1}{p_{med,t}{}^2}(\frac{c_1}{\lambda_2-1}+h_{med,t}^{pub}) - \frac{c_1(\beta_5-\alpha_2/p_{med,t})}{(\lambda_2-1)^2}\frac{d\lambda_2}{dp_{med,t}}}{\lambda_2-\rho_1} \tag{A.2.2}$$
$$-\frac{\beta_6(\frac{c_2}{\lambda_2-1}+h_{bio,t}^{pub})}{(\lambda_2-\rho_2)^2} - \frac{\frac{c_2\beta_6}{(\lambda_2-1)^2}\frac{d\lambda_2}{dp_{med,t}}}{\lambda_2-\rho_2}].$$

Substituting from (A.2.1) into (A.2.2) and multiplying by $p_{med,t}/\bar{h}_{med,t}^{prv}$ gives the expression of the elasticity $e_{m,p}^{d}$.

The knowledge output supply function $y(h_{med,t-1}^{prv}, h_{med,t}^{pub}, h_{bio,t}^{prv}, h_{ag,t}^{pp}, p_{med,t})$ can be obtained by substituting the equilibrium R&D investment (3.12) into the knowledge production function (3.2). Apply the chain rule to obtain the elasticities of private-sector knowledge supply in medical sciences with respect to $h_{med,t}^{pub}, h_{bio,t}^{prv},$ and $p_{med,t}$ as follows:

$$e_{m,m}^{s} = \beta_5\frac{\bar{h}_{med,t}^{prv}h_{med,t}^{pub}}{f(\bar{h}_{med,t}^{prv},\bullet)} + \frac{\partial f(\bar{h}_{med,t}^{prv},\bullet)}{\partial h_{med,t}^{prv}}\frac{\bar{h}_{med,t}^{prv}}{f(\bar{h}_{med,t}^{prv},\bullet)}e_{m,m}^{d},$$
$$e_{m,b}^{s} = \beta_6\frac{\bar{h}_{med,t}^{prv}h_{bio,t}^{pub}}{f(\bar{h}_{med,t}^{prv},\bullet)} + \frac{\partial f(\bar{h}_{med,t}^{prv},\bullet)}{\partial h_{med,t}^{prv}}\frac{\bar{h}_{med,t}^{prv}}{f(\bar{h}_{med,t}^{prv},\bullet)}e_{m,b}^{d}, \tag{A.2.3}$$
$$e_{m,p}^{s} = \frac{\partial f(\bar{h}_{med,t}^{prv},\bullet)}{\partial h_{med,t}^{prv}}\frac{\bar{h}_{med,t}^{prv}}{f(\bar{h}_{med,t}^{prv},\bullet)}e_{m,p}^{d}.$$

A.3 Elasticities under the Static Expectations Hypothesis

Under the static expectations hypothesis, the short-run and long-run elasticities are derived from (3.15) and (3.16), respectively. The short-run elasticities of private-sector R&D investment in medical research $h_{med,t}^{prv}$ with respect to public investment in medical research $h_{med,t}^{pub}$, public investment in biological research $h_{bio,t}^{pub}$, and knowledge price $p_{med,t}$, respectively, are:

$$\tilde{e}_{m,m}^{d} = -\frac{(\beta_5 - \alpha_2 / p_{med,t})\, h_{med,t}^{pub}}{r\beta_4(\lambda_2 - 1)\tilde{h}_{med,t}^{prv}},$$

$$\tilde{e}_{m,b}^{d} = -\frac{\beta_6 h_{bio,t}^{pub}}{r\beta_4(\lambda_2 - 1)\tilde{h}_{med,t}^{prv}},$$

$$\begin{aligned}\tilde{e}_{m,p}^{d} = \{ & \tilde{h}_{med,t-1}^{prv} \frac{d\lambda_1}{dp_{med,t}} + \frac{1}{r\beta_4(\lambda_2 - 1)^2}[\beta_2 - \frac{\alpha_1}{p_{med,t}} \\ & + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h_{med,t}^{pub} + \beta_6 h_{bio,t}^{pub}] \\ & - \frac{1}{r\beta_4(\lambda_2 - 1)}[\frac{\alpha_1}{p_{med,t}{}^2} + \frac{\alpha_2}{p_{med,t}{}^2} h_{med,t}^{pub}]\} \frac{p_{med,t}}{\tilde{h}_{med,t}^{prv}}.\end{aligned}$$

The corresponding long run elasticities are:

$$\hat{e}_{m,m}^{d} = \frac{(\beta_5 - \alpha_2 / p_{med,t})h_{med,t}^{pub}}{(\alpha_2 / p_{med,t} - \beta_3)\hat{h}_{med,t}^{prv}},$$

$$\hat{e}_{m,b}^{d} = \frac{\beta_6 h_{bio,t}^{pub}}{(\alpha_2 / p_{med,t} - \beta_3)\hat{h}_{med,t}^{prv}},$$

$$\begin{aligned}\hat{e}_{m,p}^{d} = \{ & \frac{\alpha_2 / p_{med,t}{}^2}{(\alpha_2 / p_{med,t} - \beta_3)^2}[\beta_2 - \frac{\alpha_1}{p_{med,t}} + (\beta_5 - \frac{\alpha_2}{p_{med,t}})h_{med,t}^{pub} + \beta_6 h_{bio,t}^{pub}] \\ & - \frac{1}{\alpha_2 / p_{med,t} - \beta_3}[\frac{\alpha_1}{p_{med,t}{}^2} + \frac{\alpha_2}{p_{med,t}{}^2} h_{med,t}^{pub}]\} \frac{p_{med,t}}{\tilde{h}_{med,t}^{prv}}.\end{aligned}$$

Finally, the short-run and long-run output supply elasticities are analogous to (A.2.3).

A.4 Private Rate of Return on Public-Sector R&D Investment

Consider the policy plan described in Chapter 3. The public sector announces a permanent increase in medical R&D investment by 1 $ million at period 0. The questions of interest are 1) what is the shadow value of this policy plan to the private sector, and 2) what is the "private rate of return on public R&D investment in medical sciences," by which we mean the ratio of shadow value to the marginal cost of this policy plan.

To begin, rewrite (3.17) in its open-loop form:

$$\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0) = \lambda_1^{t+1} h^{prv}_{med,-1} + (1-\lambda_1^{t+1})\hat{h}^{prv}_{med,0}, \quad t = 0,1,2\cdots \tag{A.4.1}$$

where $\boldsymbol{\omega}_0 = (h^{prv}_{med,-1}, h^{pub}_{med,0}, h^{pub}_{bio,0}, h^{pp}_{ag,0}, p_{med,0})'$ are the initial values of the exogenous state variables. The corresponding equilibrium wage is

$$\tilde{w}_{med,t}(\boldsymbol{\omega}_0) = \alpha_1 + \alpha_2[\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0) + h^{pub}_{med,0}], \quad t = 0,1,2\cdots \tag{A.4.2}$$

The value function of the problem (3.3) is

$$V^{prv}_{med,0}(\boldsymbol{\omega}_0) = \sum_{t=0}^{+\infty} r^t g(\tilde{h}^{prv}_{med,t-1}(\boldsymbol{\omega}_0), \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0) \tag{A.4.3}$$

where

$$\begin{aligned} & g(\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0) \\ = {} & \beta_1 + (\beta_2 + \beta_5 h^{pub}_{med,t} + \beta_6 h^{pub}_{bio,t} + \beta_7 h^{pp}_{ag,t} - \frac{\tilde{w}_{med,t}(\boldsymbol{\omega}_0)}{p_{med,t}})\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0) \\ & + 0.5\beta_3[\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)]^2 + 0.5\beta_4[\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0) - \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)]^2 \end{aligned} \tag{A.4.4}$$

and $\tilde{h}^{prv}_{med,-1}(\boldsymbol{\omega}_0) = h^{prv}_{med,-1}$. Differentiate the value function (A.4.3) with respect to $h^{pub}_{med,0}$ and apply the envelop condition to obtain

$$\frac{dV^{prv}_{med,0}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} = \sum_{t=0}^{+\infty} r^t [\frac{\partial g(\tilde{h}^{prv}_{med,t-1}(\boldsymbol{\omega}_0), \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0)}{\partial w_{med,t}} \frac{d\tilde{w}_{med,t}(\boldsymbol{\omega}_0)}{dh^{pub}_{med}} + \frac{\partial g(\tilde{h}^{prv}_{med,t-1}(\boldsymbol{\omega}_0), \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0)}{\partial h^{pub}_{med,t}}]. \tag{A.4.5}$$

The derivatives in the brackets are

$$\frac{\partial g(\tilde{h}^{prv}_{med,t-1}(\boldsymbol{\omega}_0), \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0)}{\partial w_{med}} = -\frac{\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)}{p_{med,t}}$$

$$\frac{\partial g(\tilde{h}^{prv}_{med,t-1}(\boldsymbol{\omega}_0), \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \tilde{w}_{med,t}(\boldsymbol{\omega}_0), \boldsymbol{\omega}_0)}{\partial h^{pub}_{med}} = \beta_5 \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0),$$

and

$$\frac{d\tilde{w}_{med,t}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} = \alpha[\frac{d\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} + 1]$$

where by (A.4.1) and (3.16)

$$\frac{d\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} = (1-\lambda_1^{i+1}) \frac{d\hat{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} = (1-\lambda_1^{i+1}) \frac{\beta_5 - \alpha_2 / p_{med,0}}{\alpha_2 / p_{med,0} - \beta_3}.$$

Substituting the last expression into (A.4.5) yields

$$\begin{aligned} \frac{dV^{prv}_{med,0}(\boldsymbol{\omega}_0)}{dh^{pub}_{med,0}} &= \sum_{t=0}^{+\infty} r^t \{-\alpha_2 \frac{\tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)}{p_{med,t}} [(1-\lambda_1^{i+1}) \frac{\beta_5 - \alpha_2 / p_{med,0}}{\alpha_2 / p_{med,0} - \beta_3} + 1] + \beta_5 \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0)\} \\ &= (\beta_5 - \alpha_2 / p_{med,0}) \sum_{t=0}^{+\infty} r^t [1 - (1-\lambda_1^{i+1}) k] \tilde{h}^{prv}_{med,t}(\boldsymbol{\omega}_0), \end{aligned} \tag{A.4.6}$$

where $k = \dfrac{\alpha_2 / p_{med,0}}{\alpha_2 / p_{med,0} - \beta_3}$. Now substitute from (A.4.1) into (A.4.6) to obtain

$$\begin{aligned}\frac{dV_{med,0}^{prv}(\boldsymbol{\omega}_0)}{dh_{med,0}^{pub}} &= (\beta_5 - \alpha_2 / p_{med,0})\sum_{t=0}^{+\infty} r^t [1-(1-\lambda_1^{\ t+1})k][\lambda_1^{\ t+1} h_{med,-1}^{prv} + (1-\lambda_1^{\ t+1})\hat{h}_{med}^{prv}], \\ &= (\beta_5 - \alpha_2 / p_{med,0})\{h_{med,-1}^{prv}\sum_{t=0}^{+\infty} r^t [\lambda_1^{\ t+1} - \lambda_1^{\ t+1}(1-\lambda_1^{\ t+1})k] \\ &\qquad + \hat{h}_{med}^{prv}\sum_{t=0}^{+\infty} r^t [(1-\lambda_1^{\ t+1}) - (1-\lambda_1^{\ t+1})^2 k]\} \\ &= (\beta_5 - \alpha_2 / p_{med,0})(m_1 h_{med,-1}^{prv} + m_2 \hat{h}_{med}^{prv}),\end{aligned} \tag{A.4.7}$$

where

$$m_1 = \frac{\lambda_1}{1-r\lambda_1} - \frac{\lambda_1(1-\lambda_1)}{(1-r\lambda_1)(1-r\lambda_1^{\ 2})}k,$$

$$m_2 = \frac{1-\lambda_1}{(1-r)(1-r\lambda_1)} - \frac{r\lambda_1(1-\lambda_1)^2}{(1-r)(1-r\lambda_1)(1-r\lambda_1^{\ 2})}k\,.$$

Dividing the shadow value by the marginal cost of the investment plan, namely, $1/(1-r)$, gives the private rate of return on public R&D investment in medical sciences.

Similarly, the private rate of return (to the medical research industry) on public R&D investment in biological sciences is

$$\frac{dV_{med,0}^{prv}(\boldsymbol{\omega}_0)}{dh_{bio,0}^{pub}}(1-r) = \beta_6 (m_1 h_{med,-1}^{prv} + m_2 \hat{h}_{med}^{prv})(1-r)\,.$$

Appendix B: Robustness of Estimation Results

This appendix presents the parameter estimates for the medical and agricultural R&D model, obtained using four alternative sets of instrument variables and using alternative discount rates. The four sets of instrument variables are: $I_{j,a} = \{z^1_{j,1t}, z^1_{,2t}, z_{j,3t}\}$, $I_{j,b} = \{z^2_{j,1t}, z^1_{j,2t}, z_{j,3t}\}$, $I_{j,c} = \{z^1_{j,1t}, z^2_{j,2t}, z_{j,3t}\}$, and $I_{j,d} = \{z^2_{j,1t}, z^2_{j,2t}, z_{j,3t}\}$, where $j = med$ for the medical research industry and $j = ag$ for the agricultural research industry. Specifically,

$$z^1_{med,1t} = (1, p_{med,t})',$$

$$z^2_{med,1t} = (1, p_{med,t}, h^{pub}_{bio,t}, h^{pp}_{ag,t},)',$$

$$z^1_{med,2t} = (1, t, h^{pub}_{med,t}h^{prv}_{med,t}, h^{pub}_{bio,t}h^{prv}_{med,t}, h^{pp}_{ag,t}h^{prv}_{med,t}, 0.5(h^{prv}_{med,t})^2, 0.5(h^{prv}_{med,t} - h^{prv}_{med,t-1})^2)',$$

$$z^2_{med,2t} = (1, t, h^{pub}_{med,t}h^{prv}_{med,t}, h^{pub}_{bio,t}h^{prv}_{med,t}, h^{pp}_{ag,t}h^{prv}_{med,t}, 0.5(h^{prv}_{med,t})^2, 0.5(h^{prv}_{med,t} - h^{prv}_{med,t-1})^2, p_{med,t}, w_{med,t})',$$

$$z_{med,3t} = (h^{pub}_{med,t}, h^{pub}_{bio,t}, h^{pp}_{ag,t}, h^{prv}_{med,t-1}, h^{prv}_{med,t}, w_{med,t}, p_{med,t})'.$$

Substituting the subscript med for ag and vice versa in the above equations gives the definitions of $z^1_{ag,1t}, z^2_{ag,1t}, z^1_{ag,2t}, z^2_{ag,2t}$, and $z_{ag,3t}$.

Tables B.1-4 contain parameter estimates for the medical research industry. The discount factor is assumed to be 0.95. The parameter estimates seem robust across the four sets of instrument variables. I conduct specification tests in the same fashion as described in Chapter 6. In all four sets of results, I fail to reject the restriction in Model 4, i.e., no knowledge spillovers from agricultural research—be it public- or private-sector—

to private-sector research. All the remaining parameters but β_3 in Model 4 have the right signs across the four tables. The estimates of β_3 in Model 4 of Tables B1, B2, and B4 are negative, but that in Model 4 of Table B3 is positive. This latter result violates the concavity assumption for technology, though the estimate of β_3 is not significantly different from zero at the 5% level. Although the parameter estimates of β_3 in Model 4 of other three tables have the right signs, only the one in Table B.2 is significantly different from zero at the 5% level. This implies that only inconclusive evidence is found for concave technology, or decreasing returns to scale in R&D investment. Fortunately, Model 4 of Table B2 yields the smallest mean squared prediction error, compared to Model 4 in other tables. Hence, I adopt it for further comparative statics analysis and simulations in Chapter 6.

Tables B.5-8 are the counterparts of Tables B.1-4 for the agricultural research industry. Restrictions in Model 5 fail to be rejected in all four tables, meaning that in the agricultural research industry, there are no adjustment costs in R&D investment, no trending technological or institutional shocks, and no knowledge spillovers from medical research. Parameter estimates have the right signs for all tables. Model 5 in Table B.6 yields the smallest mean squared prediction error, and therefore is adopted for further analysis in Chapter 6.

Table B.9 contains the parameter estimates of Model 4 in Table B2, when the discount factor is 0.90, 0.95, and 0.99. Table B.10 contains the parameter estimates of Model 5 in Table B6, when the discount factor are 0.90, 0.95, and 0.99. It is evident that the choice of the discount factor, at least in the range between 0.90 to 0.99, does not alter the parameter estimates much both in the agricultural model and in the medical model.

Table B.1 Parameter estimates with I_a : medical industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*
α_1	0.3856	0.3442	0.3828	0.3855
	(0.0153)	(0.0142)	(0.0153)	(0.0153)
α_2	0.0099	0.0109	0.0098	0.0099
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
β_1	3.5546	0.8290	3.3633	3.5970
	(0.4624)	(0.2788)	(0.4605)	(0.4480)
β_2	-0.0683	0.2609	-0.5277	-0.1090
	(0.1201)	(0.1115)	(0.0633)	(0.0477)
β_3	-0.0012	0.0031	-0.0148	-0.0023
	(0.0033)	(0.0032)	(0.0013)	(0.0015)
β_4	-0.0429	-0.0343		-0.0405
	(0.0095)	(0.0095)		(0.0069)
β_5	0.0339	0.0177	0.0577	0.0351
	(0.0083)	(0.0080)	(0.0064)	(0.0077)
β_6	0.0188	0.0070	0.0451	0.0213
	(0.0073)	(0.0072)	(0.0044)	(0.0028)
β_7	-0.0034	-0.0350	0.0249	
	(0.0091)	(0.0080)	(0.0066)	
β_8	-0.4791		-0.4434	-0.4904
	(0.0649)		(0.0644)	(0.0573)
J	11.7254	66.3003	31.9994	11.8621
P-value	0.1100	0.0000	0.0001	0.1575
d.f.	7	8	8	8

†Note: $I_a = \{z^1_{med,1t}, z^1_{med,2t}, z_{med,3t}\}$

Table B.2 Parameter estimates with I_b : medical industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*
α_1	0.3852	0.3411	0.3957	0.3837
	(0.0101)	(0.0091)	(0.0098)	(0.0088)
α_2	0.0101	0.0107	0.0097	0.0101
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
β_1	5.7310	-0.8324	5.4541	5.5571
	(0.8044)	(0.4379)	(0.8013)	(0.5765)
β_2	-0.3785	-0.1469	-0.7529	-0.3532
	(0.1209)	(0.1185)	(0.0734)	(0.0892)
β_3	-0.0056	-0.0112	-0.0168	-0.0052
	(0.0030)	(0.0030)	(0.0010)	(0.0028)
β_4	-0.0445	0.0016		-0.0457
	(0.0114)	(0.0104)		(0.0107)
β_5	0.0820	0.0797	0.1178	0.0815
	(0.0136)	(0.0135)	(0.0099)	(0.0135)
β_6	0.0251	0.0310	0.0408	0.0244
	(0.0045)	(0.0045)	(0.0021)	(0.0038)
β_7	0.0024	-0.0358	0.0126	
	(0.0077)	(0.0066)	(0.0073)	
β_8	-0.9140		-0.7617	-0.8991
	(0.0940)		(0.0855)	(0.0808)
J	15.1417	109.7443	30.3407	15.2377
P-value	0.0871	0.0000	0.0008	0.1236
d.f.	9	10	10	10

†Note: $I_b = \{z^2_{med,1t}, z^1_{med,2t}, z_{med,3t}\}$

Table B.3 Parameter estimates with I_c : medical industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*
α_1	0.3821	0.3655	0.3761	0.3804
	(0.0120)	(0.0115)	(0.0119)	(0.0118)
α_2	0.0101	0.0105	0.0100	0.0100
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
β_1	5.9335	0.5801	3.7347	6.4458
	(1.1110)	(0.5050)	(1.0619)	(0.9751)
β_2	0.1221	0.4701	-0.5651	-0.0264
	(0.1715)	(0.1589)	(0.1378)	(0.0748)
β_3	0.0067	0.0135	-0.0126	0.0033
	(0.0042)	(0.0040)	(0.0031)	(0.0023)
β_4	-0.0880	-0.0584		-0.0851
	(0.0131)	(0.0119)		(0.0127)
β_5	0.0402	0.0385	0.0879	0.0442
	(0.0093)	(0.0093)	(0.0060)	(0.0083)
β_6	0.0012	-0.0236	0.0327	0.0101
	(0.0097)	(0.0086)	(0.0085)	(0.0031)
β_7	-0.0138	-0.0660	0.0090	
	(0.0144)	(0.0107)	(0.0140)	
β_8	-0.9645		-0.4617	-1.0796
	(0.1783)		(0.1619)	(0.1323)
J	12.7201	41.9850	58.0256	13.6458
P-value	0.1757	0.0000	0.0000	0.1898
d.f.	9	10	10	10

†Note: $I_c = \{z^1_{med,1t}, z^2_{med,2t}, z_{med,3t}\}$

Table B.4 Parameter estimates with I_d : medical industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*
α_1	0.3857	0.3427	0.3962	0.3868
	(0.0084)	(0.0069)	(0.0081)	(0.0049)
α_2	0.0100	0.0108	0.0098	0.0100
	(0.0002)	(0.0002)	(0.0002)	(0.0001)
β_1	6.3430	-1.0879	5.3895	6.4315
	(0.9070)	(0.3916)	(0.8845)	(0.7246)
β_2	-0.2147	-0.3569	-0.7614	-0.2315
	(0.1299)	(0.1290)	(0.0601)	(0.0788)
β_3	-0.0012	-0.0087	-0.0175	-0.0016
	(0.0038)	(0.0037)	(0.0016)	(0.0026)
β_4	-0.0631	-0.0029		-0.0620
	(0.0133)	(0.0115)		(0.0115)
β_5	0.0691	0.1013	0.1179	0.0700
	(0.0114)	(0.0108)	(0.0049)	(0.0097)
β_6	0.0164	0.0197	0.0412	0.0173
	(0.0068)	(0.0068)	(0.0044)	(0.0040)
β_7	-0.0011	-0.0264	0.0147	
	(0.0066)	(0.0060)	(0.0057)	
β_8	-1.0215		-0.7556	-1.0292
	(0.1125)		(0.0975)	(0.1019)
J	15.0424	97.5415	37.5683	15.0688
P-value	0.1806	0.0000	0.0002	0.2377
d.f.	11	12	12	12

†Note: $I_d = \{z^2_{med,1t}, z^2_{med,2t}, z_{med,3t}\}$

Table B.5 Parameter estimates with I_a: agricultural industry[†]

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
α_1	-0.1349	-0.1726	-0.1847	-0.1705	-0.1764
	(0.0921)	(0.0869)	(0.0865)	(0.0869)	(0.0860)
α_2	0.1038	0.1083	0.1084	0.1066	0.1071
	(0.0118)	(0.0112)	(0.0112)	(0.0111)	(0.0111)
β_1	2.2946	0.8662	1.4449	1.6532	1.6990
	(1.5272)	(0.9935)	(0.9151)	(0.8754)	(0.8704)
β_2	-0.9664	-0.0625	-0.3934	-0.3969	-0.4543
	(0.8165)	(0.3577)	(0.2812)	(0.2969)	(0.2729)
β_3	-0.0502	-0.1733	-0.1041	-0.0837	-0.0770
	(0.1258)	(0.0764)	(0.0608)	(0.0545)	(0.0527)
β_4	0.2897	0.3042		0.0735	
	(0.2037)	(0.2034)		(0.1496)	
β_5	0.3540	0.2554	0.2824	0.2432	0.2585
	(0.1595)	(0.1380)	(0.1368)	(0.1378)	(0.1342)
β_6	0.0535	0.0498	0.0495	0.0535	0.0524
	(0.0061)	(0.0054)	(0.0053)	(0.0049)	(0.0043)
β_7	-0.0026	0.0021	0.0008		
	(0.0040)	(0.0012)	(0.0009)		
β_8	0.0937				
	(0.0761)				
J	5.1934	6.7100	8.9479	9.5153	9.7567
P-value	0.6364	0.5682	0.4421	0.3911	0.4621
d.f.	7	8	9	9	10

[†]Note: $I_a = \{z^1_{ag,1t}, z^1_{ag,2t}, z_{ag,3t}\}$

Table B.6 Parameter estimates with I_b: agricultural industry[†]

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
α_1	-0.1260	-0.1276	-0.1370	-0.1112	-0.1204
	(0.1290)	(0.1290)	(0.1285)	(0.1258)	(0.1252)
α_2	0.1047	0.1043	0.1062	0.1022	0.1041
	(0.0161)	(0.0161)	(0.0159)	(0.0157)	(0.0155)
β_1	0.3840	1.0433	0.9900	1.1108	1.0580
	(1.2837)	(0.7189)	(0.7157)	(0.7091)	(0.7059)
β_2	-0.0819	-0.4981	-0.4504	-0.5324	-0.4848
	(0.7295)	(0.2852)	(0.2788)	(0.2788)	(0.2723)
β_3	-0.1799	-0.1167	-0.1225	-0.1062	-0.1119
	(0.1127)	(0.0479)	(0.0473)	(0.0443)	(0.0437)
β_4	0.0175	-0.0681		-0.0685	
	(0.1626)	(0.0860)		(0.0860)	
β_5	0.3439	0.3889	0.3793	0.3879	0.3782
	(0.1232)	(0.0996)	(0.0988)	(0.0995)	(0.0988)
β_6	0.0446	0.0458	0.0464	0.0471	0.0477
	(0.0043)	(0.0038)	(0.0037)	(0.0030)	(0.0029)
β_7	0.0025	0.0005	0.0005		
	(0.0034)	(0.0008)	(0.0008)		
β_8	-0.0334				
	(0.0539)				
J	11.6386	12.0230	12.6497	12.3481	12.9833
P-value	0.2345	0.2835	0.3168	0.3381	0.3703
d.f.	9	10	11	11	12

[†]Note: $I_b = \{z^2_{ag,1t}, z^1_{ag,2t}, z_{ag,3t}\}$

Table B.7 Parameter estimates I_c : agricultural industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
α_1	-0.1476	-0.1648	-0.1792	-0.1582	-0.1645
	(0.0701)	(0.0635)	(0.0630)	(0.0635)	(0.0621)
α_2	0.1055	0.1076	0.1078	0.1050	0.1055
	(0.0087)	(0.0079)	(0.0079)	(0.0078)	(0.0078)
β_1	1.5720	0.9675	1.6037	2.0472	2.0877
	(1.4416)	(0.9959)	(0.9230)	(0.8604)	(0.8563)
β_2	-0.5117	-0.0631	-0.4777	-0.5237	-0.5922
	(0.8550)	(0.3642)	(0.2705)	(0.2947)	(0.2580)
β_3	-0.1297	-0.1847	-0.1069	-0.0640	-0.0571
	(0.1221)	(0.0768)	(0.0617)	(0.0525)	(0.0505)
β_4	0.3678	0.3691		0.0824	
	(0.2170)	(0.2170)		(0.1714)	
β_5	0.3232	0.2578	0.3051	0.2409	0.2608
	(0.1810)	(0.1416)	(0.1389)	(0.1414)	(0.1352)
β_6	0.0504	0.0491	0.0487	0.0549	0.0538
	(0.0058)	(0.0053)	(0.0053)	(0.0046)	(0.0039)
β_7	0.0004	0.0027	0.0014		
	(0.0042)	(0.0013)	(0.0010)		
β_8	0.0478				
	(0.0825)				
J	7.2991	7.6355	10.5283	12.2720	12.5034
P-value	0.6060	0.6644	0.4836	0.3435	0.4061
d.f.	9	10	11	11	12

†Note: $I_c = \{z^1_{ag,1t}, z^2_{ag,2t}, z_{ag,3t}\}$

Table B.8 Parameter estimates I_d : agricultural industry†

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
α_1	-0.1044	-0.0990	-0.1031	-0.0615	-0.0691
	(0.0940)	(0.0939)	(0.0938)	(0.0879)	(0.0857)
α_2	0.1024	0.1006	0.1017	0.0962	0.0974
	(0.0117)	(0.0115)	(0.0115)	(0.0109)	(0.0104)
β_1	0.2690	1.5021	1.7110	1.5405	1.6444
	(1.2719)	(0.6820)	(0.6303)	(0.6811)	(0.6258)
β_2	0.2074	-0.6298	-0.6078	-0.6774	-0.6589
	(0.7746)	(0.2622)	(0.2607)	(0.2588)	(0.2543)
β_3	-0.2208	-0.0997	-0.0858	-0.0868	-0.0814
	(0.1127)	(0.0399)	(0.0359)	(0.0382)	(0.0356)
β_4	0.0981	-0.0652		-0.0288	
	(0.1638)	(0.0813)		(0.0747)	
β_5	0.2767	0.3809	0.3372	0.3876	0.3637
	(0.1369)	(0.1025)	(0.0868)	(0.1023)	(0.0816)
β_6	0.0447	0.0469	0.0490	0.0496	0.0503
	(0.0043)	(0.0039)	(0.0029)	(0.0031)	(0.0025)
β_7	0.0050	0.0009	0.0006		
	(0.0036)	(0.0008)	(0.0007)		
β_8	-0.0676				
	(0.0588)				
J	12.1586	13.4776	14.1218	14.7638	14.9131
P-value	0.3518	0.3353	0.3653	0.3223	0.3841
d.f.	11	12	13	13	14

†Note: $I_d = \{z^2_{ag,1t}, z^2_{ag,2t}, z_{ag,3t}\}$

Table B.9 Robustness over alternative discount factors

	Model 4 in Table B. 2		
	r=0.90	*r=0.95*	*r=0.99*
α_1	0.3836	0.3837	0.3835
	(0.0087)	(0.0088)	(0.0089)
α_2	0.0101	0.0101	0.0101
	(0.0002)	(0.0002)	(0.0002)
β_1	5.4601	5.5571	5.6258
	(0.5806)	(0.5765)	(0.5807)
β_2	-0.3396	-0.3532	-0.3606
	(0.0834)	(0.0892)	(0.0938)
β_3	-0.0049	-0.0052	-0.0053
	(0.0027)	(0.0028)	(0.0028)
β_4	-0.0465	-0.0457	-0.0454
	(0.0110)	(0.0107)	(0.0106)
β_5	0.0788	0.0815	0.0832
	(0.0123)	(0.0135)	(0.0144)
β_6	0.0242	0.0244	0.0244
	(0.0038)	(0.0038)	(0.0038)
β_7			
β_8	-0.8780	-0.8991	-0.9165
	(0.0804)	(0.0808)	(0.0819)
J	15.3129	15.2377	15.2273
P-value	0.1211	0.1236	0.1240
d.f.	10	10	10

Table B.10 Robustness over alternative discount factors

	Model 5 in Table B.6		
	r=0.90	*r=0.95*	*r=0.99*
α_1	-0.1243	-0.1204	-0.1178
	(0.1256)	(0.1252)	(0.1249)
α_2	0.1046	0.1041	0.1038
	(0.0156)	(0.0155)	(0.0155)
β_1	1.1647	1.0580	0.9687
	(0.7121)	(0.7059)	(0.7013)
β_2	-0.4890	-0.4848	-0.4792
	(0.2777)	(0.2723)	(0.2684)
β_3	-0.1067	-0.1119	-0.1164
	(0.0441)	(0.0437)	(0.0434)
β_4			
β_5	0.3633	0.3782	0.3898
	(0.0998)	(0.0988)	(0.0983)
β_6	0.0486	0.0477	0.0470
	(0.0030)	(0.0029)	(0.0029)
β_7			
β_8			
J	13.2924	12.9833	12.7797
P-value	0.3481	0.3703	0.3853
d.f.	12	12	12

Appendix C: Supplementary Details for Chapter 5

C.1 Selected DWPI Patent Classes for the Life Sciences

B Pharmaceutical

B01 Steroids - including systems containing carbocyclic and/or heterocyclic rings fused onto the basic steroidal ring structure.

B02 Fused ring heterocyclics.

B03 Other heterocyclics.

B04 Natural products and polymers. Including testing of body fluids (other than blood typing or cell counting), pharmaceuticals or veterinary compounds of unknown structure, testing of microorganisms for pathogenicity, testing of hemicals for mutagenicity or human toxicity and fermentative production of DNA or RNA. General compositions.

B05 Other organics - aromatics, aliphatic, organo-metallics, compounds whose substituents vary such that they would be classified in several of B01 - B05.

B06 Inorganics - including fluorides for toothpastes etc.

C Agricultural Chemicals

C01 Organophosphorus; organometallic - i.e. compounds containing other than H, C, N, O, S and halogen.

C02 Heterocyclic. C03 Other organic compounds, inorganic compounds and multicomponent mixtures. Polymers and proteins.

C04 Fertilisers - including urea and phosphoric acid production. Also soil modifiers and plant growth media. Chemical aspects of compost production.

C05 Biological control - excluding veterinary medicine, but including use of microorganisms, predators and natural products.

C06 Biotechnology - including plant genetics and veterinary vaccines.

D Food, Detergents, Water Treatment and Biotechnology

D13 Other foodstuffs and treatment – including preservation of food, milk, milk products, butter substitutes, edible oils and fats, non-alcoholic beverages, artificial sweeteners, food additives and animal feed (A23B-L).

D15 Chemical or biological treatment of water, industrial waste and sewage – including purification, sterilising or testing water, scale prevention, treatment of sewage sludge, regeneration of active carbon which has been used for water treatment and impregnating water with gas e.g. CO2, but excluding plant and anti-pollution devices (C02).

D16 Fermentation industry – including fermentation equipment, brewing, yeast production, production of pharmaceuticals and other chemicals by fermentation, microbiology, production of vaccines and antibodies, cell and tissue culture and genetic engineering.

P General

P11 Soil working, planting (A01B, C).

P12 Harvesting (A01D, F).

P13 Plant culture, dairy products (A01G, H, J).

P14 Animal care (A01K, L, M).

P15 Tobacco (A24).

C.2 Formulas and Definitions for Computing the Book Values and Market Values of Life-Science Firms

Following Hall et al. (1988), a firm's book value is defined as the sum of the net plant and equipment, the inventory, and the investments in unconsolidated subsidiaries, intangibles, and others. A firm's market value is defined as the sum of the value of the common stock, the value of the preferred stock, the value of the long-term debt, and the value of short-term debt net of assets.

BOOKVALUE = BKPLNT + BKINV + TOTAL

TOTAL = TOTAL1 + TOTAL2 + INTANG

MARKET VALUE = PREFST + VCOMS + LTDEBT + STDEBT – ADJ= PREFST + NONSHARE * PCLOSE + BKDEBT + STDEBT – (CURRASST – BKINV – STLIAB + STDEBT)= PREFST + NONSHARE * PCLOSE + BKDEBT – CURRASST + BKINV + STLIAB

BKPLNT – DATA8 (Million Dollars): Property, plant, and equipment – total (net)

BKINV – DATA3 (Millions of Dollars): Book value of inventory

TOTAL1 – DATA31 (Million Dollars): Investments and advances – equity method

TOTAL2 – DATA32 (Million Dollars): Investments and advances – other

INTANG – DATA33 (Million Dollars): Intangibles

PREFST – DATA10 (Million Dollars): Preferred stock (liquidating value)

NONSHARE – DATA25 (Millions): The number of common shares outstanding

PCLOSE – DATA24 (Dollars and cents): End of calendar year stock price

STDEBT – DATA34 (Million Dollars): Debt portion of current liabilities

STLIAB – DATA5 (Million Dollars): Current liabilities – total

CURRASST – DATA4 (Million Dollars): Current assets – total

PREFD – DATA19 (Million Dollars): Dividends on the preferred stock

BKDEBT – DATA9 (Million Dollars): Book value of long-term debt