

R&D Competition in the Chemicals Industry.

Stephen R. Finger*[†]

Duke University

Job Market Paper

November 20, 2007

Abstract

This paper examines the effect of the Research and Experimentation Tax Credit on the Chemicals Industry. The subsidy program was put into place to counteract the underinvestment in private R&D caused by firms not internalizing the benefits of technological spillovers from their research. However, the rationale ignored the impact of competition in the product market. I propose and estimate a structural dynamic oligopoly model of competition in intellectual assets to capture the impact of interactions between firms in the industry. I estimate the dynamic parameters of the model using methods from Bajari, Benkard, and Levin (2007). In the data, I observe firms that persistently invest more in research and generate more innovations than other firms that are observationally similar. I build upon previous estimators by incorporating this unobserved firm-level heterogeneity using techniques from Arcidiacono and Miller (2007). I find that increased investment in R&D by more advanced firms due to the subsidy, was largely offset by decreases by smaller firms because of the substitutability of knowledge in product market. This greatly reduced the effectiveness of the policy to spur innovation and limited its impact on social welfare.

*I would like to thank Andrew Sweeting, Paul Ellickson, Chris Timmins, and Tracy Lewis for their helpful comments. All remaining errors are my own.

[†]Department of Economics, Duke University Box 90097, Durham, NC 27708. Phone: 919-660-8160. srf17@duke.edu.

1 Introduction.

Technological innovation is a key determinant to economic growth. Growth in the economy and social welfare can both be increased through endogenously produced technological advancements. Spillovers in the research process have been a main focus of the innovation literature because they enable the possibility of long run increasing returns and have also have led economists, starting with Arrow (1962), to believe that firms underinvest in R&D from a social welfare standpoint. In response to this concern, governments in many countries including the US subsidize research expenditures. The effectiveness of these policies will depend not only on the cost elasticity of R&D at the firm level, but also on competitive interactions between firms. Not only are there externalities involved in the creation of intellectual assets, but there are externalities involved in the translation of intellectual assets into financial profits. Product innovations may result in complementarities if innovations increase the horizontal differentiation between goods. However process innovations and vertical differentiation can cause “creative destruction” reducing the profits of previous discoveries. If R&D efforts are strategic substitutes then an increase in spending by one firm will be partially offset by decreases by other firms, reducing the potential impact of subsidy programs. Competitive interactions are likely to be particularly important in concentrated oligopolies, which are the typical market structures of industries with high-levels of sunk R&D costs.

In this paper I analyze the returns to the government funded subsidy program in the Chemicals Industry by estimating a structural dynamic oligopoly model. I find that the dynamic competitive effects, which are ignored in all of the existing literature, have a significant impact on the results of policies. In particular, while the subsidy currently used in the US raises research expenditures and innovation rates of the largest firms, there are offsetting reductions in the efforts of smaller firms. This reduces the potential impact on innovation by approximately one-third and the impact on social welfare by almost two-thirds.

I find that firms differ in their abilities to utilize public information for their own gain. Even in the presence of patents, the non-rival nature of knowledge implies that firms may learn from advancements by their competitors. However, I find that outside firms must have the expertise necessary to assimilate and exploit the information. This result mirrors Cohen and Levinthal

(1989)'s idea of "absorbative capacity". Therefore more advanced firms, not only have more accumulated knowledge of their own to use in the search for new innovations, but they can also more efficiently utilize external ideas for their own gain. This is one of the reasons that an across the board subsidy has heterogenous effects on different firms. The government's policy does produce social welfare gains through increased aggregate innovation, however the effect on consumer surplus is lessened due decreased investment by less advanced firms and its impact on market structure with the benefits concentrated on the largest firms.

My results do not necessarily contradict the much studied Schumpeterian hypothesis that the first order issue regarding social welfare is the pace of innovation and not inefficiencies due to market concentration. However, I find that because own and outside innovations are substitutes in the product market, the subsidy causes many firms in the industry to decrease their investment in R&D, offsetting some of the gains in the pace of innovation. This then exacerbates the inefficiencies from market concentration in the static product market, further reducing the social benefits of the policy.

The Chemicals Industry is an appropriate subject for an assessment of R&D subsidies because firms in the industry are focused on innovation. Firms invest a high percentage of their revenues in research, account for approximately one in nine patents, and receive 8% of the government research tax credits¹. In addition, chemical products are used as intermediate goods in many other manufacturing segments implying that innovations in the industry contribute to growth in other sectors of the economy as well.

My analysis is structured around a dynamic oligopoly model in the spirit of Ericson and Pakes (1995). I model the Chemical Industry as a single homogeneous market with marginal costs differentiated by firms' level of achieved innovation. My results support this modeling assumption as I find that outside knowledge is a substitute for a firm's own knowledge even when I relax the structural restrictions in the profit function. If research in the industry actually led to increased horizontal product differentiation, I would have likely found a stronger complementarity between innovations from rival firms.

Firms invest in R&D which through a stochastic process produces innovations. Firms then are able to generate profits in the product market based on their level of intellectual capital and

¹National Science Foundation (2000).

the levels of their rivals. I motivate the payoff relevant state space, and then assume that firms follow symmetric Markov-Perfect Nash equilibrium (MPNE) strategies. I estimate the primitives of the dynamic model using a two-step estimation strategy first proposed by Hotz and Miller (1993) and laid out in a multi-player, oligopoly setting in Bajari, Benkard, and Levin (2007).

In the data I observe firms that persistently invest more in research and generate more innovations than other firms that are observationally similar. Existing dynamic games estimators cannot allow for this persistent firm level heterogeneity. I demonstrate how to account for this using techniques of Arcidiacono and Miller (2007). I model the heterogeneity as an unobserved state that raises a firm's research productivity. I am able to identify the distribution of firms across states by using the empirical correlation between abnormal R&D expenditures and unusually high realized innovations over time. I show that allowing for just a two state form of the unobserved heterogeneity substantially improves the fit of the model.

I estimate the model using panel data on firms' R&D expenditures and their patenting activities. The limited nature of the data is both an impediment and a strength of the analysis. It requires me to model firms as competing in a homogeneous good market where innovations may only distinguish firms in a single dimension. This is a simplification of reality which I hope to address in future work with better data. However, the data I use is also widely available for many other industries, making it relatively straightforward for other researchers to test whether my results can be found in different settings.

1.1 Literature Review

This paper is related to the current literature in three different areas. First of all, I improve upon previous papers examining the effectiveness of R&D subsidies. Because of the difficulty in measuring social returns, past papers have used reduced form analysis to measure increased R&D expenditures under the subsidy versus the cost of the subsidy. Hall (1995) provides an overview of this literature. Papers such as Berger (1993) and Eisner, Albert, and Sullivan (1984) look at the variation in industry level R&D spending after the implementation of the subsidy. Industry aggregation does not capture the distribution of firm-level effects which may have significant impacts on both the likely research outcomes and dynamic market structure. These papers also

are not able to analyze the social welfare impact of the policy. Hall (1993) approximated the marginal increase in research using estimates of the marginal elasticity R&D spending based on firm-specific cost shifters. This type of methodology ignores the dynamic competitive effects of a research subsidy that is applied to all firms simultaneously. Finding that a single firm will increase research investment when its own costs fall, holding the costs of its rivals constant, does not imply that all firms will increase R&D spending when each of their costs declines.

My analysis addresses these issues by taking into consideration equilibrium effects as firms not only react to the changes in the cost of their own research, but also anticipate decisions by their rivals. I also find firm-level differences in marginal research efficiency due to dynamic considerations which likely cause aggregate measures of R&D to misstate changes in innovation and social welfare.

It is through recent advancements in the estimation of dynamic games that I am able to estimate my model. Following the seminal work on estimating dynamic discrete choice models, Rust (1987), the major hurdle to using dynamic oligopoly models was the computational burden in solving for the MPNE. This issue is addressed using a two-step approach that allows the parameters of a dynamic oligopoly model to be estimated without the calculation of equilibria. Hotz and Miller (1993) was one of the first papers to demonstrate this method in a single-agent setting, and number of recent papers including Arcidiacono and Miller (2007), Aguirregabiria and Mira (2007), and Pesendorfer and Schmidt-Dengler (2007) extend the estimator to multi-player scenarios. I follow Bajari, Benkard, and Levin (2007) because it allows me to model the firms' states and research decisions as continuous variables. The idea of the two-step estimators is that we assume that we observe firms acting optimally. Therefore, after directly estimating optimal actions from the data in the first step, we can recover the primitives of the model that rationalize those decisions in a second step.

I build upon the this literature by including unobserved heterogeneity in a two-step estimator. Firm level unobserved heterogeneity has been used in the labor literature to account for persistent differences in actions and macro models incorporating regime change, but has not been used in an industrial organization setting. Aguirregabiria and Mira (2007) demonstrates how unobserved market fixed effects can be included in dynamic games, but do not look at differences at the firm-level. In the past it has been difficult to maintain the computational feasibility of

dynamic oligopoly models while incorporating unobserved firm heterogeneity, but the computational simplicity of two-stage estimators combined with the advancements in Arcidiacono and Miller (2007) allow for this with minimal additional burden.

This paper also builds upon the patent data literature by being the first paper to use the data in a structural framework. Economists have long sought to use patents as a source of information on the value of innovations. However, beginning with Schmookler (1966) much evidence has been established indicating that while patents are a good indicator of innovative inputs, or R&D, but they are much less correlated with the innovative output of the research process. Improving on these measures, there have been many studies focusing on the information inherent in patent citations. Similar to citations to academic journal articles, the number of citations a patent receives is a good indicator to the value of the underlying innovation. Studies such as Trajtenberg (1990), Harhoff et al. (1999), Hall, Jaffe, and Trajtenberg (2005), and Lanjouw and Schankerman (2004) detail the informational value in the patent data set to measure the firm level innovation, but it has not yet be used in a structural setting. Given the number of theoretical papers on R&D that stress the impact of externalities on firm decisions, it is important to use this data in a setting which incorporates firm interactions. Bloom, Schankerman, and Van Reenen (2007) use the data to separately identify the presence of both technology and product market spillovers based on the implications of comparative statics. However, my analysis provides a more direct link to the actions of firms and will also enable me to conduct counterfactual policy simulations in the future.

Section 2 of this paper discusses the Chemical Industry and the current R&D subsidy that is offered in the United States. Section 3 describes the data used in the paper. The theoretical model and its sources of identification are described in Section 4. Section 5 presents estimation techniques used in the analysis and the results are then presented in Section 6. The counterfactual policy simulation and conclusions of the paper are described in Section 7.

2 Chemicals Industry.

The Chemicals Industry is a \$474 billion enterprize in the U.S. alone employing over 860,000 workers. It produces over 70,000 different substances and accounts for almost 10% of U.S.

exports². By the nature of its products and production processes, the industry is more closely tied to science than other industries³. The industry is one of the largest investors in research and development, investing an estimated \$24 billion annually and employing 89,000 R&D chemists, engineers, and technicians⁴. Firms in the industry spend a higher percentage of their sales on research than the automobile or telecommunications industries. It was fitting that in 2006 the U.S. Patent and Trademark Office awarded its seven millionth patent to DuPont because the Chemicals Industry generates approximately one in nine U.S. patents.

Chemical companies are dependent on innovations to survive. DuPont is credited for creating the first modern corporate R&D lab and for firms to develop a long-term advantage in the industry they need to continually innovate. Aboody and Lev (2001) examines the return on investment of R&D by chemical companies and finds that the majority of firms' profits result from research. In their analysis the return on physical capital essentially mirrored the industry cost of capital, while R&D was the major contributor to shareholder value. The authors concluded that physical capital was essentially a commodity and that every major producer used its equipment at close to maximum efficiency.

Arora and Gambardella (1999) notes that historically there have been considerable economies of scope involved in chemical R&D. One of the most significant areas of development in the chemical industry has been in polymer science. Throughout the 20th century, as scientists understood more and more about the molecular composition of materials they were able to more efficiently produce products with desirable physical properties. A striking feature of this aspect of the industry is that the same underlying technological base links very diverse product markets. This applies to many of the major breakthroughs in the industry⁵. The introduction of Kaminsky catalysts allowed for more efficient and precise production of a wide range of products. For instance, subtle changes in the composition of nylon can be used to make shopping bags, underwear, sweaters, stockings, tire cords, and fishing line⁶. This motivates the modeling of the industry as homogeneous market because in order for firms to improve products or processes in

²American Chemistry Council (2007)

³Landau and Rosenberg (1991).

⁴American Chemistry Council (2007)

⁵Landau and Rosenberg (1991)

⁶Arora and Gambardella (1999).

one area, they need a fundamental understanding of materials and molecular structures that is applicable to a wide range of markets.

The codifiability of chemical innovations lends itself to the use of patents to protect intellectual property. Levin et al. (1987) conducted an extensive survey of executives in different industries, asking for opinions on the use of patents as means to appropriate gains from innovation. Despite surprising results, which indicated that patents generally were not viewed as an effective means of protection, the responses from the Chemicals Industry ranked patents as the best method of appropriation, though still imperfect. One of the explanations they put forward to explain this is that comparatively clear standards can be applied to assess a chemical patent's validity and to defend against infringement. This protection would be much more clear-cut versus a patent on a complex system, such as electronic device that combined multiple technologies. The importance of patents as a means to appropriate knowledge in the Chemical Industry would suggest that the propensity to patent is high, and more importantly for my analysis, consistent among firms in the industry.

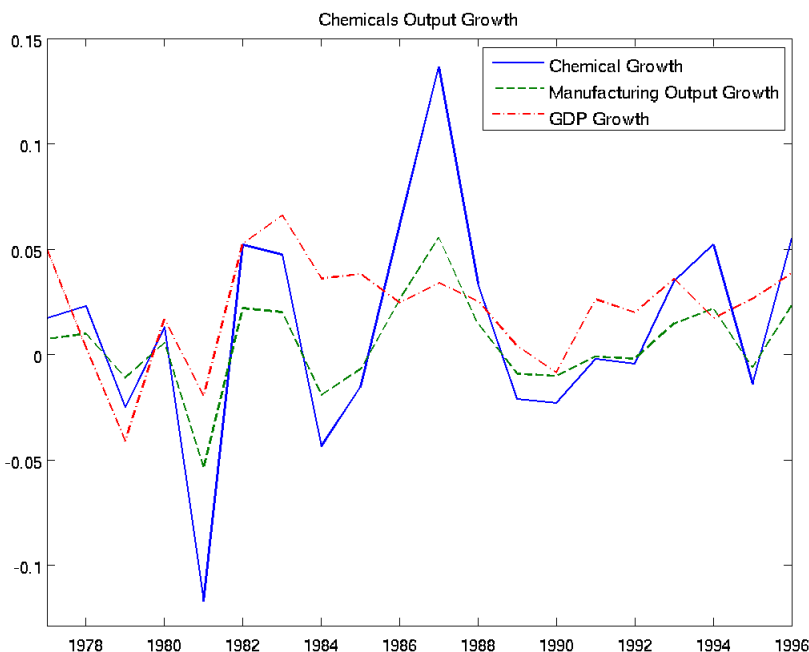
However, as with all industries patents cannot perfectly appropriate a firm's discoveries. Firms' research processes may benefit from the successful innovations of their rivals through personnel turnover, public disclosure, reverse engineering, and being informed as to successful lines of research.

Chemicals are mainly used as intermediate goods by other manufacturing industries. The American Chemistry Council estimates that 96% of manufactured goods are directly touched by chemicals. Therefore, growth in overall economy is strongly influenced by developments in the Chemicals Industry and correspondingly, "The single most important factor in the US Chemical Industry's output and profitability therefore remains the state of the US economy."⁷ Figure (1) displays the relationship between manufacturing growth and chemicals growth.

To summarize, the Chemicals Industry is an important subject to analyze because of its size, influence on the overall economy, and use of the R&E Tax Credit. It is applicable to my model because of the importance of R&D to firms' profitability; the widespread use of patents to protect intellectual property, which provides me with an accurate measure of firm level innovation; the economies of scale and scope in R&D which substantiates the modeling assumption of single

⁷Lenz and Lafrance (1996).

Figure 1: Chemicals Industry Growth Rate.



homogeneous market; and the insignificance of physical capital in generating supranormal profits, which allows me to focus on knowledge levels in the payoff relevant state space.

2.1 US R&D Tax Credit.

In 1981, the US implemented the Research and Experimentation Tax Credit as a part of the Economic Recovery Tax Act of 1981. The goal of the subsidy was to overcome the underinvestment in research by incentivizing additional R&D at the margin that otherwise would not be conducted. The policy provided a 25% tax credit on qualified R&E expenditures above a firm specific base. The base was defined as the greater of either the average R&D of the previous three-years or 50% of the current year. The rate was reduced to 20% in 1986 and the policy has been extended and repeatedly modified through the present. By only providing subsidies for research above the firm specific base, the credit in theory benefitted all firms equally.

Qualified R&E is generally defined as R&D expenditures devoted to the discovery of new scientific knowledge. Research spending qualifies if it is “undertaken for the purpose of discovering

information” that is “technological in nature” and “intended to be useful in the development of a new or improved business component.”⁸ In general, R&E expenditures cover basic and applied research, but not product development. Hall (1993) estimates that 60-75% of R&D expenditures qualify.

One of the limitations of the credit is that it has minimal effect on startup firms that do not generate substantial profits. While firms may carry tax credits forward for up to 15 years, the incentive is effectively discounted if firms cannot immediately use them. This is one of the reasons that approximately 3/4 of the dollar value of credits goes to firms with more than \$250 million in assets.⁹ The firms I analyze all have significant tax liabilities. However, this suggest that the negative impact on the firms smaller than those in my panel is likely to be even more severe than the negative effect on the less advanced firms I analyze. Not only are the smallest firms likely to face stronger competitors, but their effective subsidy rates are close to zero due to their limited tax liabilities.

3 Data.

In this model I focus on firms’ R&D investment decisions. I measure how much innovation results from their research. Then I recover the profits firms generate from their innovations that rationalize their decisions. In order to do this, I use two main data sources. I use publicly available financial data on research expenditures to measure investment and patent data to measure innovations.

I also collect aggregate manufacturing output data from the US Bureau of Labor Statistics. All prices are adjusted to 1980 constant dollars.

3.1 Financial Data.

Research expenditures come from Compustat’s records of publicly available financial data. I use calender year firm research expenditures denominated in \$1980 as my measure of R&D. All of the firms in my universe are large public companies that follow strict financial reporting

⁸U.S. Congress (1995)

⁹NSF (2000).

guidelines. All firms must follow research reporting procedures standardized in the Federal Account Standards Board's Statement 2, 1974, so I do not worry about inconsistencies.

I recover research expenditures related to chemicals business for conglomerates who also operate outside of the industry using their simple patent counts, their total R&D spending, and the proportion of their patents related to their chemicals business.

3.2 Patent Data.

I use the NBER Patent Citations Data File to measure innovations. This data set contains all 3 million issued U.S. patents granted between January 1963 and December 1999 and all 6 million citations made between 1975 and 1999. The data contains information such as Assignee Name and Assignee Number, Application and Issue Date, patent class designation, and the number of citations made and received by 1999.

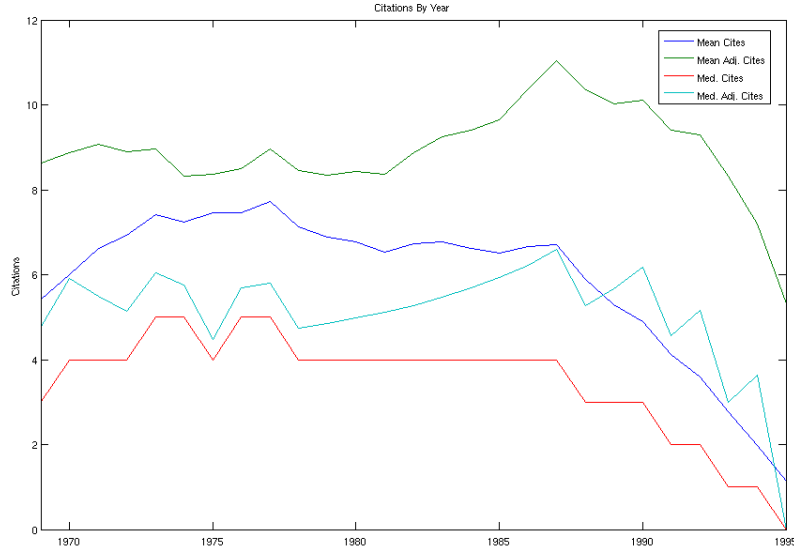
I follow the recent findings in the literature that while patent counts are noisy indicators of innovative value, received citations are much more accurate predictors. When patents are filed, the inventor must disclose any previous technology upon which her innovation builds. Additionally, the patent examiner, who is an expert in the field, reviews the previous patents in the technological area to determine if there are any other citations that are necessary to include. Given that profit-maximizing firms must expend costly R&D to generate new inventions, received citations indicate that resources continue to be invested along the same line of research. This gives some indication of the importance of the original innovation.¹⁰

I use the the citation weighted patents to obtain an aggregate measure of firm innovation in each period. Innovations are defined as $Inn_{it} = \sum_j (1 + C_{ijt+1})$, where C_{ijt} are the adjusted lifetime citations¹¹ received by patent j which was applied for in year $t + 1$, and belongs to firm i . Table (1) shows skewed distribution of citations with the mean value significantly higher than the median.

¹⁰An example of the informational value of citations is that Exxon Chemical won a \$171 Million lawsuit against Mobil for a violation of one of their Metallocene patents in 1998 which had been applied for in 1991. In my data the patent was in the 99.7th percentile of citations with an adjusted citation count of 93.6.

¹¹One of the short-comings of using patent citation data is the fact that patents continue to receive citations over time and I do not observe all of the citations each patent receives over its entire life. I follow the methodology of Hall, Jaffe, and Trajtenberg (2001) in order to account for the truncation bias. Figure (2) displays citations received over time.

Figure 2: Citations Received by Year.



I use a linear weighting structure for citations, assuming that an additional citation has the same innovative value as an additional uncited patent. Papers such as Trajtenberg (1990), have found that there may be increasing returns to the informational value of citations, however the correct weight is not identified in my model so I use a more conservative approach.

I build a one year lag in the innovation process to reflect the delay between when the company conducts research and makes its discovery and then when the patent for the resulting discovery is filed. For example, in year t , I observe firm i conduct R_{it} dollars of research from its financial data. This research leads to innovations, Inn_{it} , which the firm files patents for in the next year. Therefore, my measure of Inn_{it} corresponds to research expenditures R_{it} , even though it actually is a measure of the patents I see being applied for in year $t + 1$. This timing is consistent with the findings of Hausman, Hall, and Griliches (1984) which looked at the effect of lagged R&D expenditures on the production of new applications.

Firms build up their intellectual capital over time by combining new innovations with their previous discoveries. I construct a variable of firm knowledge, $F_{it+1} = \delta F_{it} + Inn_{it}$ which measures the accumulated knowledge of firm i at time t . Without new innovations firm knowledge depreciates due to obsolescence and institutional forgetting. I follow the accepted convention

Table 1: Data Summary.

	Firm Knowledge, F_{it}	Yearly R&D, R_{it}	Yearly Innovations, Inn_{it}	Adj. Citations per Patent
Mean	5,574.6	90.3	949.3	8.890
Minimum	8.3	1.9	0	0
25%	1,200.3	20.8	184.5	1.590
Median	3,496.2	40.9	556.0	5.137
75%	7,286.8	91.3	1,208	11.40
Max	57,983.0	921.7	13,923	607.3

setting $\delta = 0.85$.

Therefore, in each period I use data on research expenditures, R_{it} , innovations, Inn_{it} , and firm knowledge F_{it} for each firm in the industry and Manufacturing Output, Y_t to control for periods of increased demand. Although, I have patent data for 1963 to 1998, I only analyze firms' research decisions from 1976 until 1994. I choose the lower end of the range to enable an accurate estimate of the initial firm knowledge stocks. I used patents as far back as 1969 to build up the initial stocks. I cut off the analysis in 1994 to avoiding introducing too much error into estimation of the lifetime citation figures for patents at the end of the sample. Table (1) shows summary statistics of the financial and patent data.

3.3 Firm Selection.

According to the Department of Energy, there are 170 major chemicals companies operating in the US. However, this industry is dominated by very large firms. In 1985, the top four firms accounted for over 20% of production, and the top 20 firms accounted for 44%. I included 52 companies in my analysis based on the rankings by Chemical & Engineering News of top chemical producers from 1970 to 1990. The firms generated over 60% of sales in the industry in 1985, and no excluded firm generated over 0.2% of sales.

The included firms conduct the vast majority of research in the industry. In 1995, these firms conducted a total of \$6.6 billion in R&D compared to the American Chemical Council's estimate of \$7.1 billion for the entire industry. I therefore assume that I capture the firms which influence competition in intellectual assets.

4 Theoretical Model.

To evaluate the effect of the tax subsidy, it is necessary to have a theoretical model that is representative of the important features of the Chemicals Industry. Demand is largely dependent on the state of the overall economy. Firms are focused on R&D to generate profits. They strategically make decisions which have dynamic implications based on the expected actions of their rivals. Firms also tend to vary in their abilities to conduct successful research.

I model the Chemicals Industry as a homogeneous market with firms engaging in Cournot competition. Firms differ in their observed knowledge levels and unobserved research productivity levels. Innovations benefit firms by lowering their marginal costs and improving their research processes in the future. However, competitors also learn from a firm's innovations, improving their future research processes as well. Future innovations by competitors may also affect the benefit a firm is able to capture from its own breakthroughs.

I assume homogeneous Cournot competition based upon the economies of scope in R&D and largely process based innovations in the Chemicals Industry. However, I check the robustness of this assumption by also estimating an unrestricted model which takes no position on symmetric product differentiation versus process innovations or quantity versus price competition. It simply estimates how firms' profits are translated from own and outside knowledge and the results generally correspond to my results under Cournot competition. The unrestricted specification, however still maintains the assumption that competition and spillovers are symmetric, with no competitors more closely related in products or in technological distance than any others. This is an area I would like to address in the future.

Admittedly, the parameters of the profit function would change under a different specification, but it is important to point out that the estimated pace of innovation, research investment, and firm profits are not dictated by the assumed form of competition, but by the data from the industry.

Patents are used to appropriate innovations, but many studies, such as Levin et al. (1987) have concluded that even in industries with widespread patenting, knowledge is imperfectly concealed. I assume that all firms are equally capable of concealing innovations and treat all outside discoveries equally. I do not distinguish between the extent that knowledge spills over

and the potential usefulness of that knowledge. All innovations produce unintended knowledge flows, but it is then up to other firms to evaluate, assimilate and exploit the outside knowledge to their own benefit. Cohen and Levinthal (1989) describe this as firms’ “absorptive capacity”. I model this by allowing firms to vary in their abilities to utilize extramural information based on their own knowledge level and their research expenditures. Even though knowledge spills over into the industry, firms’ must be able to figure out which ideas are most valuable, and how to use exploit them for their own gain.

Firms make two decisions in each period. They choose a production level, q_i , and a research investment level R_i to maximize their lifetime profits. Production does not affect a firm’s ability to innovate, nor does it affect a firm’s production in the future. The quantity decision affects the firm’s within period profits, but does not have any dynamic implications. Conversely, current research is paid for immediately, but it only affects the firm’s marginal costs in the future through its impact on F_{it+1} . Therefore the level of production is purely a static decision and the research decision is analyzed separately in the dynamic framework.

Assumption 1 Firms follow symmetric, anonymous Markov-Perfect strategies. Therefore, in a symmetric MPNE, each firm uses the same optimal research strategy to maximize lifetime profits given its knowledge level, productivity, the state of the industry, and the expected actions of its competitors.

Assumption 2 Private shocks, ε_{it} , on productivity, research, innovation, rival knowledge, and the economy, are independently and identically distributed.

Combined, these assumptions allow me to consider each firm-year decision as a separate observation of the MPNE strategy, conditional on the observed and unobserved states. Inherent in this analysis is the assumption that the data from every year is generated from a single equilibrium.

4.1 State Space.

In each time period there are N firms competing in the market. Time periods are discrete and continue indefinitely. I do not model entry and exit as only two of the 52 firms exited the industry in my data. I assume these two exits occur exogenously and adjust the other parameters

to account for the reduction in firms.

Each firm has an accumulated knowledge level, $F_{it} \in \mathfrak{R}_+$ at the beginning of the period and also has a private research productivity state, s_{it} , which is a dummy variable that takes on value one when the firm is highly productive and zero otherwise. F_{it} is the depreciated sum of past innovations and is commonly known, while s_{it} is a private parameter that is independent of the values of other firms, but is correlated over time. The firm knows its own research productivity parameter, but does not directly observe the parameters of its competitors.

The firm observes the knowledge levels of each of its competitors, F_{-it} . In an ideal setting this would consist of an $(N - 1) \times t$ matrix capturing the knowledge levels of all of the firm's rivals for each of the previous time periods. However, due to data limitations and the curse of dimensionality in estimating the optimal policy function, this is not feasible. Below I simplify the state space of the firm's competitors through assumptions on the functional form of the cost and demand functions, and on the firm's beliefs of the evolution of rivals' knowledge levels.

In addition, firm's also condition their decisions based on the state of the overall economy, Y_t . Chemicals are mainly used as intermediate goods in the manufacturing sector, and therefore demand generally fluctuates with the business cycle.

4.2 Timing.

- At the beginning of the period, the firm knows its own knowledge level, the knowledge stocks of the rest of the firms in the industry, and the state of the overall economy.
- Each firm receives a productivity shock which determines its own research productivity level, conditional on its unobserved state in the previous period.
- Firms simultaneously choose their levels of output. Firms have no storage capacity and earn profits by selling their production on the spot market.
- Each firm receives an independently drawn research shock and then makes its research decision simultaneously with its rivals.
- Each firm then conducts R&D and innovates based in its research inputs and an innovation shock that is independent across firms and time.

- The firm and industry state vectors then adjust heading into the following period based on the innovation outcomes.

Firms make their production and investment decisions, not knowing the actions of their rivals. In addition, firms never observe the productivity levels of their rivals, however they are able to learn about the likely states by observing research decisions and innovation outcomes.

4.3 Within Period Profits.

Firms generate profits each period by producing and selling chemicals in the product market. Firms have different marginal costs of production based on their individual level of knowledge and engage in Cournot competition. I follow the Lancasterian Model where q is services delivered to the customer, so lower marginal costs, $c(F_{it})$, may result from process innovations that reduce the cost of production or from product innovations that allow the firm to increase the value of the services delivered to the customer¹². Each firm knows the knowledge levels of all of the other firms in the industry and chooses a quantity of production to maximize its profits each period.

Inverse demand is a function of the state of the US economy, Y_t defined as US Manufacturing Output, the total amount of knowledge in the industry, $(F_{it} + \sum_{-i} F_{-it})$, and the total quantity produced in the industry, $\sum_j q_{jt}$. Y_t captures changes the level of demand for chemicals from end users. Demand is dependent on the total knowledge stock of the industry because end users increase their investment in chemical intensive products as the Chemical Industry becomes more advanced. In this respect rival knowledge may be complementary to the knowledge of a given firm.

$$P_t = D(Y_t, (F_{it} + \sum_{-i} F_{-it})) - \alpha \sum_{j=1}^N q_{j,t}$$

Taking the first order conditions of the static profit function, and solving for the equilibrium quantities, it is clear that the optimal quantity $q_{it}^*(Y_t, F_{it}, F_{-it})$, is a function of the state of the economy, Y_t , the firm's knowledge level and its own marginal cost, $c(F_{it})$, and the knowledge of each of the firm's rivals, F_{-it} , through the sum of their knowledge levels, $\sum_{-i} F_{-it}$, and the

¹²Although the specification is robust to process innovations or product innovations in a vertically differentiated market, it cannot be applied to a model with horizontally differentiated goods.

sum of their individual marginal costs, $\sum_{-i} c(F_{-it})$. While quantities are choice variables in this model, they are uniquely determined through the static Cournot equilibrium.

$$q_i^*(Y_t, F_{it}, F_{-it}) = \frac{D(Y_t, F_{it} + \sum_{-i} F_{-it}) - Nc(F_{it}) + \sum_{-i} c(F_{-it})}{\alpha(N+1)}$$

Substituting the optimal quantity equation into the profit equation¹³

$$\pi_{it}^*(Y_t, F_{it}, F_{-it}) = \frac{(D(Y_t, F_{it}, \sum_{-i} F_{-it}) - Nc(F_{it}) + \sum_{-i} c(F_{-it}))^2}{\alpha(N+1)^2} = \alpha q_i^*(Y_t, F_{it}, F_{-it})^2$$

4.4 Welfare Analysis.

Under linear demand in a Cournot model, and the relationship between quantities and profits from the previous equation, total consumer surplus is equal to

$$CS_t = \frac{1}{2}\alpha Q_t^2 = \frac{1}{2}\left(\sum_{i=1}^N \pi(Y_t, F_{it}, F_{-it})^{\frac{1}{2}}\right)^2$$

Total welfare then equals

$$W_t = \sum_{i=1}^N (\pi(Y_t, F_{it}, F_{-it}) - R_{it}) + \frac{1}{2}\left(\sum_{i=1}^N \pi(Y_t, F_{it}, F_{-it})^{\frac{1}{2}}\right)^2 \quad (1)$$

This implies that consumer surplus will not only depend on the cumulative amount of knowledge in the industry, but will also be affected by how knowledge is distributed amongst firms.

4.5 Functional Form of Cost and Demand Equations.

In equilibrium, rivals' knowledge affects a firm's profits through $\sum_{-i} F_{-it}$ and $\sum_{-i} c(F_{-it})$. I specify the functional forms of the demand and cost equations, taking into consideration generally accepted features of the functions and considering implications to the payoff-relevant state space.

¹³In this paper, I consider "profits" to refer to profits from production, excluding research expenditures. However, firms maximize lifetime profits net R&D investment.

I specify the marginal cost function as

$$c(F_{jt}) = K_C - \gamma \log(F_{jt})$$

This functional form is decreasing in firm knowledge, and also experiences diminishing returns¹⁴. It also implies that the sum of the logs of rivals' knowledge levels, $L_{it} = \sum_{-i} \log(F_{jt})$, captures the information in $\sum_{-i} c(F_{jt})$ because $\sum_{-i} c(F_{jt}) = (N-1)K_C - \gamma L_{it}$.

I specify inverse demand as varying linearly in industry-wide knowledge, the state of the economy and aggregate production.

$$P_t = K_D + D_Y Y_t + D_F (F_{it} + \sum_{-i} F_{-it}) - \alpha \sum_{j=1}^N q_{jt} = K_D + D_Y Y_t + D_F F_{it} + D_F \sum_{-i} F_{-it} - \alpha \sum_{j=1}^N q_{jt}$$

Using these specifications the profit function simplifies to

$$\pi_{it}^*(Y_t, F_{it}, F_{-it}) = \frac{(K_D + D_Y Y_t + D_F F_{it} + D_F (N-1)M_{it}) - N(K_C - \gamma \log(F_{it})) + (N-1)K_C - \gamma L_{it}}{\alpha(N+1)^2} \quad (2)$$

where $M_{it} = \frac{1}{N-1} \sum_{-i} F_{-it}$. This specification allows the information in the $(N-1)$ dimension vector of rivals' knowledge, F_{-it} , to enter the profit function solely through the (M_{it}, L_{it}) .

Therefore the within period profit function is a second order polynomial of $(Y_t, F_{it}, M_{it}, \log(F_{it}), L_{it}, K)$, where $K = K_D - K_C$. Coefficients are all functions of $(D_Y, D_F, N, K_D, K_C, \gamma, \alpha)$.

Even with the simplification of the profit function, the serial correlation in the unobserved state space implies that the firm generates its beliefs on the likely productivity levels of its rivals from their entire history of actions and knowledge levels. This information is then used to generate beliefs on the future transitions of its competitors' knowledge and should be considered in the payoff-relevant state space.

However, it would obviously be impossible to replicate this process empirically because I only

¹⁴It is possible that some values of (F, γ, C) may lead to negative marginal costs. However I cannot separately identify K_D and K_C , so therefore I can set K_C to an arbitrarily high level to only permit positive costs, with a correspondingly higher value of K_D .

observe the industry back until 1976. Therefore an additional simplifying assumption must be made as to how firms generate beliefs on the productivity levels and knowledge transitions of their rivals.

Assumption 3. Firms condition their beliefs on the likely actions of their rivals using the two current rival knowledge statistics, (M_{it}, L_{it}) and their values in the previous period.¹⁵

For instance, a large increase in its rivals' mean knowledge in the previous period may be an indication of a subsequent increase in the following period. These transitions also inform the firm as to the specific distribution of its rivals given their mean and the sum of the logs of their knowledge.

Assumption 3 along with the functional forms of the cost and demand equations simplifies the payoff-relevant state space of the firm's rivals from a $(N - 1)x(t)$ matrix of continuous values to a four dimensional vector $I_{it} = (M_{it}, L_{it}, M_{it-1}, L_{it-1}) \in \mathfrak{R}_+^4$.¹⁶

Therefore each firm conditions its output and research decisions on the state, $(F_{it}, I_{it}, Y_{it}, s_{it})$, which represents the observed state, unobserved firm specific research productivity, and the firm's expectations over the values these parameters will take in future. Firms also receive an idiosyncratic research shock which reflects noise in the decision process and does not directly affect profits. Research expenditures are determined by:

$$\sigma_R(F_{it}, I_{it}, Y_{it}, s_{it}) + \varepsilon_{R,it}$$

where $\sigma_R(\cdot)$ is the firm's decision rule.

4.6 Innovation and State Transition Process.

The model consists of four dynamic states: the firm knowledge, rival knowledge, and economy states which are continuous, and the research productivity variable which takes on values $\{0,1\}$.

¹⁵I also tried estimating the optimal policy function including the rival knowledge variables from two years prior, however the coefficients were not significant even at a 68% confidence level.

¹⁶This simplification is in the spirit of Jovanovic and Rosenthal (1988) who analyze sequential games where players base their strategies on summary statistics of their rivals' states. In my case, M_{it} is the mean of their rivals' knowledge and $L_{it} = (\sum_i \log(F_{-it})) = (N - 1)\log(M_{it}) + \sum_{-i} \log(\frac{F_{-it}}{M_{it}})$ when considered with the mean, is a measure of the dispersion. More variance in knowledge levels corresponds to lower values of L_{it} .

Research productivity follows a first-order, time independent, Markov process.

$Pr(s_{it} = 1|s_{it-1}) = (s_{it-1})(\lambda_{1,1}) + (1 - s_{it-1})(\lambda_{0,1})$, where $\lambda_{j,k}$ is the probability of transitioning from productivity j to k . At the start of the period a firm receives a private productivity shock $\varepsilon_{sit} \sim U(0, 1)$, independent of the shocks of other firms and of the shocks on the other parameters, determining its productivity level.

$$s_{it} = \mathbf{1}(\varepsilon_{sit} \leq s_{it-1}\lambda_{11} + (1 - s_{it-1})\lambda_{01})$$

Firm knowledge transitions according to the innovation function. Once the firm chooses its level of R&D expenditures, it conducts research. Innovations are achieved through a stochastic process based on the research inputs. $G(Inn_{it}|R_{it}, F_{it}, I_{it}, s_{it}) = Pr(x \leq Inn_{it}|R_{it}, F_{it}, I_{it}, s_{it})$, where Inn_{it} is the amount of new innovation firm i is able to generate in period t . This specification assumes that outside innovations all benefit the firm equally, independent of the identity of owner. This ignores any effect of technological or geographical proximity on the potential benefits of spillovers. After the firm decides on its R&D, it receives a private innovation shock, $\varepsilon_{Inn,it} \sim U(0, 1)$, reflecting the success of its research.

$$Inn_{it} = G^{-1}(\varepsilon_{Inn,it}|R_{it}, F_{it}, I_{it}, s_{it})$$

The depreciated value of past innovations combine with new innovations to result in the firm knowledge in the following period. $F_{it+1} = \delta F_{it} + Inn_{it}$, where δ is the knowledge depreciation rate.

The firm has rational expectations over the rival state transition. $H(I_{it+1}|I_{it}, F_{it}, R_{it}) = Pr(x \leq I_{it+1}; I_{it}, F_{it})$. While the industry transition depends on the strategies of all of the firm's rivals, the firm assumes that its rivals will always follow the same equilibrium strategies and it uses rational expectations to generate accurate beliefs on the expected future states of the industry. The firm's choice of research expenditure does not affect directly affect the industry transition, however the firm must consider how its investment will affect its knowledge in the forward period and how that will affect future transitions. Each period there is a two-dimensional industry transition shock, $\varepsilon_I \sim N(0, \Omega_I)$ that determines the state of the industry in the future

period.

The state of the overall economy follows an AR(1) process. Growth is dependent on the previous growth rate and exogenous shocks to the economy. It is assumed that actions within the Chemicals Industry do not have a substantive effect on manufacturing growth.

4.7 Equilibrium Concept.

In each period firms choose their research expenditures and quantities of chemicals to produce. Following with other literature on dynamic oligopoly models, I assume firms follow symmetric, anonymous, Markov-perfect strategies,¹⁷ σ_{it} . Strategies are formally functions mapping the state space, $(F_{it}, I_{it}, s_{it}, Y_t)$ and research shocks, ε_R , into choices of (q_{it}, R_{it}) .

Therefore firms' strategies make up a MPNE, where each firm uses the same optimal strategy to maximize lifetime profits given its knowledge level, productivity, the rival knowledge variables, the state of the economy, and the expected actions of its competitors.

Each firm follows a strategy, $\sigma_{q,R}$, to maximize its lifetime discounted profits, which determine its value function,

$$V(FIY s_{i0}|\sigma_{q,R}) = E \sum_{t=1}^{\infty} \beta^{t-1} (\pi(FIY s_{it}, \sigma_{q,R}) - R(FIY s_{it}, \sigma_{q,R})) =$$

$$\pi(FIY s_{i0}, \sigma_{q,R}) - R(FIY s_{i0}, \sigma_{q,R}) + \beta \int V(FIY s_{i1}|\sigma_{q,R}) dP(FIY s_{i1}|FIY s_{i0}, \sigma_{q,R}) \quad (3)$$

where $P(\cdot)$ is the conditional probability of transitioning between states $FIY s_{i0}$ and $FIY s_{i1}$. Note that $V(FIY s_i|\sigma_{q,R})$, is actually the value function conditional on the firm following the strategy $\sigma_{q,R}$, and not necessarily the unconditional value function.

Production quantities, $q(F_{it}, I_{it}, Y_t, s_{it})$, are included in the optimal strategy, but do not affect the efficiency of research, so they do not affect state transition probabilities or future profits. Therefore, optimal quantities are determined each period through a static Cournot equilibrium.

A research strategy, σ_R qualifies as a symmetric Markov-perfect Nash equilibrium, if it is the

¹⁷The unobserved productivity state relaxes the symmetry assumption as it is applied in other papers. Although firms with identical observed and unobserved states are restricted to follow the same strategies, firms that appear observationally identical to the econometrician may pursue different strategies for reasons other than i.i.d. shocks.

optimal strategy for a firm given that all other firms also follow the same strategy:

$$V(F_{it}, I_{it}, Y_t, s_{it} | \sigma_{R_i}^*, \sigma_{R_{-i}}^*) \geq V(F_{it}, I_{it}, Y_t, s_{it} | \sigma_{R_i}', \sigma_{R_{-i}}^*) \quad \forall F_{it}, I_{it}, Y_t, s_{it}, \sigma_{R_i}'$$

This condition allows me to create moments to use in my empirical estimator of the primitives of the profit function.

5 Estimation.

I estimate the primitives of the model using a two-step estimation strategy first proposed by Hotz and Miller (1993) and laid out in a multi-player setting in Bajari, Benkard, and Levin (2007). I also incorporate persistent, but not permanent, firm level unobserved heterogeneity using techniques from Arcidiacono and Miller (2007). The benefit the two-step approach is that it allows the parameters of a dynamic model of imperfect competition to be estimated without the calculation of equilibria. Unobserved heterogeneity is included with relatively little additional complexity by modeling the productivity variable as hidden Markov chain with transitions governed by a first-order Markov process.

The estimation proceeds in two stages.

1. In the first stage the parameters defining the optimal policy function and all of the transitions are recovered.
 - I jointly estimate the parameters of the optimal policy function, the innovation function and the unobserved productivity state which maximize the joint likelihood of the research and innovation observations in the data. I accomplish this using the EM algorithm.
 - I assume the firm has rational expectations over the transitions of the rival knowledge and economy states and these functions are estimated as simple AR1 processes.
2. In the second stage, I use the estimates of the state transition functions to form the basis of a simulated minimum distance estimator to recover the primitives of the profit function.

- I first generate projected state paths for firms starting in different initial states using the optimal policy function and using alternative policy functions based on my first stage estimates. These paths form the basis of the conditional value functions.
- I then build a distance function based on the MPNE moment inequalities where the conditional value functions from following the optimal policy must exceed the profits from following any of the alternative policies.

5.1 First-Stage Estimates.

I use the innovation outcomes and the optimal policy decisions to identify the underlying parameters of the hidden Markov process which determines the unobserved productivity state in the first stage of the estimation.¹⁸

I estimate the parameters that maximize the joint likelihood of the research and innovation observations. This is done by first defining the likelihood function conditional on the unobserved state and computing the probability of each data pair coming from each of the states. Then the unconditional likelihood is calculated integrating out the unobserved state.

Because I observe the outcome of the stochastic innovation process and precise research values that are generated from distributional strategies, I can learn about the likely unobserved productivity state of the firm. Firms who generate more innovation than I otherwise would expect given their observed state, are more likely to have high productivity. Unobserved research productivity increases a firm's ability to innovate, implying that the optimal research expenditures should also vary with the unobserved state. Therefore, I can also infer that firms who invest at different levels than I otherwise would expect them, have greater likelihoods of being in the high productivity state.¹⁹

¹⁸ It is only possible to estimate the parameters of the unobserved state in the first stage because productivity does not directly affect profits. Both current and future profits are independent of productivity, conditional on the observed states of the firm. If this condition did not hold, it would be necessary to completely solve the second stage of the estimation procedure for each iteration of the EM algorithm. Given, the second stage is by far the more computationally intensive part of the process, having to repeatedly calculate this stage would make the combination of the BBL and AM methods infeasible.

¹⁹Ex ante it is unclear whether high productivity firms necessarily should invest more or less than low productivity firms. Without knowing the innovation parameters and the profit parameters it could be that the "wealth effects" of the additional innovations they are able to generate would outweigh the "substitution effects" due to more efficient processes causing high productivity firms to invest less money in R&D. However, by estimating the parameters of the policy function, the innovation function and the unobserved state together, I am able to work

Additionally, I allow for persistence in the unobserved state, ($\lambda_{jj} \neq \lambda_{jk}$), so information on the likely productivity state of a firm in one year, contributes to the likelihood of the firm's state in preceding or following years. Therefore, when integrating out the unobserved state in the likelihood function, the likelihood of an observation is dependent on all of the observations for a given firm. This makes directly maximizing the likelihood function prohibitively difficult.

Arcidiacono and Jones (2003) show that by estimating parameters of the model using the Expectation-Maximization (EM) algorithm, the separability of observations is restored.

Joint Likelihood Function.

I estimate the parameters ($\theta_{Inn}, \theta_R, \lambda$), that jointly maximize the likelihood of

$$\max_{\theta_{Inn}, \theta_R, \lambda} \sum_{i=1}^N \sum_{t=1}^T L(Inn_{it}, R_{it} | F_{it}, I_{it}, Y_t; \theta_{Inn}, \theta_R, \lambda) \quad (4)$$

where $L(\cdot)$ is the likelihood of observing the research and innovation variables for firm i in period t , conditional on the observed states, given the estimated the parameters of the policy, innovation, and productivity transition functions.

The joint likelihood of observing firm i investing R_{it} in research and generating Inn_{it} conditional on the state ($F_{it}, I_{it}, Y_t, s_{it}$) given the parameters (θ_{Inn}, θ_R) is defined as $\mathcal{L}(Inn_{it}, R_{it} | F_{it}, I_{it}, Y_t, s_{it}; \theta_{Inn}, \theta_R)$. The unconditional likelihood is then computed by integrating the conditional likelihoods over the probabilities of firms being in each unobserved state given the data and ($\theta_{Inn}, \theta_R, \lambda$). Abbreviating $\mathcal{L}(Inn_{it}, R_{it} | F_{it}, I_{it}, Y_t, s_{it} = s; \theta_{Inn}, \theta_R, \lambda)$ by \mathcal{L}_{sit} , the likelihood of observation it is:

$$L(Inn_{it}, R_{it} | F_{it}, I_{it}, Y_t; \theta_{Inn}, \theta_R, \lambda) = q_{it} \log(\mathcal{L}_{1it}) + (1 - q_{it}) \log(\mathcal{L}_{0it}) \quad (5)$$

where $q_{it} = E[s_{it} = 1 | Inn_i, R_i, F_i, I_i, Y] = q_t(Inn_i, R_i, F_i, I_i, Y; \theta_{Inn}, \theta_R, \lambda)$, is the probability that firm i is in the high state at time t , conditional on the data for firm i from all periods, ie $Inn_i = (Inn_{i1}, Inn_{i2}, \dots, Inn_{iT})$.

This is where persistence in the unobserved state complicates calculations. The data from other periods, though not from other firms, is informative to the likely state in a given period.

around this issue.

$q_t(Inn_i, R_i, F_i, I_i, Y; \theta_{Inn}, \theta_R, \lambda)$ is computed using Bayes' Rule, based on, L_{st} , the joint probability of firm i being in the high state in period t and observing firm i 's choice and innovation paths (Inn_i, R_i) , conditional on the state variables in each period, and the parameter values $\theta_{Inn}, \theta_R, \lambda$. L_{st} is defined by:

$$L_{st}(Inn_i, R_i | F_i, I_i, Y; \theta_{Inn}, \theta_R, \lambda) =$$

$$\sum_{s_1=0}^1 \cdots \sum_{s_{t-1}=0}^1 \sum_{s_{t+1}=0}^1 \cdots \sum_{s_T=0}^1 \left(\prod_{r=2, r \neq tr \neq t+1}^T \lambda_{s_{r-1}, s_r} \mathcal{L}_{s_r, ir} \right) (\lambda_{s_1} \mathcal{L}_{s_1, i1} \lambda_{s_{t-1}, s} \mathcal{L}_{s, it} \lambda_{s_{t+1}} \mathcal{L}_{s_{t+1}, it+1}) \quad (6)$$

Summing over both the high and low states at time t leads to the likelihood of observing the firm's R&D choices and innovations, (Inn_i, R_i) . Therefore the probability of firm i being in the high state in period t , given the parameters and conditional on all of the data for firm i is:

$$q_t(Inn_i, R_i, F_i, I_i, Y; \theta_{Inn}, \theta_R, \lambda) \equiv \frac{L_{1t}}{L_{0t} + L_{1t}} \quad (7)$$

The estimates of the state probabilities are used to update the productivity transition parameters, λ .

Conditional Likelihood Functions.

With the ability to calculate the probabilities of unobserved states along with the transition parameters, the only terms needed to estimate Equation (5) are conditional likelihoods of research and innovation conditional on the unobserved state. Using the definition of conditional probabilities, \mathcal{L}_{sit} can be decomposed into the product $\mathcal{L}_{Inn, sit} \mathcal{L}_{R, sit}$, where $\mathcal{L}_{Inn, sit}$ is the likelihood of generating Inn_{it} , conditional on the research decision, R_{it} and $\mathcal{L}_{R, sit}$ is the likelihood of the observed research decision. Both likelihoods are conditional on the firm being the productivity state s .

The expected log likelihood is therefore

$$\sum_{i=1}^N \sum_{s=0}^1 \sum_{t=1}^T q_{sit} [\log(\mathcal{L}_{Inn}(Inn_{it} | R_{it}, FI_{it}, Y_t, s_{it} = s; \theta_{Inn})) + \log(\mathcal{L}_R(R_{it} | FI_{it}, Y_t, s_{it} = s; \theta_R))] \quad (8)$$

I assume innovations follow a Gamma Distribution. I use a continuous distribution even

though patents and citations take on discrete values, because when I adjust them to account for truncation bias, they become continuous variables. The gamma distribution is applicable because it provides weight to strictly positive values and it is skewed permitting greater mass for smaller returns while still allowing for significant discoveries to occur. In addition total innovations can be interpreted as the sum of individual research projects that each follow the exponential distribution, with mean equal to the scale parameter of the gamma distribution. Following from the idea of free disposal, I restrict the estimated parameters so that expected innovations are non-decreasing in both firm knowledge and research expenditures. The conditional likelihood of the innovation function is therefore equal to the density function of the gamma distribution.

$$\mathcal{L}_{Inn}(Inn_{it}|R_{it}, F_{it}, I_{it}, Y_t, s_{it} = s; \theta_{Inn}) = g(Inn_{it}|R_{it}, F_{it}, I_{it}, Y_t, s_{it} = s; \theta_{Inn}).$$

Because I model research as a continuous choice and the knowledge as a continuous state, I am unable to nonparametrically obtain estimates of the optimal policy function that converge in probability. It would also not be possible to use a nonparametric estimator in conjunction with the EM algorithm to refine my estimates of the unobserved states. I therefore assume I can approximate the optimal policy function with a normal distribution where the explanatory variables are a functions of the state space, (F, I, Y, s) .

$\mathcal{L}_R(R_{it}|F_{it}, I_{it}, Y_t, s_{it} = s; \theta_R) = \varphi(\log(R_{it})|F_{it}, I_{it}, Y_t, s_{it} = s; \theta_R)$, where $\varphi(\cdot)$ is the density function of the normal distribution.

EM Algorithm.

After generating initial guesses for the conditional likelihoods I estimate parameter values of the innovation and policy functions, by iterating between maximizing the overall likelihood function (8), given the previous values of q_{it} , and then updating the values of the unobserved state probabilities using equation (7). I continue to iterate between the three steps until I achieve convergence in all three parameters $(\hat{\theta}_{Inn}, \hat{\theta}_R, \hat{\lambda})$. Once I have recovered the parameters of the hidden Markov chain that represents the productivity state, I can treat the state as if it is observed and simulate it in the second stage.

Rival Knowledge and Economy Transition.

I estimate both the rival knowledge and economy transitions using least squares regressions. I estimate the two rival knowledge variables transitions separately and then use the residuals to

determine the covariance parameters of Ω_I . The rival knowledge transitions are defined by:

$$I_{it+1} = I_{it} (\Delta_I(F_{it}, I_{it}, Y_t, \hat{\theta}_I) + \varepsilon_{I,t}) \quad (9)$$

where $\varepsilon_{I,t} \sim N(0, \hat{\Omega}_I)$.

The economy state is assumed to evolve independently of the Chemicals Industry. Therefore its transitions are defined by

$$Y_{t+1} = Y_t (\Delta_Y(Y_t, \hat{\theta}_Y) + \varepsilon_{Y,t}) \quad (10)$$

where $\varepsilon_{Y,t} \sim N(0, \hat{\sigma}_Y^2)$.

5.2 Second-Stage Estimates.

The basis of this estimation procedure is that

$$V(F, I, Y, s | \sigma_R^*, \theta_\Pi) \geq V(F, I, Y, s | \sigma_R', \theta_\Pi) \forall F, I, Y, s \text{ \& } \sigma_R'. \quad (11)$$

Using the first-stage estimates, I simulate the equilibrium conditional value functions under strategy σ_R , for a firm starting in an initial state, $(FIY s_0)$.

$$V(FIY s_0 | \sigma_R; \theta_\Pi) = \mathbf{E} \left(\sum_{t=0}^{\infty} \beta^t \pi(FIY s_t), -R(\sigma_R, FIY s_t, \varepsilon_{R,t}), ; \theta_\Pi \right) | FIY s_0; \theta_\Pi,$$

where the expectation is over the values of $\varepsilon_t = (\varepsilon_{R,t}, \varepsilon_{Inn,t}, \varepsilon_{s,t}, \varepsilon_{I,t}, \varepsilon_{Y,t})$ for all t .

Under the MPNE, the value function conditional on following the optimal policy should exceed the value function from any alternative policy. This forms the basis of moment inequalities which I use to create a distance function in order to estimate the primitives of the profit function.

Starting States $(FIY s_0)$.

This model contains continuous state variables, $F, I, Y \in \mathbf{R}_+^6$ and a discrete state $s \in \{0, 1\}$.

Equilibrium strategies must be defined and optimal for all possible states. However, I only simulate value functions over a finite set of starting points. I set up a vector of starting states $FIY_{s_0} = (F, I, Y, s)$ where FIY_{s_0} is $\#Starts = 250$ randomly drawn states based on the observations in the data.

Alternative Research Decisions σ'_R .

I specify alternative research decision rules in order to construct the inequalities. I build value functions based on these alternative decision rules beginning with the same initial states and experiencing the same private and industry shocks as under σ_R^* . The alternative decision rules are proportional adjustments of σ_R^* . I draw, $\#Alts = 25$, values of $\nu_{alt} i.i.d. \sim N(1, .3) > 0$. An alternative strategy $\sigma_R(\nu_{alt}, FIY_{s_t}, \varepsilon_{R,t}) = \nu_{alt} \sigma_R^*(FIY_{s_t}, \varepsilon_{R,t})$.

Simulations.

To obtain the value function under different strategies, I solve for the firm's expected lifetime discounted profits, where the expectation is take over research, innovation, productivity, and industry shocks.

I forward simulate the projected state paths using $\#Shk = 2500$ simulated error paths of length $\#Yrs = 75$, for each starting state and alternative policy function, σ_R . The value function is then the sum of the discounted profits.

In each period, I determine the selected levels of R&D given the state of each firm, the policy function and the errors. Then based on $\hat{\theta}_{Inn}$ the R&D interacts with the states and ε_{Inn} to produce innovation leading to firm knowledge in the next period. The state along with ε_I describes the transition of rival knowledge and the economy evolves according to its estimated process.

The value function for a given starting state and policy function is calculated as the mean over all shocks of the sum of discounted profits from years 0 to T .

$$V(FIY_{s_0}, \sigma_R; \theta_{II}) = \mathbf{W}(FIY_{s_0}, \sigma_R) \theta_{II}, \quad (12)$$

where $\mathbf{W}(FIY_{s_0}, \sigma_R)$ is the functional form of the profit function from equation (2) computed for each path that began at FIY_{s_0} and used policy function σ_R .

Typically dynamic oligopoly models normalize their parameters by using observed prices in an

initial stage. I do not do this and as a result within period profits, $\pi(FIY_{s_t}) - R(\sigma_R, FIY_{s_t}, \varepsilon_{R,t})$ are essentially flow utilities based on knowledge levels. Therefore in comparing value functions based on optimal and alternative policy functions, any positive monotonic transformation of a utility function yields the same ordinal ranking of alternatives. I normalize the scale of the profit function by setting the coefficient on R&D expenditures to -1. Because I observe actual dollar values for research investment, this fixes the units of the coefficients on knowledge state variables in the profit function to 1980 \$millions.

Estimation of the Profit Function.

The projected paths for the $\#Starts$ starting states, $\#Alts$ different alternative strategies, and $\#Shocks$ error shocks form ($\#Starts \times \#Alts$) inequalities, after taking expectations over the error paths. These are used to generate the distance function,

$$Distance(\theta_{\Pi}) =$$

$$\sum_{(FIY_{s_0}, Alts)} \mathbf{1}(V(FIY_{s_0}, \sigma_{R'}; \theta_{\Pi}) \geq V(FIY_{s_0}, \sigma_{R^*}; \theta_{\Pi})) (V(FIY_{s_0}, \sigma_{R'}; \theta_{\Pi}) - V(FIY_{s_0}, \sigma_{R^*}; \theta_{\Pi}))^2 \quad (13)$$

I then estimate $\hat{\theta}_{\Pi} = \arg \min Distance(\theta_{\Pi})$. Although the distance function contains the discontinuous indicator function, the form of the penalties lead it to have well-defined analytic gradients. At points of discontinuity, the right-side of the equation is equal to zero. At any points where the right-side is strictly greater than zero, the indicator is continuous. Therefore I can use efficient optimization routines to solve for θ_{Π} .

It is important to note that α , the slope of the inverse demand function, is not separately identified from the other parameters of the profit function because it only appears in the denominator of Equation (2). Therefore, I only identify the coefficients on the terms of the polynomial incorporating restrictions from Equation (2) based on the parameters (D_Y, D_F, N, K, γ) . Even though I estimate 18 coefficients, the restrictions limit me to four degrees of freedom. Consumer surplus and social welfare are calculated using Equation (1) and the profit function estimates.

I also estimate the coefficients of the profit function without any of the restrictions in order to test the robustness of the specification.

6 Estimation Results.

The model is comprised of a number of different functions which I estimate. They involve first stage estimates of the policy and transition functions, estimates of the profit function, and then finally estimates from the counterfactual simulations. I first present the results of each stage individually and discuss how they relate to the theoretical innovation literature. Then I close the section by summarizing how the estimates of the different equations fit together to describe how research and development competition occurs in the Chemicals Industry and implications for the industry structure.

6.1 First Stage Policy and Transition Estimates.

The optimal policy function, the innovation function, and productivity state transition parameters are all intrinsically linked.

Table 2: Productivity Transition Parameters.

λ^1	0.2997
λ_{11}	0.7908
λ_{01}	0.4357

Table (2) displays the estimated productivity transition parameters. Parameter λ_{11} clearly demonstrates that there is persistence in the unobserved state parameter. This result helps explain why there is significant knowledge dispersion both at the end of the sample, and it explains how the Chemicals Industry could have reached the level of dispersion at the beginning of the sample. The heterogeneity of firms not only occurs because of idiosyncratic shocks as proposed in the Ericson and Pakes model, but because once firms become highly productive, they tend to stay that way.

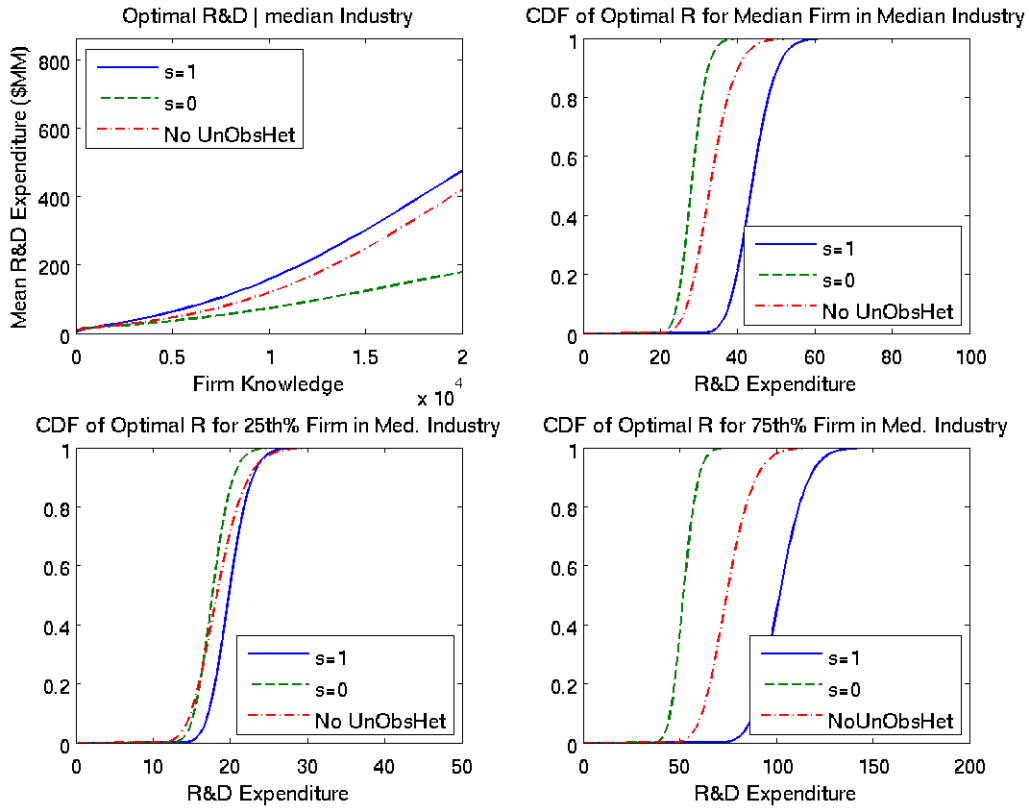
Table(3) displays selected parameters of the optimal policy function. I use a flexible functional form of the firm and rival knowledge variables, so the coefficients on individual parameters are difficult to interpret. Highly productive firms invest more in research than other firms. The margin between the two types of firms is increasing in firm knowledge and decreasing in rivals' knowledge. Figures (3A-D) also displays these results.

Figure (3A) shows that research expenditures are increasing in the knowledge of the firm.

Table 3: Optimal Policy Function Estimates.

	Full Specification	SE**	No Unobserved Het.	SE**
Subsidy Effects				
$1(\text{SubsidyYears}) * xlog(F)$	0.0772	0.0050	0.0434	0.0083
$1(\text{SubsidyYears}) * xlog(M)$	-0.0703	0.0057	-0.0365	0.0075
High Productivity Adjustments				
$1(S = 1)xlog(F)$	0.3068	0.0072	-	-
$1(S = 1)xlog(M)$	-0.2416	0.0069	-	-
R^2	0.8980		0.8841	
Durbin-Watson	0.3398		0.1404	
ρ_t	0.7779		0.9259	
*Since 1981 all firms have received federal tax subsidies on based on research expenditures.				
**Standard errors ignore estimation error of unobserved state transitions.				

Figure 3: Optimal Policy Function.



Firms with a cost advantage invest more in R&D to maintain and increase their position. Without examining the other parts of the model it is unclear whether this is derived from greater research efficiency in innovation process or through nondecreasing returns in the product market. It is also important to note that research declines with knowledge for highly advanced firms. This ensures the stationarity model allowing for the use of my estimator.

Table 4: Innovation Function Estimates.

	Full Specification	SE*	No Unobserved Het.	SE*
Spillover Effects				
M x log(R) / 5000	-1.08E-08	1.53E-05	-1.69E-08	4.34E-06
M x log(F) / 5000	0.4732	0.0171	0.5200	0.0720
High Productivity Effects				
1(S=1)	0.4729	0.0405	-	-
1(S=1) x log(R)	-0.5992	0.0630	-	-
1(S=1) x log(F)	0.2013	0.0225	-	-
1(S=1) x F x log(R) / 5000	-0.7655	0.5952	-	-
1(S=1) x R x log(F) / 1000	0.1078	0.0108	-	-
Dispersion	106.35	2.83	115.45	4.56
R^2	0.8397		0.8208	
Durbin-Watson	0.5322		0.4008	
ρ_t	0.5721		0.7902	
$\rho_{R,Inn}$	-0.0126		0.0332	

Table (4) displays selected parameters of the innovation function. The first key point from this function is that innovations are increasing with rival knowledge indicating that spillovers are present and significant in the research process. Firms are unable to fully appropriate their discoveries, which likely leads to underinvestment in R&D from a social welfare standpoint. It is also interesting to note that the coefficient on the interaction between rival knowledge and research expenditures is not significant. It is not enough for firms to spend money on R&D to be able to access outside knowledge. However, as also shown in Figures (4 C&D) the coefficient on rival knowledge interacted with firm knowledge is significant. This coincides with the findings of Cohen and Levinthal (1989) that firms need their own expertise to be able to evaluate and assimilate outside innovations. This would seem to suggest that there will be increasing dominance in the industry, as large firms can utilize innovations from smaller firms, but small firms have a harder time using the innovations of the larger firms.

As expected, innovation is increasing with firm knowledge and research expenditures. Re-

Figure 4: Innovation Function.

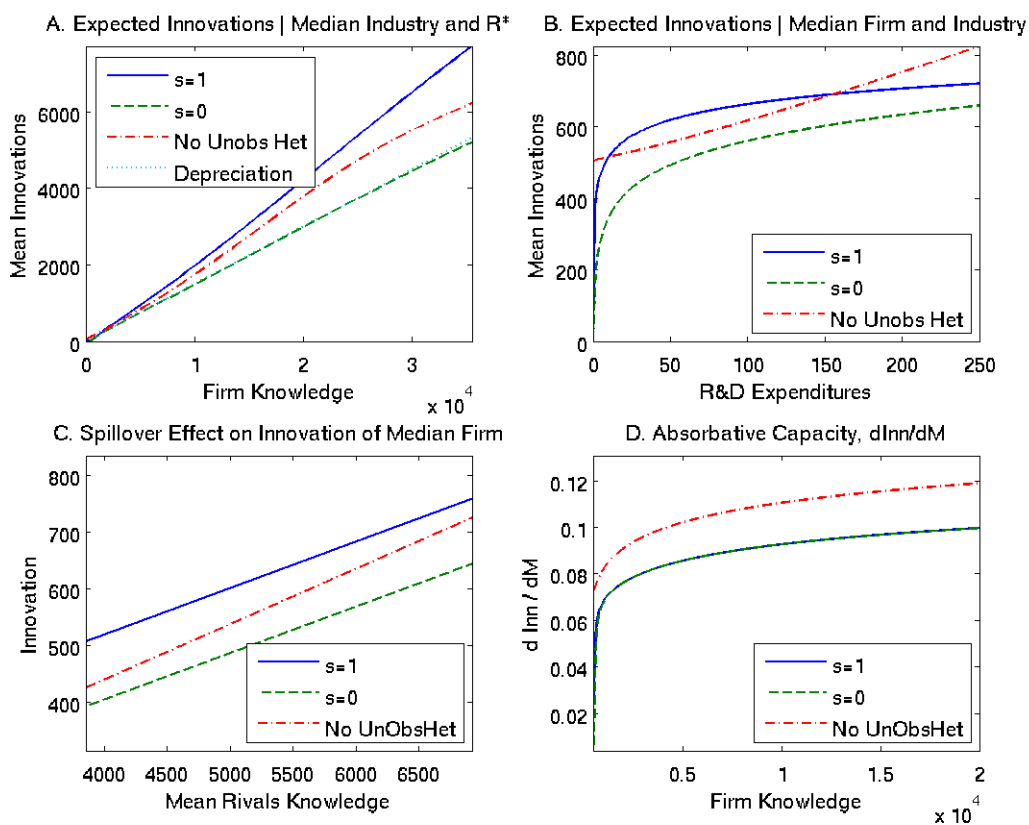
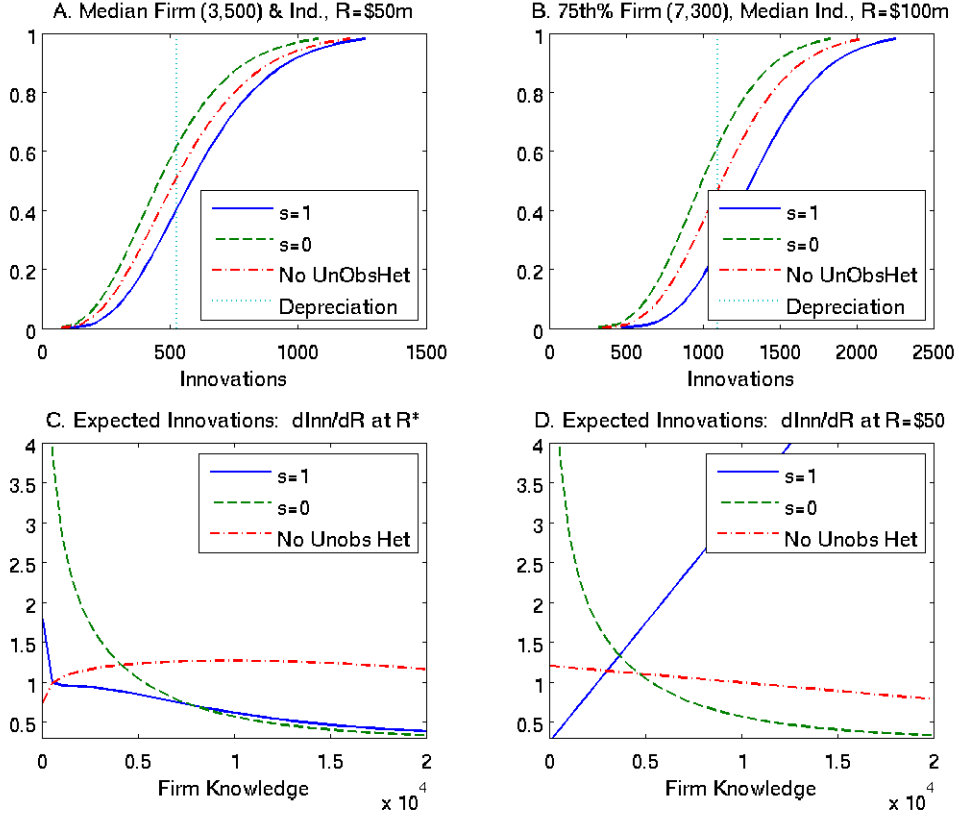


Figure 5: Innovation Function.



search displays significant decreasing returns. This can be seen in Figure (4B) where there is almost no benefit to additional research above \$60M for either type of firm. This further motivates the need for firms to consider current and future actions of their competitors in their current R&D decisions because it is prohibitively costly to make large adjustments to firm knowledge in any given period.

The main difference between innovations in the high productivity state versus the low productivity state is the complementarity of knowledge and research for highly productive firms. Figure(5D) demonstrates this stark contrast. The return to research is increasing in knowledge for high state firms while decreasing for low state firms. These results coincides with the findings in the optimal policy function that differences in R&D investment between firms in the two states is increasing in the knowledge level of the firms.

The differences in research efficiency combined with expenditure patterns lead to low pro-

ductivity firms not being able to generate enough innovations to counteract depreciation, while highly productive firms tend to increase their level of knowledge over time. This is demonstrated in Figures (4A,5A&D), where the expected innovations for firms in the low state are less than the depreciation level for most knowledge levels.

Figure (5C) shows that at the margin, the last dollar of research is more productive for smaller firms rather than larger firms, in both the high and low states. Smaller firms may underinvest to a greater degree than larger firms due to the heterogeneity in spillover effects. This shows that smaller firms would be able to generate greater additional innovations from an unexpected government R&D grant. However, this does not necessarily imply that smaller firms would benefit more than larger firms from anticipated research grants or subsidies. The same factors that lead smaller firms to invest less than larger firms would likely cause the same distribution of expenditures under symmetric policies.

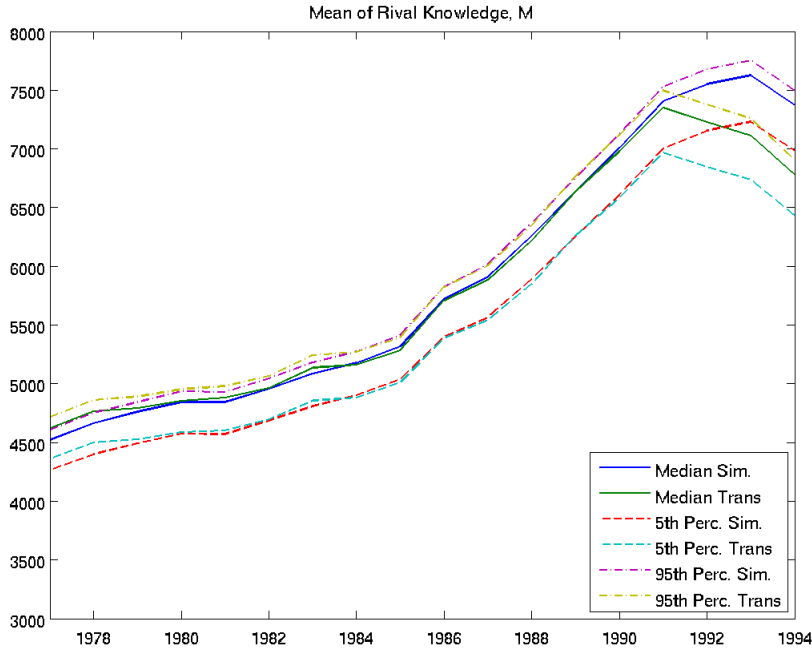
I show that the estimated transition function of rival knowledge is internally consistent with the optimal policy and innovation functions in Figure (6). The figure displays the yearly transitions of the mean rival knowledge variable using the estimated transition function (Trans.) and simulated with the optimal policy and innovation functions (Sim.). Both the median transitions and the confidence interval around likely transitions demonstrate the consistency of the functions.

Effect of Unobserved Heterogeneity.

The most important impact of including the unobserved productivity state relates to the return to R&D. Figure(4B) displays that without incorporating the unobserved state, the effect of research on innovations does not decline. This result is contrary to the estimated innovation function with the unobserved state and contradicts assumptions used in other innovation papers. Without the unobserved state, estimates of return to research are likely to be biased because of the simultaneity between innovations and firms' decisions. Failing to account for this, it would be difficult to explain why firms do not invest more in R&D, and it would also permit large adjustments in firm knowledge in given periods.

The productivity state also significantly reduces the dispersion parameter in the innovation function as shown in Table (4). Including the unobserved state lowers the variance in innovation outcomes by 8.6%.

Figure 6: Internal Consistency of Rival Knowledge Transitions.



6.2 Second Stage Profit Function Estimates.

Table (5) displays the parameter estimates of the profit function for two different specifications. The first column displays the results when the coefficients on the profit function polynomial are restricted by the terms, (D_F, D_Y, γ, K) , from the theoretical model presented in the previous section. The second specification derives its general functional form as a second order polynomial of the different knowledge variables from the theoretical model in equation (6). However, it does not enforce the equality conditions on the coefficients. It is simply a reduced form profit function translating own and outside knowledge into profits that rationalize observed research expenditures.

The estimated individual parameters of the model are difficult to interpret because of the flexible form of the profit function. Many of the terms of the two specifications are opposite in sign. However, looking at the differences between profits of firms in different states, and the analytic derivatives of the two functions in Table (6), it does not appear that the Cournot competition assumption is driving the results. For the three derivatives where the signs are

Table 5: Profit Function Estimates.

Parameters	Restricted Function	Unrestricted
Firm Knowledge, F		
F	0.0355	0.9885
log F	251.79	1446.3
$(\log F)^2$	0.064	-50.786
F(log F) / 1000	0.0090	-78.457
$F^2 / 1000^2$	0.0013	1.5707
Mean Rival Knowledge, M		
M	1.738	88.217
$M^2 / 1000^2$	3.03	212.04
Sum of Log of Rival Knowledge, L		
L	-5.04	-5089.6
L^2	2.54E-05	-1.0889
Interactions		
F x M / 1000 ²	0.062	-8.341
F x L / 1000	-0.0002	-0.2552
F x Y	0.0159	-0.0028
log(F) x M	0.0004	0.1699
log(F) x L	-0.001	-9.825
log(F) x Y	112.99	2578.7
M x L / 1000	-0.0088	2.2534
M x Y	0.780	-90.708
L x Y	-2.260	5934.9
R	-1	-1
<i>Distance</i>	1.63E+07	5.36E+06

Table 6: Profit Function Estimates.

	Restricted Fn.	10% to 90% C.I.		Unrestricted	10% to 90% C.I.	
$\Pi(\text{MedianFirm}) - \Pi(25\text{th}\%Firm)$	517.7	146.1	2231.2	801.7	68.4	2015.5
$\Pi(75\text{th}\%Firm) - \Pi(\text{MedianFirm})$	470.5	154.4	1269.7	673.3	38.5	1708.9
$\frac{\delta Profit}{\delta F} \text{MedianFirm}$	0.158	0.048	0.591	0.246	0.011	0.523
$\frac{\delta Profit}{\delta F} 75\text{th}\%Firm$	0.103	0.030	0.222	0.129	-0.007	0.282
$\frac{\delta Profit}{\delta F} 25\text{th}\%Firm$	0.362	0.095	1.598	0.551	0.040	2.164
$\frac{\delta Profit}{\delta M}$	2.509	0.807	7.097	-1.131	-6.338	3.862
$\frac{\delta Profit}{\delta L}$	-7.410	-16.911	-1.723	289.28	-2256.1	1208.6
$\frac{\delta Profit}{\delta Y/100}$	42.561	-638.28	203.49	27729	-38969	128645
$\frac{\delta^2 Profit}{\delta F^2}$	-3.03E-05	-1.83E-04	-7.76E-06	-5.25E-05	-3.38E-04	7.70E-06
$\frac{\delta^2 Profit}{\delta F \delta M}$	-1.82E-07	-4.60E-05	1.30E-06	-2.44E-05	-3.03E-04	5.04E-05
$\frac{\delta^2 Profit}{\delta F \delta L}$	-5.42E-07	-1.29E-04	-1.43E-07	-3.46E-03	-1.51E-02	1.43E-02

different for the two specifications, the estimates in the unrestricted model are insignificant.

The derivatives with respect to rival knowledge demonstrate that firm and rival knowledge are substitutes in the product market. Although rival knowledge increases industry demand and firm profits, the marginal return to firm knowledge is decreasing in rival knowledge for both specifications. This suggests that horizontal differentiation is not the main focus of innovations. Innovations for rivals is not relaxing product market competition, but causing creative destruction as the benefits a firm can generate from its own innovation, fall with innovations by rivals.

Given the observed research intensity in the industry, this result coincides with the theoretical works of Symeonidis (2003) and Qiu (1997). They found that R&D expenditures were greater under quantity competition than under price competition in product and process innovations, respectively.

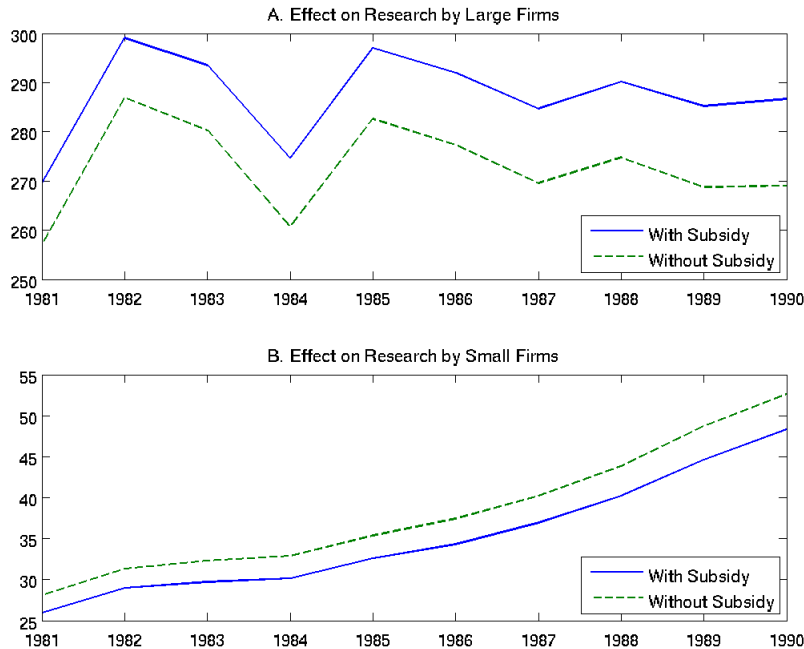
Profits are also decreasing with respect to the sum of the logs of rivals' knowledge in both specifications. This also coincides with the typical results from the Cournot model. Given an aggregate amount of knowledge in the industry, a firm's profits increase with greater size dispersion. As you move away from a situation where all firms are the same size, larger firms compete less intensely in the product market as they take profits on inframarginal goods.

The median firm is able to generate approximately \$158,000 in profits from an additional unit of knowledge. When conducting the optimal level of research, this implies the firm generates \$115,500 in the product market for each additional million dollars in R&D. Given that knowledge depreciates at 15% and firms discount profits at 5%, a marginal dollar of research is expected to contribute \$0.67 directly in discounted profits. It also increases the research efficiency of the firm in the future periods, allowing firms to fully recover research costs. In comparison, a marginal dollar of research generates \$1.21 and \$0.34 for the 25th percentile and 75th percentile firms, respectively. This is another key motivation of using a dynamic structural model to analyze competition in the industry. It would be difficult to construct a static model to rationalize why firms of different sizes have varying marginal returns. The heterogeneous dynamic returns to innovation, mainly through differences in absorptive capacity, explains this result.

6.3 Conclusions.

1. Externalities are present in the innovation process, motivating the modeling competition as a dynamic game and suggesting that firms underinvest in R&D from a social welfare standpoint.
2. Absorptive capacity is strongly influenced by firm knowledge, but is independent of current research expenditures. This, along with increasing returns in the profit function motivates expenditures to be increasing in firm knowledge. Smaller firms are less able to utilize outside innovations and have a more difficult time protecting their own advances, reducing their incentives for R&D.
3. The main difference between high and low productivity firms is the complementarity between firm knowledge and research for high productivity firms. This causes differences in optimal expenditures for firms in the two productivity states to be increasing in firm knowledge.
4. Firm and rival knowledge are substitutes in the product market. Although rival knowledge increases industry demand and firm profits, the marginal return to firm knowledge is decreasing in rival knowledge suggesting that Cournot competition in homogeneous goods is the correct specification.
5. A marginal dollar of R&D generates different marginal profits for firms of different sizes. However, discounted marginal profits directly attributable to the marginal research do not alone justify the expense. The dynamic return component, contributing to the firm's ability to innovate in the future makes up this difference.
6. Differences in firms level of expenditures, abilities to innovate at the margin, and their abilities to generate profits all suggest that further work is necessary in analyzing appropriate subsidy programs to increase industry R&D.

Figure 7: Tax Policy Simulation.



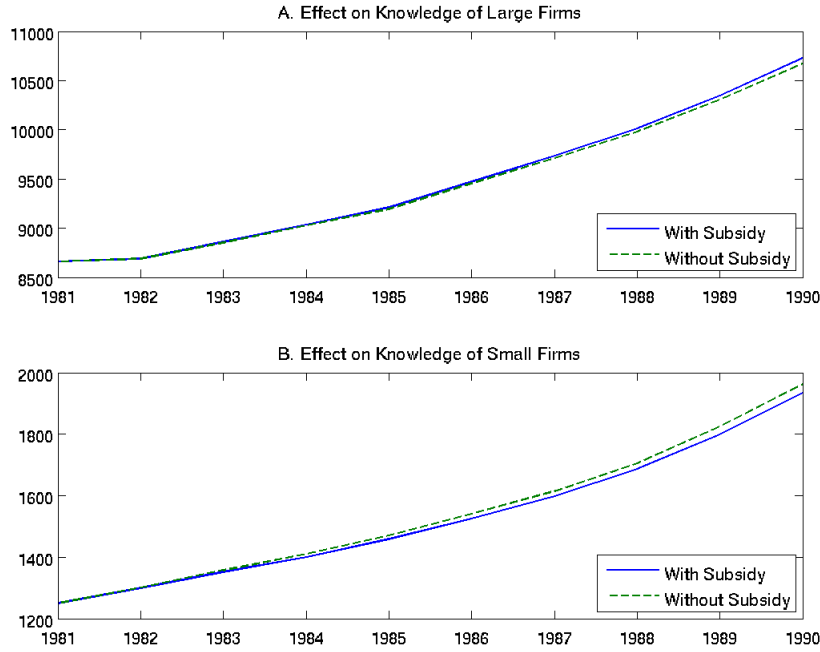
7 R&E Tax Credit Analysis.

I use the implementation of the R&E Tax Credit of 1981 as a natural experiment. Given the parameters of optimal policy function, both before and after the implementation of the subsidy, I simulate how the Chemicals Industry would have involved with and without the subsidy from 1981 to 1990. I consider how the non-symmetric reactions of firms affected the dynamics of the industry and analyze whether or not the subsidy enhanced social welfare.

Figures (7,8) shows the differences in response by large and small firms. I define large firms as firms above the median level of knowledge in 1981. Conducting the simulation, these are the firms which increase their R&D expenditures enough to take advantage of the subsidy. On average, smaller firms are only able to receive credits in 3 out of 10 years, while larger firms receive credits in over 6 years.

On average, the subsidy generates \$170.4M in additional R&D each year. This represents an increase of 2.13%. To accomplish this, the government forfeits \$246.2M in revenue. This implies that a dollar of tax credit produces a \$0.69 increase in R&D. This is lower than the estimates

Figure 8: Tax Policy Simulation.

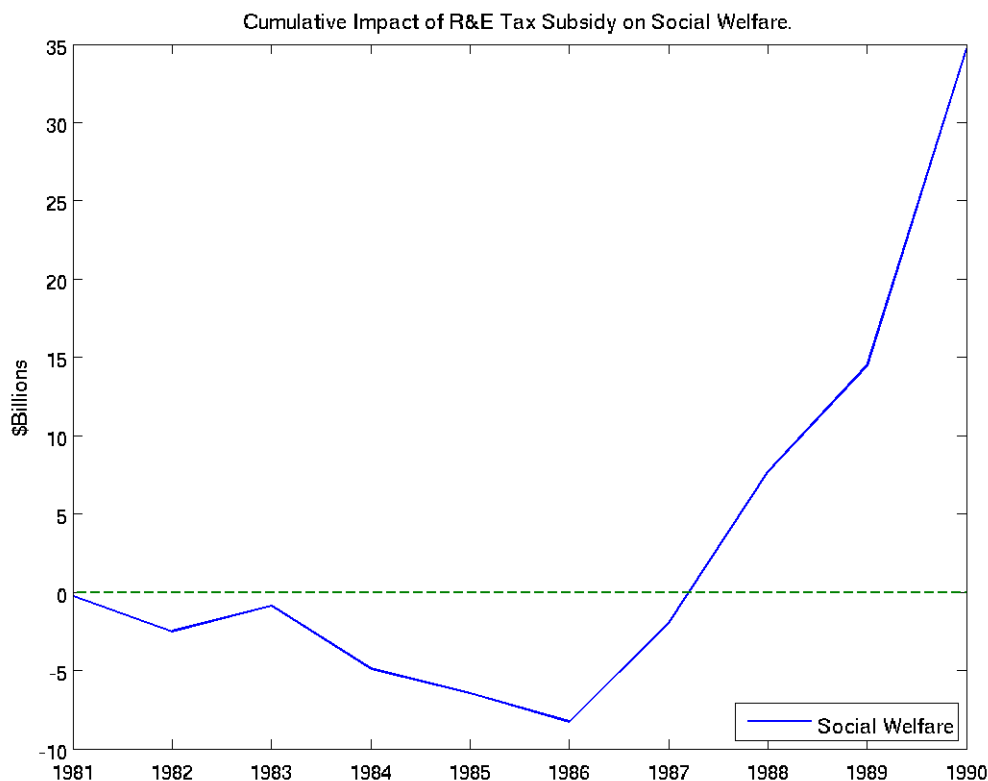


obtained by Hall (1993) and Berger (1993) which obtained values close to 1. However, when I only examine the actions of large firms, I obtain an estimate of \$1.06, suggesting the other methods are ignoring the negative impact of the subsidy on investment by the followers in the industry.

The aggregate amount of knowledge in the industry increases by 0.121% with the subsidy. However, the mean marginal cost actually increases versus how the industry would have involved due to the redistribution of investment and innovation. This mitigates much of increase in consumer surplus, as advancements are concentrated on the leaders of the industry who are less concerned about gaining market share, and more focused on generating greater profits from inframarginal units. Consumer surplus rises on average by 0.655% a year²⁰, compared to an increase in industry profits of \$581M or 1.97% a year. The cumulative effects of the subsidy raised consumer surplus 4.15% in 1990, however, if smaller firms had not decreased their investment

²⁰When I use the unrestricted profit function estimates, there is actually a decrease in consumer surplus with the tax subsidy caused by more severe diminishing returns to knowledge. However, the consumer surplus calculation is dependent on the Cournot assumption, making this figure difficult to interpret.

Figure 9: Tax Policy Simulation.



over the decade, surplus would have increased 10.6%.

The effect of the tax policy on social welfare over time is shown in Figure(9). In the initial years after the policy is implemented social welfare declines because the increased R&D spending takes time to generate enough profits to offset its cost. In addition, as the market becomes more concentrated consumer surplus declines. It is not until 1988 that cumulative effects become positive as large firms become more efficient in both their research and in the chemical production process.

While the combined affects of profits and consumer surplus indicate that the subsidy is social welfare enhancing after eight years. It also emphasizes the how the competitive effects of the policy concentrate the benefits on the industry leaders.

The simulations reflect the fact that differences in absorptive capacity create heterogenous responses from firms. Of the two sources of externalities, smaller firms' research decision are more

affected by creative destruction than knowledge spillovers in the innovation process. However, as firms become more advanced, they not only are able to capitalize on their own developments, but they better able to assimilate successful ideas from their rivals. Therefore, under the subsidy, the impact of new information on their research processes dominates the creative destruction from rival innovations that reduce the marginal benefits of their discoveries.

7.1 Conclusion and Extensions.

This paper demonstrates the importance in considering dynamic competitive effects in analyzing R&D subsidies. I find that increased expenditures and innovations from large firms are largely offset by smaller firms who's marginal benefit from innovating is decreased by the actions of their rivals.

While the subsidy does increase social welfare, the gains are made mainly by large firms at the expense of smaller firms. Although the government planned to implement a policy which seemingly benefitted all firms equally due to firm-specific base levels, competitive effects cause large firms to prosper to a greater degree. This type of asymmetric impact should be considered when developing future policies.

I only analyze the policy's effect on the Chemicals Industry to focus on the interactions between rival firms, however, I believe these same effects should be present in other industries. Given the availability of the data I use my analysis, the methods presented here could be used to measure the competitive impacts of the subsidy in other areas.

My methods depend on the assumption of homogenous good Cournot competition in calculating welfare effects. The estimates of the unrestricted profit function suggest that this is not an invalid simplification. Even in the unrestricted case, the marginal benefit a firm generates from a discovery is decreasing in the technological state of its rivals. This indicates that it is more appropriate to model process innovations or vertical differentiation than horizontally differentiated goods. However, it is likely that true research competition in the Chemicals Industry is a combination of both product differentiation and process improvements. I would like to be able to relax the assumption that each firm is impacted symmetrically by each of its rivals in both the product and technological settings. With more detailed data I would like to address

these issues by including measures of the product mix and types of innovations for each firm.

Finding that the effectiveness of the tax credit is reduced by the decreased in R&D by smaller firms suggests that alternative policies may be more efficient in meeting the government's goals. One of the benefits of the structural model is that it allows for the testing of counterfactual policies. Using the parameters of the innovation and profit functions, I can recover equilibrium investment strategies under different subsidy regimes. This paper suggests that it would be useful to test the welfare implications of progressive subsidies that are targeted to smaller firms. While there may be political issues that would make it problematic to enact "asymmetric" subsidies, it would provide a basis to understand whether or not the difference in performance would justify the political difficulties.

References

- ABOODY, D., AND B. LEV (2001): “R&D Productivity in the Chemical Industry,” *Working Paper*. New York University, NY.
- AGUIRREGABIRIA, V., AND P. MIRA (2007): “Sequential Estimation of Dynamic Discrete Games,” *Econometrica*, 75(1), 1–54.
- AMERICAN CHEMISTRY COUNCIL (2007): “The Business of Chemistry: Essential to Our Quality of Life and the US Economy,” [http://www.americanchemistry.com/sacc/bin.asp?CID = 288DID = 731DOC = FILE.PDF](http://www.americanchemistry.com/sacc/bin.asp?CID=288DID=731DOC=FILE.PDF).
- ARCIDIACONO, P., AND J. JONES (2003): “Notes and Comments Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm,” *Econometrica*, 71(3), 933–946.
- ARCIDIACONO, P., AND R. MILLER (2007): “CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity,” *Working Paper*. Duke University.
- ARORA, A., AND A. GAMBARDELLA (1999): “Chemicals,” *U.S. Industry in 2000: Studies in Competitive Performance*, pp. 45–74.
- ARROW, K. (1962): “Economic Welfare and the Allocation of Resources for Invention,” *The Rate and Direction of Inventive Activity*, 13, 609–625.
- BAJARI, P., C. BENKARD, AND J. LEVIN (2007): “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75(5), 1331–1370.
- BERGER, P. (1993): “Explicit and Implicit Tax Effects of the R & D Tax Credit,” *Journal of Accounting Research*, 31(2), 131–171.
- BLOOM, N., M. SCHANKERMAN, AND J. VAN REENEN (2007): “Identifying Technology Spillovers and Product Market Rivalry,” *NBER Working Paper No. 13060*.
- COHEN, W., AND D. LEVINTHAL (1989): “Innovation and Learning: The Two Faces of R & D,” *The Economic Journal*, 99(397), 569–596.

- EISNER, R., S. ALBERT, AND M. SULLIVAN (1984): “The New Incremental Tax Credit for R&D: Incentive or Disincentive,” *National Tax Journal*, 37(2), 171–183.
- ERICSON, R., AND A. PAKES (1995): “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *The Review of Economic Studies*, 62(1), 53–82.
- HALL, B. (1993): “R&D Tax Policy During the Eighties: Success or Failure?,” *Tax Policy and the Economy*, 7, 1–36.
- (1995): *The Effectiveness of Research and Experimentation Tax Credits: Critical Literature Review and Research Design*. Report to Office of Technology Assessment. Congress of the U.S.
- HALL, B., A. JAFFE, AND M. TRAJTENBERG (2001): “The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools,” .
- HALL, B., A. JAFFE, AND M. TRAJTENBERG (2005): “Market Value and Patent Citations,” *Rand Journal of Economics*, 36(1), 16–38.
- HARHOFF ET AL. (1999): “Citation Frequency and the Value of Patented Inventions,” *The Review of Economics and Statistics*, 81(3), 511–515.
- HAUSMAN, J., B. HALL, AND Z. GRILICHES (1984): “Econometric Models for Count Data with an Application to the Patents-R & D Relationship,” *Econometrica*, 52(4), 909–938.
- HOTZ, V., AND R. MILLER (1993): “Conditional Choice Probabilities and the Estimation of Dynamic Models,” *The Review of Economic Studies*, 60(3), 497–529.
- JOVANOVIĆ, B., AND R. ROSENTHAL (1988): “Anonymous sequential games,” *Journal of Mathematical Economics*, 17.
- LANDAU, R., AND N. ROSENBERG (1991): “Innovation in the Chemical Processing Industries,” *Technology and Economics: Papers Commemorating Ralph Landau’s Service to the National Academy of Engineering*.
- LANJOUW, J., AND M. SCHANKERMAN (2004): “Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators*,” *The Economic Journal*, 114(495), 441–465.

- LENZ, A., AND J. LAFRANCE (1996): “Meeting the Challenge: US Industry Faces the 21st Century: the Chemical Industry,” .
- LEVIN ET AL. (1987): “Appropriating the Returns from Industrial Research and Development,” *Brookings Papers on Economic Activity*, 1987(3), 783–831.
- NATIONAL SCIENCE FOUNDATION (2000): “Science and Engineering Indicators,” <http://www.nsf.gov/statistics/seind00/>.
- PESENDORFER, M., AND P. SCHMIDT-DENGLER (2007): “Asymptotic Least Squares Estimators for Dynamic Games,” *Review of Economic Studies*.
- QIU, L. (1997): “On the Dynamic Efficiency of Bertrand and Cournot Equilibria,” *Journal of Economic Theory*, 75, 213–229.
- RUST, J. (1987): “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 55(5), 999–1033.
- SCHMOOKLER, J. (1966): “Innovation and Economic Growth,” *Cambridge, Mass.*
- SYMEONIDIS, G. (2003): “Comparing Cournot and Bertrand equilibria in a differentiated duopoly with product R&D,” *International Journal of Industrial Organization*, 21(1), 39–55.
- TRAJTENBERG, M. (1990): “A Penny for Your Quotes: Patent Citations and the Value of Innovations,” *The RAND Journal of Economics*, 21(1), 172–187.
- U.S. CONGRESS (1995): “Tax Policy: Additional Information on the Research Tax Credit,” GAO/T-GGD-95-161.