Three Essays in Applied Microeconomics

by

Diego Rodriguez

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1993

Author .....
Department of Economics
May 28, 1993

Certified by .....
Oliver D. Hart
Professor of Economics
Thesis Supervisor

Accepted by .....
Richard E. Eckaus
Chairman, Departmental Committee on Graduate Students

Three Essays in Applied Microeconomics

by

Diego Rodriguez

Submitted to the Department of Economics
on May 28, 1993, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis dissertation comprises three papers in two different areas of Applied Microeconomics: Industrial Organization and Applied Econometrics.

The first essay *"The Growth and Diffusion of Knowledge and the Theory of the Firm"*, considers a model of endogenous knowledge spillovers. Innovation often takes the form of knowledge that cannot be protected by patents. An entrepreneur with access to a research opportunity needs a research team to exploit this opportunity. The members of the team become competitors for the sale information, that is the output of research. In this trade-off between efficiency in research and appropriability of innovation, the entrepreneur offers an optimal contract to the researcher so as to affect the researcher's incentives to become a competitor. Contractible variables that affect those incentives are the size of the team, profit sharing agreements, ownership of physical assets (and their strategic investment) and the allocation of control rights. The organization of the firm is viewed as a preemption game between the entrepreneur and the employees.

The second essay *Decentralization and the Management of Competition* (joint with Dimitrios Vayanos), develops a theory of delegation of authority in large corporations. In particular, the paper studies the costs and the benefits of decentralizing decision-making authority to competing divisions of a firm. We argue that when a division is given the right to market its product to customers of the other division, its manager has an incentive to increase product quality. At the same time divisions locate their products too close together. We find that top management may or may not want to restrict competition among divisions and show that even if competition is desirable, it may not be credible due to top management's temptation to intervene "ex-post".

The third and last essay, *Semiparametric Measurement of Environmental Effects"* (joint with Thomas M. Stoker), discusses the application of semiparametric methods to the characterization of the effect of pollution on housing prices. Semiparametric methods are likely to be superior to parametric and nonparametric alternatives, in hedonic price applications, because they combine flexibility in functional form of

moments with simple interpretability of results. The ideas are applied to the Boston Housing Data of Harrison and Rubinfeld, using the semiparametric framework in Rodriguez and Stoker [13]. It appears that previous findings from this data are modified when the partial index model framework is used. The partial index structure allows us to identify the features in the data responsible for misspecification of the previous (parametric) studies that used this data base.

Thesis Supervisor: Oliver D. Hart
Title: Professor of Economics

A mis padres, Angela y Carlos.

Contents

List of Figures

List of Tables

Acknowledgements

A large number of people have helped to make life easier and research smoother in these years at Cambridge. There are too many to be remembered here, but I am thankful to all of them.

M.I.T. is notorious for the intelligence and initiative of its students. I was fortunate to get to know and become friends with many of them. I would like to thank very specially to Dimitri Vayanos, for many enlighting conversations, wise advice and friendly support. Walter Novaes, Fernando Branco, Luigi Zingales, Sara Fisher, Glen Ellison, Rabi Abraham, Jeff Zwiebel, Andres Almazan were excellent in providing intellectual stimulation.

I will keep a specially good memory of some of my teachers at M.I.T. In particular, Bob Solow, Peter Diamond, and Whitney Newey. But it is with my advisors with whom I contracted an important debt in the last years. Oliver Hart, Michael Piore and Tom Stoker have been a great group of advisors. I have learnt very important lessons from all of them and they have helped me to understand what is the essence of good research. I am grateful to Oliver for emphasizing the need to keep highest standards in relevance and rigour at the same time. I am specially grateful to him for his encouragement in a moment of difficulty. To Michael for inducing me to explore contact with practitioners to understand the issues and to develop alternative ideas to conceptualize phenomena. From a more personal point of view, Michael has been most helpful in the last years. Tom has been an always cheerful source of encouragement and ideas. It is to a large extent thanks to him that I didn't fall more often in the main problem of graduate student life: to lose perspective. I got uncountable insights

on Econometrics from conversations with Tom.

I would like to thank very specially Jean Tirole for his generous advice. Before I talked to Jean about my research, I was fortunate to profit from his excellent teaching (in particular the topics course in I/O offered in the fall of 1990). Each conversation with Jean has helped me to improve my research.

My last year at M.I.T. has been made much more enjoyable than the previous ones thanks to the love and friendship of Swetlana Gaffron. Her spontaneous charm and direct style have been dearest companions. It will be a challenge to honor my debt with her in the coming years.

Finally, I would like to acknowledge the constant and patient support from my parents during all these years. It is their stimulating influence what induced love of knowledge in me in the first place. It is the liveliness they put in my education what started my curiosity to travel and to come to America to study. It is to them to whom this thesis is dedicated.

Introduction

This doctoral dissertation is composed of three essays. The first two ("Diffusion of Knowledge and the Theory of the Firm" and "Decentralization and the Management of Competition" -joint with Dimitrios Vayanos) essays in Industrial Organization. The third one ("Semiparametric Measurement of Environmental Effects." -joint with Thomas M. Stoker) is an study in Applied Econometrics.

The essay "Diffusion of Knowledge and the Theory of the Firm" studies the problem of how agents capitalize on innovation that cannot be protected through patents. It is generally understood that there does not exist a market for innovations in organizational form of firms. Similarly, it is commonly observed in industries that information some of the necessary for the realization of profit opportunities cannot be specified in contracts[1]. In particular this implies the impossibility of establishing property rights on that information.

The paper studies the optimal contract offered by an agent (entrepreneur) that introduces this kind of innovation. The entrepreneur faces a trade-off in the number of agents that he will introduce in order to develop the innovation. The reason is that, since the innovation cannot be patented, the entrepreneur cannot *prevent* the trade of the innovation by any agent that is informed about it, to a third party (nor can he commit not to sale the innovation to third parties). This has implications on the size of the team that will develop the innovation.

[1]An example of this is the expertise on *general* knowledge about the industry that improves superior prediction of industry parameters.

It is clear that making agents claimants of the profit stream derived from the innovation is necessary to prevent its diffusion (and the profit dilution that follows from it). Under liquidity constraints of agents there are costs of selling rights to profit streams. It follows that the non-patentability of innovation imposes a restriction on the size of the team. In addition, we show the role of physical assets as a substitute mechanism to protect knowledge diffusion. The strategic investment of physical assets affects the payoffs of informed agents that want to collude with third parties that are not part of the original contract with the entrepreneur. It therefore might improve appropriability. In order to exploit the benefits derived from the strategic use of physical assets it is necessary that they have become already sunk and linked to the profit stream at the time when the contract between entrepreneur and employees. This has implications on the ex ante allocation of property rights of physical assets.

The second essay ("Decentralization and the Management of Competition") explores the nature of authority delegation in large firms. The paper gives a central role to the endogenous determination of information flows inside firms. As opposed to the team theory literature, that considers the implications of costly transmission of information, in this paper those costs don't exist here. Information transmission, for us, will affect agent's incentives. The paper belongs to the current of literature that has focused on Incomplete Contracting (that is associated with the names of Oliver Williamson, Harold Demsetz and Oliver Hart, among others).

Setting aside the question of asset ownership, we characterize the organization of the firm by the degree of information-sharing between a principal and an agent (i.e. the firm is more *centralized* if more information is being shared). Information-sharing affects the incentives of the agent because it affects the "appropriability" conditions at the time profits are split. In particular, we find that a principal might choose to gather *less* information in order to commit to a distributive policy ex post.

We develop this question in a particular framework where the firm is formed by two product divisions and a third party that, for exogenous reasons, retains all

property rights over physical assets (named the general office). The managers of divisions engage in two "moral hazard" activities (that are not observable to the general office): product quality and "design". Design is a horizontal differentiation parameter that determines the degree to which the two products are substitutes for consumers. In this framework, there are two main questions we study: the first one ignores the problem of commitment by the general office to particular ex post policies of distribution; it concerns rather with the degree of competition that the general office will introduce in the relationship between the two divisions. There is trade-off in introducing strong competition among divisions: competition increases the incentives to perform high product quality, but it also induces division managers to engage in rent seeking activities that are in their own interest, but inefficient from the perspective of the firm as a whole. The paper therefore derives the optimal degree of competition between divisions. Competition takes the form of bargaining between divisions to deal with customers. We assume efficient bargaining and the role of competition is to determine the status quo of the bargaining game.

In particular, regulation of competition by the general office is determined by the extent to which the general office *intervenes* in the allocation process of customers to divisions. In the two extreme cases, the general office mandates the allocation of customers to divisions or lets the divisions bargain for customers without constraints. This brings us to the second question of the paper: firstable, intervention by the general office requires that the latter is *informed* about customer characteristics. In addition, the general office is always tempted to intervene ex post and regulate customer allocation (to avoid potential competition among division that would dissipate profits). This brings to our claim: the general office might prefer not to have access to some pieces of information about divisional affairs in order to commit not to intervene ex post in the competition among divisions.

The third essay (joint with Tom Stoker) applies semiparametric techniques developed in Rodriguez Stoker (1992) to the problem of measuring effects of environmental policy.

Chapter 1

The Growth and Diffusion of Knowledge and the Theory of the Firm

Abstract

Innovation often takes the form of knowledge that cannot be protected by patents. An entrepreneur with access to a research opportunity needs a research team to exploit this opportunity. The members of the team become competitors for the sale information, that is the output of research. In this trade-off between efficiency in research and appropriability of innovation, the entrepreneur offers an optimal contract to the researcher so as to affect the researcher's incentives to become a competitor. Contractible variables that affect those incentives are the size of the team, profit sharing agreements, ownership of physical assets (and their strategic investment) and the allocation of control rights. The organization of the firm is viewed as a preemption game between the entrepreneur and the employees.

1.1 Introduction

It is well known that the imperfect ability to protect product and process innovation with patents and trade secret gives a public good characteristic to the knowledge necessary to generate those innovations. Empirical and theoretical results in the Industrial Organization literature have systematically found that this violation of the property rights of the innovator lead to inefficiencies in the allocation of productive factors towards research and development (Spence (1984), Griliches (1991)).

The theoretical literature that has focused on the so called problem of technological spillovers has in general studied the case of expropriation through *imitation in the marketplace.* This means that in these models there is typically a firm that introduces a new product in the market and after a time lag there is an imitator that is able to "backward engineer" the new technology without infringement of an existent patent. The presence of this imitator and potential entrant is taken as an exogenous factor.

Without questioning the relevance of these results for the understanding of the agent's incentives to engage in research and development, we will focus on a different source of technological spillovers that has not been dealt with in the Industrial Organization literature. Namely, we will be interested in the role that members of a research team play in the diffusion of the firm's innovation to firm's outsiders. In particular, we will study the factors that affect the economic incentives of researchers to diffuse information and the scope for the optimal control of these factors.

The point of departure of this paper is extensive field research on an american industry characterized by drastic ongoing technological innovation. Namely, the set of firms engaged in the commercial application of genetic engineering techniques. (For a description of the R&D process in Biotechnology see (Pisano and Mang (1992), Office of Technology Assessment, (1984 and 1991) and Rodriguez (1992)).

With that body of evidence we build a stylized description of the research process in high technology sectors that is used to lay out the assumptions of our formal

framework. The paper is intended as a contribution to the theory of the firm, and more particularly to the theories of organization of industrial research.

In a large fraction of industries, activity directed to the discovery of new products is not undertaken by an isolated agent. Rather, technological innovation is the outcome of a team of researchers that coordinate their activity. In sectors where the research is closely related to basic science, the research team is mainly engaged in the production of new knowledge that will eventually lead to the discovery and design of viable products. At this very early stages of product creation, the ability to enforce contracts that restrict the communication of the information relevant for product discovery is very limited.

The problem of diffusion of knowledge outside the firm that generated often takes the form of *diffusion of learning rents.* Consider the following case: an agent derives a new process (called X) to produce diagnostic kits for a form of cancer. The new process X, reduces the cost of the kits to half of the cost of the available best practice. This innovator forsees that there are related processes, that derive from the same research program, that will lead to other new products (diagnostic kits for other forms of cancer, eventually a drug with therapeutic effects). Call Y to these second innovations.

The agent is in general able to patent the new process X. In order to satrt its commercialization, he will hire a number of agents (that will study the reliability of the new method, will manufacture it...). This further exploration of the new process X leads to *learning by doing* of the second generation of innovations, Y. But it was unfeasible to patent Y before hiring human resources, since Y was at that time only a possibility, rather than a tangible contribution. Once human resources are brought to commercially exploit X, all agents working on project X can also forsee product Y. Since at that stage property rights on Y cannot be stablished, Y becomes common "property" (common knowledge) to all agents that know X. In particular, the knowledge about Y becomes possibly the object of trade by all agents that work

on product X.

A central claim of this paper is that the above description of the research process is the rule rather than the exception in high-technology sectors.

There are two main ways in which members of a research team can induce the diffusion of innovation generated by the team. The first one is through direct communication of some of the innovation-generating information to which they have access as team members. The second is through the job turnover of the members of the research team. This means that team members leave the firm and develop projects in a different firm that they start or in an existing firm. These projects could not have been undertaken if the seceding researcher had not been a member of the team that he left, where he had access to knowledge and where he learnt skills that he could not have acquired otherwise.

The problem of the optimal design of contracts for the maximization of the profits of an entrepreneur, under either of these alternative mechanisms of diffusion will in general be very similar. Indeed, if it is the case that the decision to leave a firm cannot be stipulated by a contract, then both of these scenaria are formally identical. [1]

A number of recent contributions have provided systematic case studies from industries that show the agent's inability to protect from expropriation the knowledge derived by research. These include, among others, (Rumelt, 1984), (Teece, 1986), (Levin et al. 1987), (Mishina, 1991) and (Pisano and Mang, 1992).

Mang and Pisano (1992) describe the research process in its early stages in Biotechnology as follows:

[1] Such an assumtion would appear to be too extreme. In this paper we will focus on the case where information affecting learning rents can be spread without costs and without leaving the firm. We argue that the analysis of the organization of research, under the assumption that diffusion requires employees' turnover would be formally very similar, and that our results would extend to that case as well.

The product development cycle in biopharmaceuticals can be viewed as a sequence of investments in information about both the product and the process. By product information, we mean information about the characteristics of and application of the product.

Various types of product information is generated throughout the development cycle. For example early research can be thought of as search for information about the molecule, its structure and its likely behavior. Additional research generates information about which versions of the molecule are likely to have desirable therapeutic effects. Pre-clinical development is designed to generate more information about how the molecule behaves in animals and what its toxicity and efficacy are likely to be in humans. [...]

To deal with the inability to patent this collection of information that leads to new products and processes, employees in Biotechnology (and high technology sectors in general) are required to sign non disclosure agreements as they become members of the firm. Yet these contracts are apparently insufficient by themselves to prevent the knowledge from diffusing outside the firm. The following quotation was taken from the corporate report of firm in Biotechnology. Similar statements can be found in the reports of most Biotechnology companies:

Patents and Proprietary Rights

[...] Much of the Company's Know-how and Technology is not patentable. To protect its rights, the Company requires all employees, consultants, advisors and collaborators to enter into confidentiality agreements with [the Company] . There can be no assurance, however, that these agreements will provide meaningful protection for the company's trade secrets, know-how, or other proprietary information in the event of any unauthorized use or disclosure.

To study the issue of the organization of research under incontractibility of knowledge in a formal framework, we will consider the case of an agent (the entrepreneur) that has a research project in an early stage. For instance, the entrepreneur is an expert in a field of science and the only one who knows that research in a particular biological compound might eventually lead to the development of a treatment for some kind of cancer. We call this information available only to the entrepreneur at an initial stage the *learning capital.*

The exploration of the particular compound can be undertaken by the entrepreneur alone or by a research team. The second implies higher efficiency in research (like the earlier discovery of the medication). If the entrepreneur brings other agents (researchers) to the project, these will necessarily have access to the information that made the project possible (the learning capital). The team members will then have the choice of transfering this information to firm outsiders, or to preserve secrecy so that the information remains available only to firm members. This decision by the members of the team will in general affect the degree of competition for the profits arising from the result of research that the team faces.

All this implies that there is an important difference between the two kinds of spillovers described above. In the case where diffusion stems from backward engineering of products by an imitator, the existence of the imitator has to be taken as given. The incumbent will be able to affect the entry decision of the potential entrant by engaging in the strategic investment of assets that affect profits of the entrant (Spence-Dixit, 1984), (Fudemberg-Tirole, 1984).

In the case where diffusion derives from the communication or turnover of employees, the existence of imitators should be considered as an endogenous variable that depends on the design of the research team. Namely, it will depend on those variables that affect the incentives of agents to communicate the learning capital to firm outsiders. These are the size of the team, the profit sharing agreement among team members and, similarly to the case of imitation in the market, on the strate-

gic investment in assets. Therefore we will interpret the problem of design of the R&D entrepreneurial firm as a preemption game between the entrepreneur and the members of the team.

To study this problem of the organization of the research activity, we will use the framework of the moral hazard in teams literature (Holmstrom, 1981). This means that we will assume that the profits of the firm are contractible (we rule out the ability of the owner of the physical assets to cream-off a fraction or the whole of the profits of the firm, as it is the case in the incomplete contracts literature –see Hart (1987)). Firm's profits and equity can therefore be used as input in agents' incentive schemes. Moreover, there is an action that is to be taken by all agents. This action cannot be specified in a contract. But in our setting this action is not an effort level chosen by the agent. Rather it corresponds to the communication of (and trade for) the learning capital to agents that had previously no access to it.

The analysis shows that the fact that some innovations cannot be specified in contracts has a large number of implications on the organization of research and development. Our main results can be summarized as follows.

The model implies that incentive schemes can be interpreted as factors that preempt entry to an industry. Under employment contracts that offer flat wages to employees, the incentives to collude against the firm are strong, since it represents a positive gain and no loss to the agent that sells information. This suggests a different understanding of incentive contracts written in firms, with different normative and positive implications.

We have find as well that the inability to control the diffusion of knowledge that leads to innovation implies a restriction to the size of the research team. Research is conducted efficiently only if researchers are brought to the project. Yet increasing the size of the team dilutes the ownership of the income stream and increases the incentives of individual members to collude against the firm. This trade-off between efficiency and appropriability of rents is shown to be crucially affected by the nature

of competition (bertrand, bertrand with differentiated products) in the industry. In particular, if competition in the market for innovation is severe (little product differentiation) the difference between the profits of the monopolist and the sum of the duopoly profits will be large. A team member that sells firm's information to outsiders induces entry and competition in the industry. Therefore bertrand competition will imply that the payoff from collusion against the firm (a function of the profits of the duopolist) is relatively small and that apropriation of rents from innovation is more likely.

Furthermore, we show the role of sunk investments in physical assets and the scale of the firm, for appropriating rents from knowledge-based innovation. The result is similar to the preemption games [Fudemberg and Tirole, (1984)], but in our case the strategic investments in physical assets are jointly determined with incentives schemes, size of the firm and payoffs from collusion, the result on strategic investment departs from the Stackelberg equilibrium. The ability to sink an investment in physical assets that affect the payoffs of a potential competitor allows the entrepreneur to improve on the extraction of rents from the knowledge that she is introducing. This implies that the problem of organization of research should be interpreted as a preemption game between the entrepreneur and the employess.

Finally, we extend the model to analyze the role of the allocation control rights among agents over ex post actions that are to be taken in the firm (this refers to the decision on investments in downstream activities that cannot be done until research takes place and more information relevant investment has been created), in relationship with the problem of knowledge diffusion and entry.

Our analysis yields empirical implications that are unique with respect to the related literature. In addition, the model has strong social welfare implications that relate to the relevance of policy variables like the size and the implementation of subsidies to research and development and the regulation of the contracts that affect the turnover of jobs by employees.

The paper is formally related to the literature on entry and strategic investments for preemption, and with that on the theory of the firm. More specifically, with the theories of moral hazard in teams. We believe that results of this kind, that derive from modelling jointly the organization of the firm, technological spillovers and preemptive investments, are not closely related to any previous contribution.

In particular, the idea that an important reason to explain the existence of firms is the fact that some form of intelectual property rights cannot be protected (and that asset ownership, control rights and incentive schemes are optimally set up so as to influence the diffusion of this knowledge) is seen as the main contribution of the paper.

The paper is organized as follows: section two lays out the model and solves for the optimal contract offered by the entrepreneur, in the case where the productive investments are purely in human capital. Section three extends the model to introduce *ex ante* investments in physical capital by the entrepreneur, investments that are to be sunk before any other agent becomes informed about the learning capital. Section four studies the case of *ex post* investments, that are to be sunk after research (and therefore communication) takes place. The paper closes with a discussion of extensions and policy implications in Section five. Most of the proofs of the lemmas and results are to be found in the Appendix.

1.2 The Model

The model derives the optimal contract offered by an entrepreneur to an employee. The contract specifies the incentive schemes affecting the researcher and the entrepreneur, a transfer from the researcher to the entrepreneur and the size of the research team (taken for simplicity to be either one or two). In the extension of the model that includes ex post investment opportunities, the contract specifies as well the allocation of control rights among team members and exclusivity clauses that restrict agent's ownership of equity from two firms.

The model is divided in two main parts and an extension. The first part (Section 3) considers the case where physical assets play no role in the outcome of research. In the second case (Section 4) physical assets affect players' payoffs. The cost of those assets can be sunk at the initial time of the strategic interaction.

The extension (Section 5) studies the case where the physical asset that affects payoffs cannot be introduced at the initial stage of the game (e.g., because it is needed to conduct a downstream activity, like the manufacturing or marketing of the good resulting from research). Investment can only take place after the outcome of the research effort has been realized.

The starting point of the model is a situation where a research opportunity can be undertaken. If research is successful, it leads to the ability to produce a good with value to consumers. Conducting research requires access to an specific piece of information (the learning capital). Only the entrepreneur "owns" (knows) the learning capital initially. The entrepreneur has to decide whether to bring a researcher to the project, therefore sharing with the latter the learning capital. The entrepreneur cannot enforce contracts restricting the communication of the learning capital to third agents.

1.2.1 Players

There are two relevant players:

Entrepreneur An agent who is a monopolist over the information capital at an initial point on time, t=1. (The timing is defined in subsection 3.1.3.).

Researcher an agent chosen from a competitive market of potential agents [2] that do not have the learning capital at t=1, but that are able to innovate conditioned on having access to it.

[2]This assumption of perfect competition in the market for potential researchers allows us to assign all the bargaining power for the distribution of surplus to the tendering party, simplifying the analysis. The results are robust to more general specifications

1.2.2 Timing

t=1 The entrepreneur can make a take it or leave it offer to the researcher of a contract that specifies:

Profit Sharing: the distribution of the revenue from innovation between the researcher (b_R) and the entrepreneur ($(1 - b_R)$).[3]

Transfer A: from the researcher (if any) to the originator.

The contract, if offered, is either accepted or rejected by the researcher.

t=2 Costless research effort takes place.

t=3 The (non-licentiable) innovation can sold by any informed agent in a competitive market of downstream firms.

$t = 1$	$t = 2$	$t = 3$
contract offered	research	innovation sold

1.2.3 Contractibility and information structure

A central assumption of the paper is the following:

ASSUMPTION A

A.1: The information that allows the research effort (and that we called learning capital) cannot be recognized by the court (i.e., diffusion of information by agents is not verifiable).

[3]We are assuming that only balanced mechanisms can be used, i.e., it is not possible to "break the budget constraint" in order to make all agents residual claimants (given that they are assumed to be risk neutral). In general these contracts are known to be non-robust to collusion. We avoid the issue here and assume that only equity holding agreements can be written (where equity can have either voting and non-voting power).

A.2: Trade for this information can be privately enforced. [4]

A.3: Information is symmetric. All variables and actions are observable and common knowledge to all agents.

1.2.4 Profit functions of firms

The profits of a firm from the sale of the innovation at $t = 3$ are expressed as: $X(i)$, where X depends on the number of firms producing the good and i is the number of agents that did research in the firm.

In particular we will assume:

$X(i) = M(i)$ if only one firm produces the good.

$X(i) = D(i)$ if two firms produce the good.

$X(i) = 0$ if more than two firms produces the good.[5]

These functions satisfy:

$$X(i) > X(j) \;\; iff \;\; i > j = 1$$

$$M(i) \geq 2.D(i) > 0 \;\; \forall i > 0$$

NOTATION Profit functions will be henceforth expressed as follows:

$$X(i) = X \;\; if \;\; i = 2 \;\; and \;\; X(i) = X_i \;\; if \;\; i = 1, \;\; for X = \{M, D\}$$

[4]In this respect we follow the collusion literature, (Tirole, 1991) assuming that collusive contracts can be enforced.

[5]this assumption simplifies the analysis but is somewhat artificial. It could on the other hand be avoided easily without damaging the results of the paper.

It is clear that once the size of the team is equal to one, the possibility of diffusion and hence of entry disappears. This is why we don't need to consider $D(i = 1)$.

The assumption that $M_1 \leq M$ should be interpreted as follows. If two agents are engaged in research at $t = 2$ then innovation takes place with probability one. If only one agent does research, then innovation takes place with probability α, and expected profits of the firm at $t = 1$ equal $\alpha.M$, lower or equal than M.

1.3 Human Capital as the only productive factor

This section solves the optimal organization of R&D in the case where physical assets do not affect the productivity of research (i.e., under the assumptions layed out in Section 2). Physical assets will be introduced in the following two sections.

1.3.1 Result 1: Incentive schemes to preempt entry

Result 1 contains two insights.

The first one is that it shows how profit sharing agreements can be used as entry deterrent to an industry. If insiders to a firm engaged in research are not claimants to a part of the income stream, they have incentives to diffuse the relevant information outside the firm, since they are not affected by the decrease in profits that follows their collusive agreement with firm's outsiders.

Result 1 shows how the nature of competition in the industry affects this trade-off between efficiency (large team) and appropriability (incentives to collude against the firm). If the profit-destroying effect of competition is intense, then a team member that colludes against the firm obtains a relatively small reward from the entrant. At the same time the income of this agent arising from his share of the initial firm's profits is much decreased.

The second one is that the inability to perfectly protect the learning capital of a

firm with trade secrets or patents imposes in general a restriction in the size of the research team in a firm. When the incentives of team members to collude against the team are strong, the entrepreneur will choose between achieving duopoly profits (notice that any revenue to the employee from collusion will be internalized by the entrepreneur through the transfer A in the contract signed at $t = 1$). The entrepreneur has therefore to choose between receiving the total of the duopoly profits of the industry and the monopoly profits realized through inefficient research (when the size of the research team is restricted to be equal to one). If the synergy between team members is low enough, the researcher will prefer to keep the size of the research team small.

Lemma 3.1.1 derives the players'payoffs from the continuation game at $t = 3$. If one agent unilaterally sells the innovation to third agents, she is able to make a take it or leave it offer. Hence the duopoly firms of the entrant enter in the payoff of the seller. If both agents sell the learning capital, then there are three firms in the market and all firms make zero profits.

Lemma 3.1.1

The payoffs to players from the continuation game at $t = 3$, as a function of the profit sharing agreement are given by the following matrix (where the columns depend on the action of the entrepreneur and the rows to the actions of the researcher.

R / E	*do not sell*	*sell*
sell	$[b_R.M]$, $[(1-b_R).M]$	$[b_R.D]$, $[(1-b_R).D+D]$
do not sell	$[b_R.D+D]$,$[(1-b_R).D]$	[0],[0]

Result 1 shows the conditions under which the entrepreneur will bring a researcher to the team and still will control the spread of information and sustain the monopoly position over the innovation. For this it is sufficient that the no diffusion constraints are satisfied for both agents, i.e., that:

$$\beta.M \geq \beta.D + D, \;\; \beta \in \{b_R, 1 - b_R\}, \;\; b_R \in [0,1]$$

If the no diffusion constraint is not satisfied for any $b_R \in [0,1]$, then the entrepreneur chooses between:

1) efficiency in research (bring a researcher to the team) with diffusion of innovation to another firm and entry, and,

2) inefficiency (one agent research team), preservation of secrecy and monopolization of the industry.

Result 1

If $3.D \leq M$ then:

$$b_R^* \in [\frac{D}{M-D}, \frac{M-2.D}{M-D}] \text{ and } A = b_R^*.M$$

If $3.D > M$, then if $M_1 < 2.D$ the researcher is hired with $b_R^* = W = 0$ and entry takes place.

Finally, if $3.D > M$ and if $M_1 > 2.D$: then there will be only one member in the research team.

Liquidity Constraints

So far we have assumed that agents don't face liquidity constraints. This subsection considers the effect of liquidity constraints on the optimal contract. Since this case should be considered more realistic than the one without liquidity constraints, and since liquidity constraints are found to have implications on welfare, this case deserves special attention.

Consider the extreme opposite case, where the researcher has no wealth and he cannot borrow in the credit market, nor from the entrepreneur, so that the initial transfer A is equal to zero.

This has two implications. The first one is that the entrepreneur is not anymore able to internalize the rents derived from ex post collusive agreements between the employees and third parties. The second one is that the researcher will be unable to purchase equity from the entrepreneur's firm. Setting up an incentive scheme for the researcher implies therefore a direct loss of profits for the entrepreneur.

This implies that only in the case where the entrant's profit will be zero (as in the case where competition for the sale of the innovation is a Bertrand price game without product differentiation), will the entrepreneur be able to make monopoly profits with under efficiency in research.

Under liquidity constraints of the researcher the constraints that set a limit in the size of the firm are stronger. The research team will more often be a one member team, with negative implications on welfare.

Implications about incentive schemes

Result 1 has a clear implication on the correlation between the diffusion of knowledge (and the rate of entry in an industry) and the distribution of profits across firm's insiders. Namely, this theory of incentive schemes argues that faster diffusion and entry implies a more unequal distribution of profits across firm members, and vice versa. (This prediction is in turn related to the nature of competition and the scope for product differentiation in the industry, as explained above).

The main implication of result 1 is that it identifies a direct and potentially testable link between the nature of competition in an industry (resulting from structural variables like the elasticity of product differentiation) and the organization of the firm (meaning the incentive schemes and size). In order to interpret the implications of result 1 as a function of more primitive parameters, the following example derives

the ratios $M/M-D$ and $M-2.D/M-D$ in a model of Bertrand competition with differentiated products. If the no-diffusion constraints are satisfied and the ratio $M/M-D$ is lower than $M-2.D/M-D$, then these two ratios define an interval. The *size* of this interval should be interpreted as a measure of the appropriability conditions that the entrepreneur faces.

Result 1 has shown how the ability to perform research efficiently (in a team) without the risk of technological spillover depends on the difference between monopoly and duopoly profits. In particular, it depends on whether the inequality $D/(M-D) < (M-2.D)/(M-D)$ is satisfied. Factors that make this interval be broader are factors that favor the "cohesion" of the firm, i.e., factors that reduce agent's incentives to collude against the firm.

Consider a model of horizontal differentiation where consumers are distributed uniformly over the interval [-1,1] and have willingness to pay for one unit for the good of the equal to θ. There is a transportation cost per unit of distance equal to t. Production costs are zero.

If there is only one firm in the economy it will locate at the midpoint of the interval (0). Profits of the monopolist are given by: $\pi_m = (\theta)^2/2.t$.

If there are two firms, they will locate at opposite interval extremes, (-1) and (1) respectively. Their equilibrium profits can easily be shown to be symmetric and equal to 2.t.

Then the ratios D/(M-D) and (M-2.D)/(M-D) are given by:

$$\frac{2.t}{(\theta^2/2.t)-2.t} \quad and \quad \frac{(\theta^2/2.t)-4.t}{(\theta^2/2.t)-2.t}$$

It is easy to show that as t increases (i.e., as the differentiation of the products increases), the interval so defined becomes smaller, both because the left extreme increases and the right extreme decreases. I.e., it becomes more difficult to retain

knowledge in the firm.

It follows as well that as θ (the willingness to pay, i.e., demand) increases, then the interval becomes strictly larger (i.e., firm cohesiveness increases). This is because as demand is higher the opportunity cost of competition is higher (competition becomes relatively more profit-destructive) therefore making it more likely that no agent spreads information outside the firm.

Moreover, result 1 implies that if the parameter that determines product differentiation (t above) decreases over time (as it is likely the case, if consumers become more informed about the new products over time) or as the production costs reduce over time (through learning by doing) the conditions of appropriability of innovation become more favorable for the innovating firm. This implies a nonstationarity in the evolution of firm sizes over time, and in particular it implies rates of firm's entry that decrease over time and increases in the size of the firm with the evolution of the industry.

1.4 Investment in Physical and Human Capital

This section extends the model in 3.1 to analyze the role of investments in physical assets on the organization of the research team. The physical asset should be interpreted as equipment that is needed to perform research, like laboratory supplies, or information technology that increases the productivity of the research team, like data base managers. It could also be interpreted as factors that decrease the transaction costs of the trade for the innovation that is to generated, like the quantity and quality of patent lawyer services that are hired in the firm.

1.4.1 Timing

The timing of the game is the same as the one described in (2.1.2), except that we introduce two additional stages.

$t = 0$: We consider the case where the entrepreneur is able to sink the cost of a physical asset K, at time $t = 0$, before offering an employment contract to the researcher at $t = 1$.

$t = 4$: After the innovation has been sold to downstream firms (at $t = 3$), the buyers of the innovation can make an invest in physical capital K_c.

The complete timing in the case of ex ante physical assets is summarized in table 4.1.

$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 4$
K	contract offered	research	innovation sold	K_c

Table 4.1. Timing with ex ante investments

1.4.2 Profit functions

Introducing physical assets in the model requires to specify how physical capital affects the profit functions of the firms. Assumption K states these conditions. [6]

ASSUMPTION K:

K.1: The Monopolist profit functions $M(K)$ and $M_1(K)$ are strictly concave functions of the investment level K, with strict maxima at K^* and K_1^* respectively.

K.2: Let $D(K, K_c)$ be the profits of a duopolist, with K being the investment of the firm and K_c the investment of the competitor.

$D(K, K_c)$ is strictly concave in K, and reaches an strict maximum and $\frac{\partial D(K,K_c)}{\partial K_c} < 0 \ \forall K_c$.

[6]The following K.1-K.4 assumptions would be satisfied by a standard "Cournot" quantity game with linear demand and linear or quadratic costs of capacity.

K.3: Furthermore we will assume that the reaction function of a duopolist, defined by:

$$\psi(K_c) \equiv \{K : \frac{\partial D(K, K_c)}{\partial K} = 0\}$$

is downward sloping, continuously differentiable and that $\psi(K)$ has a unique fixed point in $K_N > 0$. i.e., $\psi(K_N) = K_N$, (K_N is therefore the unique Nash equilibrium of a two-player, symmetric simultaneous investment game).

The reaction function satisfies: $\mid \psi'(K_N) \mid < 1$, i.e., the Nash equilibrium is stable. And that there is finite value of K, $\bar{K}$, such that: $D(\psi(\bar{K}), \bar{K}) = 0$

K.4: For all K, $M(K) \geq D(K, \psi(K)) + D(\psi(K), K)$

Notice that K.3 ($\psi'(K) < 0$) restricts the generality of the duopoly functions. In particular we are assuming that the variables K and K_c are strategic substitutes.

1.4.3 Result 2: organization of the firm as a preemption game

The ability to use the investment in physical assets allows the entrepreneur to play a preemption game with the employees to the research team. This ability has as a sine qua non condistion that the entrepreneur is the agent that effectively **sinks** the cost of the investment at $t = 0$, before the employment contract is offered.

The investment in physical assets affects the incentives of the agents to collude against the firm because it reduces the profits of entry. The level of investment that the entrepreneur needs to sink to deter entry determines the profit sharing arrangement that is set in the entrepreneur's firm. It is the entrepreneur's introduction of assets previous to the communication of the learning capital and its ownership what allows the entrepreneur to improve the appropriability conditions.

In particular, result 2 shows that the entrepreneur does not need to overinvest

to the level where entrant's profits are negative to effectively deter entry (as in the Dixit-Spence preemption game). The deterring investment is the one that makes entrant's profits *sufficiently low* so as to reduce the incentives of selling the learning capital outside the firm. This implies lower inefficiency and rent dissipation of the preemptive activity than in the models where the presence of the potential entrant is exogenous.

Lemmas 4.3.1 4.3.2 and 4.3.3 derive results needed to prove Result 2. Lemma 4.3.1 derives the payoffs to players from the continuation game at $t = 3$, given the profit sharing agreement b_R and the sunk investment in physical capital K.

Lemma 4.3.1:

The payoffs of the continuation game at $t = 3$ when a researcher was introduced at $t = 1$ are given by the following matrix:

R/E	*do not sell*	*sell*
sell	$b_R.M$, $(1 - b_R).M$	$b_R.D_0$, $(1 - b_R).D_0 + D_1$
do not sell	$b_R.D_0 + D_1$,$(1 - b_R).D_0$	0,0

where:

$$D_0 \equiv D(K, \psi(K)) \ \ and \ \ D_1 \equiv D(\psi(K), K)$$

Lemma 4.3.2 and 4.3.3 derive the profits of the entrepreneur at the $t = o$ stage of the game, as a function only of the investment level K. It derives from the fact that there is a second investment stage at $t = 4$, where both firms (in the case that entry took place) invest simultaneously.

Lemma 4.3.2:

The entrepreneur choses at $t = 0$ and investment level greater or equal than the

Nash equilibrium level of the simultaneous investment game (K_N).

Proof: in Appendix.

Lemma 4.3.3 determines the investment level that maximizes the profits of the entrepreneur, subject to the constraint that no collusion for the trade of information takes place. It shows that there is an interval of investment levels $[\hat{K}, \bar{K}]$ that preempt diffusion of the learning capital, and hence entry. The optimal preemting investment $\hat{K}$ is the one that maximizes the profits of the entrepreneur within the interval $[\hat{K}, \bar{K}]$. This turns to be the minimum investment level that belongs to the interval, and will in general be lower than $\bar{K}$, the investment level of the incumbent that sets the profits of the entrant equal to zero.

Lemma 4.3.3: *Optimal preempting investment*

$$Let\ P \equiv \{K \geq K^* : M(K) \geq D(K, \psi(K)) + 2.D(\psi(K), K)\}$$

And let $\hat{K} \equiv argmax M(K) s.t. K \in \{P\}$

Then $\hat{K}$ *always exists and* $\hat{K} = min\ P$

Proof: in Appendix.

From the previous Lemma and Assumption K it follows that the entrepreneur can always follow one of three strategies: to preempt entry by investing $\hat{K}$ (and set the profit sharing agreement accordingly), to accommodate entry, or to restrict the research team to himself and therefore preserve necessarily the learning capital under secrecy. This decision is described by Result 2:

Result 2:

If $M(\hat{K}) \geq max(2.D(K_N, K_N), M_1(K_1^*))$ *then:*

- *the entrepreneur sets* $K = \hat{K}$ *at* $t = 0$
- $b_R^* \in [\Gamma_1, \Gamma_2]$, $A^* = b_R^*.M(\hat{K}$

  where $\Gamma_1(K) = D(\psi(K), K)/[M(K) - D(K, \psi(K)]$

  and $\Gamma_2(K) = [M(K) - 2.D(\psi(K), K)]/[M(K) - D(K, \psi(K)]$

If $2.D(K_N, K_N) > max(M(\hat{K}), M_1(K_1^*))$ *then:*

- *the entrepreneur sets* $K = 0$ *at* $t = 0$
- $b_R^* = A^* = 0$

If $M_1(K_1^*) > max(D(K_N, KN), M_1(K_1^*))$ *then:*

No contract is offered and the size of the research team equals 1

The role of Physical Assets:

Result 2 shows the effect of *large scale* in physical assets of a firm to induce agents to keep secrecy and therefore retain firms's competitive advantage.

The role of the entrepreneur is precisely that of committing and anticipating the physical capital, and to organize the firm so as to optimize between efficiency and appropiability.

Result 2 should be compared to the findings of the Incomplete Contracts literature on asset ownership (see Hart 1990 for a review of this literature). Ownership of physical assets matters in an incomplete contracts framework because agents cannot contract on the profits of the firm and because agent's effort investments show *specificities* to physical assets (the productivity of those investments depends on access to

the access to the particular assets for which property rights will contracted upon). If an agent's investment is relatively important, the agent should be given ownership, since this is the only way in which the agent will be compensated for exerting effort.

In our framework physical assets play exactly the same role as in the strategic preemption literature. The agent that does not own assets needs to purchase those assets and then *compete* with the incumbent in order to make profits. In our case, the incumbent is the entrepreneur and the entrant is the employee. The employment relation in the R&D startup is therefore interpreted as a preemption game.

Notice that our approach hinges on the inability to protect knowledge through contracts, but does not assume the ability of the owner to cream-off profits, nor the specificity of the investments in human capital.

1.5 Ex post investment in Physical Capital

This section introduces in the model physical assets that affect the profit functions of sale for innovation (the functions M(.) and D(.,.) defined above) that cannot be introduced before the time when the initial contract is offered ($t = 1$). Rather, the investment can only be decided after the time when research yielded its outcome.

The rational for this is that presumably there will be in general great uncertainty about the exact nature of the innovation and therefore also about the exact vintage of capital that is necessary to commercialize the new product or process. This section deals with the case where the investment can not be sunk until that uncertainty in the research outcome has unfolded.

The study of this case is indeed relevant because it makes the definition of control rights a central variable in the organization of the firm, and yields an interpretation of control rights that is different to that in the incomplete contracts literature.

The decision about whether this ex post investment in physical capital should take place is left to the agents that have control rights in the firm (this will be precisely

the definition of control rights). If an agent has all control rights of a firm, then this agent chooses the investment level K at $t = 3$. If two agents have control right, the investment level is chosen in unanimity by both agents.

The question asked in this extension of the model is precisely how the allocation of control rights (together with other variables described in the initial contract in subsection 2.1.1) affect the entrepreneur's ability to appropriate the rent deriving from the learning capital that she introduced.

Control rights matter in the model because they affect the costs and benefits of a particular agent, of colluding against the firm to which he belongs. In particular, an agent with veto power over the determination of the preemptive investment is able to delay certain actions that would diminish the profitability of an entrant. By taking those actions against the firm's interest, this agent would thereby be increasing the payoffs derived from the sale of information (i.e., from colluding against her own firm).

This will imply that in order to decrease the power of firm members to punish the firm in favor of outsiders, control rights should be concentrated on a subset of the firm members. We will see that agents with control rights will be claimants to a higher fraction of the firm's profits.

1.5.1 Timing

The timing is defined as in 2.1.1 except that:

$t = 3$: Innovation is sold in a competitive market. The researcher and the entrepreneur are able to sink the physical capital K.

$t = 4$: All agents with access to the innovation are able to sink an investment in physical capital K_c.

$t = 1$	$t = 2$	$t = 3$	$t = 4$
contract offered	research	K, innovation sold	K_c

1.5.2 Result 3: Concentration of control rights in one agent

Lemma 2.3.1 shows the payoffs to the researcher and the entrepreneur from the continuation game at $t = 3$, under the timing structure described in 2.3.1. The main difference with Lemma 2.2.1 is that now the investment level K cannot be taken as a parameter, but is a variable that is determined at this stage of the game.

Whether the level of investment in physical capital is an action in the strategy of **one** or **both** players depends on how the control rights are assigned (on whether they are given to one agent or they are shared by both).

The benefits of having control rights for an agent come from the ability that they give the agent to increase duopoly profits, buy not taking the best action for the interest of the firm.

To see this compare the no-diffusion-constraint of an agent that holds a fraction β of the (non voting) equity of the firm, when the agent has control rights and when she does not.

If the agent holds no control rights, then she chooses between retaining knowledge and diffusing (selling) it. But in either case the investment level is given to the agent. The no diffusion constraint is then, as before:

$$\beta.M(K) \geq \beta.D(K, \psi(K), K) + D(\psi(K), K)$$

If the agent holds control rights, then if she decides to spread information, she will affect K (over which she now has veto power) so as to maximize total profits. Total profits of a colluding firm members are: $\beta..D(K, \psi(K)) + D(\psi(K), K)$ which

is maximized by setting the investment at $t = 4$ equal to zero (see Lemma 3.2 above, and recall that there is a second investment opportunity).

The no-diffusion constraint under control rights of the agent becomes:

$$\beta.M(K) \geq \beta.D(K_N, K_N) + D(K_N, K_N)$$

Result 3 proves that this second inequality will in general not be satisfied when the previous one is, and that when the latter one is satisfied, then the first one also is satisfied. Equivalently, under a non-restrictive allocation of control rights (both agents have control rights) given that profits have to be divided, it is more likely that one agent finds more profitable to work against the firm (use her power to maximize collusive profits).

Lemma 2.3.1: The payoffs to both agents from the continuation game at $t = 3$ are given by the matrix in Lemma x. For the agents that have been assigned control rights, K belongs to their strategy space.

Result 3:

For the entrepreneur, the strategy of concentrating control rights in one agent (in the initial contract at $t = 1$) dominates the one of giving control rights to two agents.

Proof: In Appendix.

Result 3 yields a theory of hierarchies in firms, that is essentially different to the existing theories proposed in the literature. Hierarchy means here an organization where control rights are *concentrated* among a subset of members of the organization. The theory has not spelled what are the costs of concentrating the allocation of control rights over firm's actions among few agents, but only the benefits (make the firm more robust against collusion).

Yet it seems plausible that an argument against concentrating control rights in *one* agent would be a lack of robustness against mistakes in decision making, when all the power is given to a single agent, as well as the risk of transfering a great deal of bargaining power to an agent, in cases where decision making is characterized by a learning curve (i.e., when the decision maker becomes herself indispensable for the the exercise of power in the firm).

As opposed to the theory of the organization of the firm of Grossman and Hart (1986) and Hart and Moore (1990), our theory does not identify control rights with ownership of physical assets. In our model control rights mean the right to decide on ex post actions (actions that could or could not be specified in a contract: what can be contracted upon is the identity of the agent that will set the value of the variable at some point). This means that our results apply generally to firms where ownership and control are separated.

1.6 Conclusions

The precise mechanics of the diffusion of knowledge about innovations in products and processes has received little attention in the economics literature. Rather, the process of technological spillovers has in general been taken as a simple case of imitation by backward engineering of new products, after they arrive to the marketplace.

Moreover, the way that research and development has been characterized in the Industrial Organization literature is mainly by assuming that the output of R&D are patentable new products and processes.

This two premises taken together have allowed authors that have modelled the strategic interaction of agents engaged in innovation (patent races models, the strategic licensing literature) to take in general a "black box" approach to the organization of the firm. The firm in these models was characterized as an individual, rather than as an organization.

Without disputing that the goal of Research is the design and discovery of new products, we argue that the main output of research and development is *knowledge*, that will eventually be used in the discovery of patentable products. Frequently this knowledge, specially when it refers to know-how in new production processes, is not susceptible of being patented and can therefore be communicated without legal restrictions.

This description of the research process has as an implication the fact that it brings the organization of the firm in the early stages of an industry as the main determinant of the long term structure and performance of that industry. It also points to the factors that determine how the firm will be organized.

To make our case for this general claim, we have considered an economic setting where there is an agent that has superior information (a project with commercial potential), that needs human resources to exploit the project efficiently, but who is unable to protect the project from competition by patents. In this case, the main source of competition to the entrepreneur is precisely the employees that become informed about the nature of the project.

The role of the entrepreneur is therefore to set up the organization of the firm so as to optimize in the trade-off between efficiency (hire human resources) and appropriability (avoid the spread of knowledge about the project). This problem of organization means the optimal decision about the number of insiders in the firm (the entrepreneur plus the employees), the incentive schemes that regulate their retribution and the structure of asset ownership and control rights under which both the entrepreneur and the researcher will operate.

In order to model the microstructure of spillovers of knowledge, the firm has been described as a contractual devise to compromise between efficiency in production and the limitation of the diffusion of knowledge to the hands of competitors (appropriability). The analysis has relied on a number of assumptions, some of which deserve

a careful test on their content of truth and their sensibility when departures from them take place.

This is particularly true of the assumption on balanced budget: the insiders to the firm own a fraction of the firm that together adds up to the profits of the firm. The extensive use of bonus in compensation schemes gives already evidence on the inadequacy of this assumption. More importantly, we have avoided in the analysis the issue that is likely to be an important factor in explaining cases where the budget constraint cannot be "broken", namely, collusion.

The ability of innovators to affect the speed of diffusion of their ideas is likely to depend on the nature of the information. In this paper we have considered one of these possible paths of communication. We have pointed to information that is likely to be spread across most of the members of the firm and that can be communicated easily.

A case in point for our case is the information the strategic plan of the firm: a document that summarizes the set of short term decisions to be taken in a firm (exact timing of investments, future direction of research, hiring of human resources). A fraction of this information will be available to agents involved in taking these plans to conclusion. The communication of this information to competitors is not a costly undertaking and it clearly benefits competitors and damages the firm.

Yet there are other instances where the information that is the source of competitive advantage is more *technical*, i.e., only a few members of the firm are really able to use and communicate it to other agents, and moreover its communication is a lengthy process that requires probably to build a common language before it can be transmitted to agents that didn't generate the new knowledge. In such a case the diffusion of information has a sine qua non condition the migration of workers from one firm to another.

We see this case as a particularly relevant one (for which many examples come to

mind). To model the organization and strategies of the firm under this alternative assumption on the nature of knowledge and its diffusion would lead probably to results different to the ones derived here (migrating from one firm to another is a verifiable action; monetary punishments –like having to forfeit a rent– can be imposed on the employee that leaves one firm for another).

We see an interesting research question precisely in identifying technologies characterized by these and other forms of communication of the knowledge on which they are based, and deriving implications on the organization of the firms and industries that use those technologies, so that a testable body of knowledge about the spread of innovation would emerge.

1.7 APPENDIX.

Proof of Lemma 2.2.2

Proof of Lemma 2.2.3 If $\hat{K}$ preempts entry, it has to be the case that there is a profit sharing agreement b_R such that the no diffusion constraints for the researcher and the entrepreneur are satisfied:

$$b_R.M(K) \geq b_R.D(K, \psi(K)) + D(\psi(K), K)$$

$$(1 - b_R).M(K) \geq (1 - b_R).D(K, \psi(K)) + D(\psi(K), K)$$

Taken together, these two inequalities define the set of preemptive investments.

From K.4, there exits an investment level $\bar{K}$ such that $D(\psi(\bar{K}, \bar{K}) < 0$. Hence $\bar{K}$ belongs to the set of preemptive investments, since:

$$M(\bar{K}) \geq D(\bar{K}, \psi(\bar{K})) + 2.D(\psi(\bar{K}), \bar{K}) = D(\bar{K}, \psi(\bar{K}))$$

By K.1, M(K) is maximized at K^*. Hence if K^* belongs to the set P, then $\hat{K} = K^*$

Proof of Lemma 4.3.3 That the set P is nonempty is clear from assumption K.3: the investment level $\bar{K}$ is defined by $D(\psi(\bar{K}), \bar{K}) = 0$. Since $M(K) \geq D(K, \psi(K)) + D(\psi(K), K)$, then $M(\bar{K}) \geq D(\bar{K}, \psi(\bar{K}))$.

The optimal K in P being the minimum in P derives from the fact that if K belongs to P then $K \geq K^*$, by the definition of P, and from the fact that $M(K)$ is strictly concave.

Proof of Result 3:

Consider the case where control rights are given to both the researcher and the entrepreneur. Then the no diffusion constraint for both agents is given by:

$$\beta.M(K^*) \geq \beta.D(K_N, K_N) + D(K_N, K_N)$$

for $\beta = \{b_R, (1 - b_R)\}$.

This is so because an agent with control rights can veto the efficient investment K^* (efficient conditioned on no diffusion and monopolization, see assumption K.1 in 4.2 above). To veto K^* implies that the firm will be engaged in a simultaneous investment game at $t = 4$, see timing in 4.1). In this case, incumbent's and entrant's profits will be equal, with a value of $D(K_N, K_N)$.

This implies that monopoly profits $M(K^*)$ can be achieved if and only if:

$$\frac{D(K_N, K_N}{M(K^*) - D(K_N, K_N)} \leq \frac{M(K^*) - 2.D(K_N, K_N)}{M(K^*) - D(K_N, K_N)}$$

Consider now the case where only one agent (say, the entrepreneur) holds control rights over the investment level K. This implies that the researcher can now choose between preserving secrecy or selling the learning capital of the firm, but that she should take the investment level as given. This yields the following no-diffusion constraints:

for the researcher:

$$b_R.M(K^*) \geq b_R.D(K^*, \psi(K^*)) + D(\psi(K^*), K^*)$$

And for the entrepreneur:

$$(1 - b_R).M(K^*) \geq (1 - b_R).D(K_N, K_N) + D(K_N, K_N)$$

This implies that $M(K^*)$ can be attained if only if:

$$\frac{D(\psi(K^*),K^*)}{M(K^*)-D(K^*,\psi(K^*))} \leq \frac{M(K^*)-2.D(K_N,K_N)}{M(K^*)-D(K_N,K_N)}$$

The proposition follows from the comparison of these two intervals. Notice that their upper extreme is equal in both intervals and that the lower interval is *smaller or equal* when control agents are concentrated in one agent.

To show that:

$$\frac{D(\psi(K^*),K^*)}{M(K^*)-D(K^*,\psi(K^*))} \leq \frac{D(K_N,K_N}{M(K^*)-D(K_N,K_N)}$$

notice that: $M(K^*) \geq 2.D(K_N,K_N)$ (by hypothesis)

and that: $M(K^*) \geq D(K_N,K_N)+D(K_N,K_N),\ \forall K$

The latter inequality follows from:

$$\frac{\partial D(K,\psi(K))+D(\psi(K),K)}{\partial K} =$$

$$\frac{\partial D(K,\psi(K))}{\partial K}+\frac{\partial D(K,\psi(K))}{\partial K_c}.\psi'(K)+\frac{\partial D(\psi(K),K)}{\partial K}.\psi'(K)+\frac{\partial D(\psi(K),K)}{\partial K_c}$$

evaluated at $K=K_N$ equals:

$$\frac{\partial D(K_N,K_N)}{\partial K_c}.(1-\mid \psi'(K_N)\mid)$$

which is negative from K.2 and K.3 in assumption K.

and $D(K,\psi(K))+D(\psi(K),K)$ being strictly concave.

Bibliography

[1] Fudemberg, D. and J.Tirole. 1984. The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look. *American Economic Review Papers and Proceedings* 74: 361-68.

[2] Griliches, Z. 1991 Search for R&D Spillovers. *NBER Working Paper Series* number 3768.

[3] Ghemawat, P. 1991. Commitment: the Dynamics of Strategy. *New York: Free Press*

[4] Grossman, S. and O. Hart.1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy* 94 691-719.

[5] Hart, O. 1987. Incomplete Contracts and the Theory of the Firm. *Journal of Law, Economics and Organization* vol.4 1 119-139.

[6] Hart, O. and J. Moore. 1990. Incomplete Contracts and the Theory of the Firm. *Journal of Political Economy* vol. 98 6 1119-1158

[7] Holmstrom, B. Moral Hazard in Teams. *Bell Journal of Economics* 13 324-340.

[8] Holmstrom, B. and J. Tirole. 1987. The Theory of the Firm. In *Handbook of Industrial Organization*, ed. R. Schmalensee and R. Willig. Amsterdam: North-Holland.

[9] Levin, R., A. Klevorick, R. Nelson and S. Winter. 1987. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 3 783-820.

[10] Mishina, K. 1991. "Knowhow Spillovers and Industry Evolution". *Working Paper. Harvard Business School.*

[11] Office of Technology Assessment. 1984. Commercial Biotechnology: an International Analysis. *Washington D.C. U.S. Congress.*

[12] Office of Technology Assessment. 1991. Biotechnology in a Global Economy. *Washington D.C. U.S. Congress.*

[13] Pisano, G. and P. Mang. 1992. Manufacturing, Firm Boundaries and the Protection of Intellectual Property. *Harvard Business School.* Working Paper.

[14] Rodriguez, D. 1992. The Speed of Innovation and Diffusion of Know-how in Biotechnology. *mimeo, M.I.T.*

[15] Rumelt, R. 1984. Towards a Strategic Theory of the Firm. In R. Lamb ed. *Competitive Strategic Management:* 566-570. Englewood Cliffs, NJ: Prentice-Hall.

[16] Shan, W. 1990. An Empirical Analysis of Organizational Strategies by Entrepreneurial High Technology Firms. *Strategic Management Journal* 11 129-139.

[17] Spence, M. 1984. Cost Reduction, Competition and Industry Structure. *Econometrica* 52 101-122.

[18] Teece, D. 1986. Profitting from Technological Innovation: Implications for Integration, Collaboration, Licensing and Public Policy. *Research Policy* vol. 15 6 285-305

[19] Tirole, J. 1989. The Theory of Industrial Organization. *The M. I. T Press*

Chapter 2

Decentralization and the Management of Competition.

Abstract

This paper studies the costs and the benefits of decentralizing decision-making authority to competing divisions of a firm. We argue that when a division is given the right to market its product to customers of the other division, its manager has an incentive to increase product quality. At the same time divisions locate their products too close together. We find that top management may or may not want to restrict competition among divisions and show that even if competition is desirable, it may not be credible due to top management's temptation to intervene "ex-post".

2.1 Introduction.

Almost any large firm is a multiproduct firm. Some of its products are, in addition, horizontally or vertically differentiated i.e. are targeted to different customers but are at the same time close substitutes for some of these customers. The design and pricing of these products require thus considerable coordination among the firm's units.

Lack of coordination in product design was a major concern of top management at GM, before the radical reform of the company in the early 20's.[1] According to Sloan [8], GM was extremely decentralized and was offering, as a result, too many competing products.

> *In the middle [segment of the market], where we were concentrated with duplication, we did not know what we were trying to do, except to sell cars which, in a sense, took volume from each other.*[2] [3]

At the same time large firms may seek to promote competition among their units. For instance, in the last decade IBM drastically decentralized its internal structure allowing units to bargain on transfer prices and to trade with outside partners.[4] In addition, under decentralization:

> *[Units] are free to compete with each other for business, even if this means cannibalizing the sales of other IBM units.*[5]

As the above examples suggest, the existence of units offering substitute products within a firm raises a number of questions for organizational design. When should

[1] See Sloan [8] chapter 4.

[2] Sloan [8] p.60

[3] Another example of a firm centralizing decision-making to reduce competition among units is described in Smith [9] (see section 2 below). A major commercial bank in California centralized its operations, removing branch managers' autonomy to sell loans to customers. Branch managers had, instead, to refer customers to a center that coordinated trade.

[4] See Ferguson [1] for a detailed description of the evolution of IBM.

[5] Ferguson p.61 and 62.

top management control the design and pricing of these products, thus centralizing decision-making, and when is it better to allow the units decide which products to make and how to sell them? In the case where it is optimal to let the units decide (and eventually compete) what prevents top management from intervening and restricting competition when the latter threatens to dissipate too many of the firm's profits?

In this paper we develop a simple model to address the above questions. We consider a firm which is composed by top management and two divisions. Divisions design horizontally differentiated products, characterized by their "location" and their "quality", and sell them to customers. For simplicity, we focus on the allocation of decision rights at the price setting stage. More precisely, top management can restrict the set of customers with whom a division can deal, thus prohibiting competition for customers, or can alternatively allow both divisions to approach all customers.

We argue that when a division is given the right to deal with all customers, it offers a higher quality product. Indeed, by increasing the quality of its product it not only can charge its own customers higher prices, but it also becomes a better competitor of the other division. Using competition as a threat, it can extract concessions from the other division, such as dealing with some of its customers.[6] However, higher quality comes at a cost. A division aiming also at attracting the customers of the other division, will opportunistically locate its product too close to the product of the other division.

Top management will allow divisions to compete, i.e. will *decentralize* decision making, when the benefits due to higher product quality outweigh the costs of inefficient locations. We argue that top management's information about customers' characteristics (i.e. their locations) crucially determines the credibility of its commitment to competition among divisions. If it knows customers' locations it can efficiently allocate them to divisions (since in equilibrium it also knows the characteristics of both products) without letting the divisions engage in bargaining and

[6]In the model we assume, for simplicity, that divisions can make side-transfers so that concessions can simply be these transfers.

eventually compete.[7]

We show that centralization of decision-making may or may not be optimal. If divisional managers are very risk-averse, centralization dominates decentralization since the benefit of increased incentives is small compared to the cost of inefficient locations. Interestingly, the same conclusion holds when divisional managers' risk-aversion is low. Indeed, high-powered incentive schemes for improving product quality can be provided at a small cost under centralization. Decentralization thus brings small additional benefits and inefficient locations. Decentralization is a better solution if the set of customers who desire both products is small so that divisional incentives to locate products closer are small and locations are close to their first-best values. Costs due to inefficient locations are in the second order, while product quality increases in the first-order.

As we argued above, when a division is given the right to deal with all customers it will offer a higher quality product. Indeed, by becoming a better competitor of the other division, it can extract some concessions from it. Taking surplus away from the other division does not have any social value and if divisional managers receive very high-powered incentive schemes so that they become residual claimants, they will overprovide quality. The idea that competing agents may overinvest in some activities is familiar in industrial organization; firms may offer too many products because they are only "stealing" each other's business and may spend too much on R&D (Tirole [10] Chs. 7 and 10). In our model competition is valuable because agents' incentives are initially *too low*.

The channel through which competition affects incentives in this paper is different than the one studied in the principal-agent literature, namely that competition improves the information available for contracting. Papers in that literature have shown that the performance of an agent can better be assessed if performance measures of competing agents are available and that, as a result, this agent can be given

[7]In our model bargaining is costless, for simplicity. However, top management prefers to intervene even for a very small cost of bargaining (which can easily be introduced in the model).

a higher-powered incentive scheme.[8] Even if the principal does not have access to performance measures of competing agents, competition may still improve the information available for contracting by making the agent's profits less variable (Hart [2], Scharfstein [6]).

Rey and Tirole [5] also address the question of whether a principal (in their model a manufacturer) prefers to restrict competition among agents (retailers). A monopoly retailer will charge a price which is too high form the manufacturer's viewpoint, as long as he does not buy from the manufacturer at marginal cost. Competition allows the manufacturer to resolve this moral-hazard problem (since retailers will charge the price at which they are buying from the manufacturer) and at the same time provides insurance to retailers making their profits less variable, as in Hart [2]. When retailers compete they will thus charge lower prices (keeping the price at which they are buying from the manufacturer constant) and will on average supply the same quality. By contrast in our paper, when divisions are allowed to compete, they offer higher quality products and charge higher prices as a result. A second difference with Rey and Tirole [5] is that the cost of competition in their model is the inability of retailers to price-discriminate, while in our model is the choice of inefficient locations by the divisions.

Our paper is also related to a recent paper by Holmstrom and Tirole [4] studying transfer pricing within firms. They compare a regime of "exchange autonomy" where top management allows divisions to trade with outside partners to a regime where divisions can only trade internally (but can refuse to trade).[9] They show that if divisions are allowed to trade with outside partners, they offer higher quality products but they may choose a product design which is general purpose and not specific to the needs of the other division. Although studying different issues, both papers share the feature that if divisions have more decision rights they will have more incentives but they may also take decisions which are not optimal from the firm's viewpoint.

[8] For summaries of the literature on relative performance evaluation see, for instance, Holmstrom and Tirole [3] and Tirole [10], Ch. "The Theory of the Firm".

[9] They also study the case where one division can order the other division to trade.

The remainder of the paper is organized as follows: In section 2 we describe the model and in section 3 we derive the first-best. In sections 4 and 5 we study the outcomes under decentralization and centralization. Section 6 compares the two organizational forms. All proofs are in the appendix.

2.2 The Model.

2.2.1 Supply and Demand.

We consider a firm which is composed by top management and two divisions. Each division makes one product. The potential buyers of these products (the "customers") are uniformly distributed in the interval $[0,1]$. The product of division i ($i = 1,2$) (referred to as product i) is characterized by its location, l_i, and its "quality", v_i. We assume that "transportation" costs are linear, so that the valuation of a customer located at x for product i is

$$v_i(x) = v_i - t|x - l_i| \tag{2.1}$$

We also assume that the location of each customer is known to the divisions and that the divisions can charge customer-specific prices.[10]

Product 1 can be located anywhere in $[0,1/2]$ and product 2 can be located anywhere in $[1/2,1]$. The manager of division i can improve the quality of product i be exerting effort e_i at a non-monetary cost $C(e_i)$. We will assume that:

$$v_i = v_0 + e_i, \qquad v_0 > 0, \tag{2.2}$$

and:

$$C(e_i) = \frac{1}{2}Ce_i^2 \tag{2.3}$$

[10]We plan to study the case where customers' locations are not known and divisions cannot price-discriminate in a later version of this paper.

The marginal cost of each of the two products is zero.

For simplicity, we will restrict the parameter space assuming that only one product cannot cover the whole market even if the manager of the corresponding division exerts first best effort. Since the optimal effort level when the whole market is covered by one product is $\frac{1}{C}$ and the optimal location in order to cover the market is $\frac{1}{2}$, our assumption ("assumption A") is equivalent to:

$$\frac{v_0}{t} + \frac{1}{Ct} \leq \frac{1}{2}$$

In words, product quality when effort is zero, v_0, has to be small compared to transportation costs, t, and the cost of exerting effort and improving quality, C, has to be large.

2.2.2 Information Structure and Contracts.

We assume that a divisional manager's effort as well as the actions that he undertakes to determine the location of his product are observed by the other divisional manager. However, neither of these actions is verifiable, i.e. divisional managers cannot write short-term contracts with each other that bind them to exert a given effort level or to choose a particular location for their product. This will naturally be the case if, as we will assume, top management does not perfectly understand product design (quality and location) and cannot enforce such contracts. The non-verifiability of actions will imply in particular that divisions may not be able to coordinate on an efficient choice of product locations. Finally, divisions are assumed to know the location of all customers.

Top management is assumed to be less informed about both divisions' environments than each divisional manager. In particular by not perfectly understanding product design, top management does not observe effort levels (product quality) and product location. We however assume that it knows the location of customers. Since

in equilibrium it also knows product quality and product location, it can infer which product is better for a given customer. We briefly consider the case where top management does not know customers' locations in section 6.

Divisional profits are partly generated by the revenues from dealing with customers. We assume that they contain an additional component which is random, normally distributed and independent across divisions,[11] and are thus a noisy signal of the actions taken by divisional managers. Profits are the only variables in our model which are verifiable by outside parties (such as courts).

We assume that divisional managers are given incentive schemes which depend only the profits of their own division (and not on the profits of the other division) (This assumption is more innocuous than it may seem. See the analysis of the centralized firm in Section 5.) are linear and have the same slope which is between 0 and 1. Denoting by π_i the profits of division i, the manager of division i thus receives:

$$A\pi_i + B_i \tag{2.4}$$

We assume that both managers are risk-averse and have a negative exponential VNM utility function with coefficient of absolute risk-aversion equal to α. Their reservation levels are normalized to zero.

We consider two organizational forms, the decentralized firm and the centralized firm. Top management allows divisions to compete for a given customer in the decentralized firm while it allocates customers to divisions in the centralized firm. We assume that under decentralization divisions bargain efficiently so that competition only defines the status quo in the bargaining game, and that bargaining power is equal. Since the incentive schemes given to divisional managers have the same slope, customers are allocated to divisions efficiently from the firm's viewpoint.

The timing is thus:

[11] It can be due, for example to other divisional operations.

Decentralized Firm:

Centralized Firm:

To illustrate our model, we consider the example of commercial banks in a large city. There is demand for bank loans but different individuals are interested in loans with different characteristics (business are interested in short-term lending, entrepreneurs in funding a start-up venture, consumers need cash for acquiring durable goods).[12]

Managers sell loans to customers and can undertake two kinds of investments in order to become more efficient:

- **"Supply" Effort:** They can acquire a skill that is not specific to any kind of loan, but rather improves the chances of striking deals in good terms (like learning state of the art software to gather information in a timely fashion, or information technology that improves readiness of on-line help to customers, whatever the kind of loan they want to purchase). (The supply effort corresponds to the investment e above.)
- **"Demand" Effort:** They can become specialized in being particularly efficient in customizing loans to a particular segment of the market (by learning detailed information on a particular sector of the population or following more closely the trends of an industry. (This corresponds to the location choice, l, above.)

[12] Heterogeneity is *large* and it is therefore difficult to segment the market in non-arbitrary intervals of potential customers.

The bank has two divisions in the city, that deal directly with customers. Each division is run by one manager. Each manager has expertise in the business and is able to assess the investment levels of the other manager. Top management is however unable to measure those investment levels.[13]

Top management can centralize or decentralize the firm's operations. gathering more or less precise information about customers.

Under decentralization, divisional managers are not required to obtain approval from top management to strike deals with customers. Under centralization divisional managers transmit information on a potential deal with a customer to top management, who either approves the deal or allocates the customer to the other division.[14]

Smith [9] provides a case study of the reorganization of several divisions in a major investment bank in California. She examines in particular[15] the transition from relative autonomy to a regime of centralization in a division of the bank. The manager of the division (referred to as L.) is being deprived of authority over some actions that she previously controlled. The control is being recovered by the next superior level in the hierarchy, the area management group (referred to as AMG.)

> *Branch top management centralized lending into specialized divisions, reflecting the bank's commitment to targeting and profiting from stratified market segments. The differentiation of the lending function into specialized units removed a significant source of L.'s authority. She no longer had authority to make loans, nor did she manage loan personnel. [...] L. forwarded [loan application] to the appropriate consumer, real state or commercial loan center.[...] Her role was [now] strictly one of referral.*[16]

[13]A manager can infer the ability (investment) of the other manager because customers transmit information from one office to another.

[14]The center is unable to separate the market geographically if, for instance, business have multiple locations.

[15]See Smith [9], pp.95-101.

[16]In addition, it becomes clear that the transfer of authority required a redefinition of the information flowing from lower level (L.) to upper level (the AMG). The information flow is defined in a

Even when L. did tabulate statistical information, the role she once played in evaluating and acting on that information had changed. Whereas formerly L. would have tracked branch information and used it to develop an integrated picture of branch performance, the job of tracking various kinds of information had been transferred into specialized sections in the AMG.[17]

The new PPCE and the extensive documentation it contained provided the area manager with the information required for indirect management: in other words, it gave area managers more control over individual branches.

2.3 The First Best.

In this section we derive the effort levels and product locations that maximize total welfare. Total welfare (which is equal to top management's payoff since the reservation levels of the divisional managers are normalized to zero) is:

$$\int_0^1 max(v_1(x), v_2(x), 0)dx - \frac{1}{2}C(e_1^2 + e_2^2)$$

$$= \int_0^1 max(v_1 - t|x - l_1|, v_2 - t|x - l_2|, 0)dx - \frac{1}{2}C(e_1^2 + e_2^2) \qquad (2.1)$$

Proposition 3.1, proven in Appendix A, gives us the first best effort levels, e_1^* and e_2^*, and locations l_1^* and l_2^*.

Proposition 3.1. *The maximum of expression 2.1 is achieved by choosing:*

$$l_1^* = \frac{1}{4} \qquad\qquad l_2^* = \frac{3}{4}.$$

If $2v_0/t + 1/Ct < 1/2$, e_1^* *and* e_2^* *are given by:*

document of the corporation, the Performance Evaluation Plan (PPCE in the text).

[17]Smith [9] p.97

$$e_1^* = e_2^* = v_0 / \left(\frac{Ct}{2} - 1 \right)$$

and the firm does not sell to all customers (i.e. the market is not "covered").

If $2v_0/t + 1/Ct \geq 1/2$, e_1^* *and* e_2^* *are given by:*

$$e_1^* = e_2^* = \frac{1}{2C}.$$

and the market is covered.

Assumption A makes the problem sufficiently concave so that it is not optimal to require different effort levels (and product locations) from the two managers.

If the market is not covered (i.e. if product quality is low compared to transportation costs and the cost of improving it is large), total welfare is maximum for any locations which imply non-overlapping sets of customers willing to buy a given product and which are not too close to the extremes, 0 and 1. (So that each product is located in the middle of its market segment - the set of customers who are buying the product.) 1/4 and 3/4 always belong to the set of optimal locations. If the market is covered, 1/4 and 3/4 are the unique optimal locations and each product is located in the middle of its market segment. ($[0, 1/2]$ for product 1 and [1/2,1] for product 2.) In addition, since more customers buy the products if the market is covered, the optimal effort levels are higher than if the market is not covered.

2.4 The Decentralized Firm.

We now come to the analysis of the decentralized firm. In section 2 we assumed that top management allows divisions to compete for a given customer. We then assumed that divisions bargain efficiently so that competition only defines the status quo of

the bargaining game, and that bargaining power is equal.

Considering a customer located at x and assuming, for example, that $v_1(x) > v_2(x) \geq 0$, it is easy to see that division 1 deals with the customer and receives $v_1(x) - v_2(x)/2$ while division 2 receives $v_2(x)/2$.[18]

Aggregating over customers, division 1's profits are given by:

$$\int_{S_1} \left(max(v_1(x), 0) - \frac{1}{2} max(v_2(x), 0) \right) dx + \int_{[0,1]\setminus S_1} \frac{1}{2} max(v_1(x), 0) dx =$$

$$\int_{S_1} \left(max(v_1 - t|x - l_1|, 0) - \frac{1}{2} max(v_2 - t|x - l_2|, 0) \right) dx$$

$$+ \int_{[0,1]\setminus S_1} \frac{1}{2} max(v_1 - t|x - l_1|, 0) dx \tag{2.1}$$

where the set S_1 is defined by:

$$S_1 = \{x \in [0,1] : v_1(x) > v_2(x)\} = \{x \in [0,1] : v_1 - t|x - l_1| > v_2 - t|x - l_2|\} \tag{2.2}$$

and division 2's profits are given by the symmetric expression.

Using equation 2.4 which gives managers' incentive schemes, we get for the payoff of the manager of division 1:

$$A \left(\int_{S_1} \left(max(v_1 - t|x - l_1|, 0) - \frac{1}{2} max(v_2 - t|x - l_2|, 0) \right) dx \right.$$

$$\left. + \int_{[0,1]\setminus S_1} \frac{1}{2} max(v_1 - t|x - l_1|, 0) dx \right) + B_1 - \frac{1}{2} C e_1^2 \tag{2.3}$$

To determine the outcome under decentralization we proceed in two steps. In Proposition 4.1, proven in appendix B, we characterize pure-strategy equilibria of the

[18] If there is competition, division 1 serves the customer charging the price $v_1(x) - v_2(x)$. Its payoff is thus $v_1(x) - v_2(x)$ while the payoff of division 2 is 0. Since bargaining power is equal, i.e. each divisional manager makes a take-it-or-leave-it offer to the other divisional manager with probability 1/2, it is clear that the payoffs of the two divisions are given by the previous expressions.

effort and location game taking incentive schemes as given. We then determine the optimal value of A, the slope of the incentive schemes.

Proposition 4.1. *There exists a pure-strategy equilibrium in which locations and effort levels are given by:*

case I: $2v_0/t + A/Ct < 1/2$

$$l_1 = \frac{1}{4} \qquad\qquad l_2 = \frac{3}{4},$$

$$e_1 = e_2 = v_0 / \left(\frac{Ct}{2A} - 1\right)$$

and the market is not covered. In any other pure-strategy equilibrium effort levels and payoffs are the same.

case II: $2v_0/t + A/Ct \geq 1/2$

$$l_1 = 1 - l_2 = \left(\frac{1}{4} + \frac{v_0}{2t}\right) / \left(\frac{3}{2} - \frac{A}{Ct}\right) \geq \frac{1}{4},$$

$$e_1 = e_2 = (2A/C)l_1 = \left(\frac{t}{2} + v_0\right) / \left(\frac{3Ct}{2A} - 1\right)$$

and the market is covered. This is the unique pure-strategy equilibrium.

In case I product quality, v_0, is low compared to transportation costs, t, the cost of improving it, C, is high and the benefit, A, that divisional managers get is small. The market is not fully covered and the divisions do not have incentives to offer products that appeal to overlapping sets of customers.

By contrast, in case II divisions offer products which appeal to overlapping sets of customers and whose distance is too small compared to the first best ($l_1 \geq 1/4$, $l_2 \leq 3/4$). The private gain of each division to locate its product closer to the product

of the other division is larger than the social gain. Indeed, by locating its product closer, the division better "penetrates" the market of the other division and becomes a more serious potential competitor.[19] The incentive to locate closer and becoming a more serious potential competitor is clearly larger if there is a significant overlap already (v_0 high, C small, A large).

Effort levels are higher in case II than in case I. Indeed, the manager of one division by exerting more effort supplies a product which is valued more by the customers of his division (who now have a larger mass) as well as by the customers of the other division (so that this product can better compete with the product of the other division).

It is straightforward to show that the top management's payoff is given by:

$$\int_0^1 max(v_1 - t|x - l_1|, v_2 - t|x - l_2|, 0)dx - \frac{1}{2}C(e_1^2 + e_2^2) - A^2\alpha\sigma_u^2 \qquad (2.4)$$

for the values of e_1, e_2, l_1, l_2 determined in proposition 4.1. Expression 2.4 is derived from expression 2.1 by subtracting the costs of having the divisional managers bearing risk. The slope of the incentive scheme, $A \in [0, 1]$, is determined by the maximization of expression 2.4. Expression 2.4 is not generally concave in A since the responsiveness of effort to A increases with A.[20] Although the characterization of the optimal A is involved, it is easy to show (see appendix B) that A decreases with managers' risk aversion, α, and the variance of the noise, σ_u^2. Comparative statics w.r.t. v_0, t and C are ambiguous.

2.5 The Centralized Firm.

In this section we study the centralized firm in which the general office allocates customers to divisions. Each division charges its customers their valuations.

[19] Of course, (price) competition does not take place in equilibrium.

[20] If A is higher, increasing A will induce the managers to exert more extra effort since they will receive the benefit of selling a better product to a larger market.

Since the sets:

$$S_1 = \{x \in [0,1] : v_1 - t|x - l_1| > v_2 - t|x - l_2|\}$$

$$S_2 = \{x \in [0,1] : v_1 - t|x - l_1| = v_2 - t|x - l_2|\}$$

and:

$$S_3 = \{x \in [0,1] : v_1 - t|x - l_1| < v_2 - t|x - l_2|\}$$

are ordered, in the sense that $\forall(x_1, x_2, x_3) \in S_1 \times S_2 \times S_3 \quad x_1 < x_2 < x_3$[21], top management can achieve its maximum payoff by choosing a dividing point $\bar{x} \in [0,1]$ such that customers with locations $< \bar{x}$ deal with division 1 and customers with locations $\geq \bar{x}$ deal with division 2. For simplicity we will consider only these strategies for top management.

We will rule out a class of equilibria that exhibit an extreme lack of coordination. In these equilibria top management allocates all customers to one division, say division 1, "shutting" division 2. The manager of division 1 is then very motivated while the manager of division 2 does not exert any effort and may locate his product close enough to product 1 so that all customers prefer product 1 to product 2.

In proposition 5.1, proven in appendix C, we characterize the remaining pure-strategy equilibria.

Proposition 5.1. *There exists a pure-strategy equilibrium in which locations and effort levels are given by:*

case I: $2v_0/t + A/Ct < 1/2$

$$l_1 = \frac{1}{4} \qquad\qquad l_2 = \frac{3}{4},$$

[21] This is easy to check, using $l_1 \leq l_2$.

$$e_1 = e_2 = v_0 / \left(\frac{Ct}{2A} - 1 \right)$$

and the market is not covered. In any other pure-strategy equilibrium, effort levels and payoffs are the same.

case II: $2v_0/t + A/Ct \geq 1/2$

$$l_1 = \frac{1}{4} \qquad\qquad l_2 = \frac{3}{4},$$

$$e_1 = e_2 = \frac{A}{2C}$$

and the market is covered. This is the unique pure-strategy equilibrium.

The outcome in case I is exactly the same as in proposition 4.1. In case II where divisional products appeal to overlapping sets of customers, divisions choose first best locations for their products. By being unable to deal with the customers of the other division their private gain to locate their product closer to the product to the other division coincides with the social gain. At the same time effort levels in case II are lower than in proposition 4.1. When deciding to exert more effort or not, the manager of a division considers only the increase in valuation of the customers of his own division and not of the customers of the other division.

Top management' payoff is again given by expression 2.4 for the values of e_1, e_2, l_1, l_2 determined in proposition 5.1. In appendix C we determine A and show, as before, that it decreases with α and σ_u^2.

Before leaving this section we discuss our assumption that divisional managers' incentive schemes depend only on the profits of their own division and not on profits of the other division. Consider the extreme case where each manager's incentive scheme depended only on the *sum* of the profits of the two divisions. In this case, a manager

would not attribute any value to his division becoming a more serious competitor of the other division and to profits being transferred from that division to his division. Locations and effort levels in both the decentralized and the centralized firm would then be given by proposition 5.1. Since managers would have to be compensated for bearing more risk, giving incentive schemes that depend on joint profits is dominated by centralizing and giving incentive schemes that depend on individual profits.

2.6 Comparison of Organizational Forms.

In this section we compare the decentralized and the centralized firm. As the analysis of the previous sections showed, decentralization brings more incentives and higher quality products together with inefficient locations for these products. Due to the non-concavity of expression 2.4 w.r.t A in both the decentralized and the centralized case, the analysis of top management's welfare for general values of the parameters v_0, t, C, α and σ_u^2 is somewhat involved. However, the following intuitive results hold:

- If divisional managers are very risk-averse (or the variance of the noise is large) so that the slope of the incentive schemes is small, centralization improves welfare. Indeed, the benefit of increased incentives is small compared to the cost of inefficient locations (provided that even if divisional managers exert zero effort, products, when optimally located, appeal to overlapping sets of customers).[22]

- If divisional managers' risk-aversion is small (or the variance of the noise is small) centralization improves welfare. Indeed, high-powered incentives for improving product quality can be provided at a small cost in the centralized firm. Decentralization thus brings small additional benefits (there may even be over-investment compared to the first best if incentives remain high-powered) at the cost of inefficient locations.

[22]This result holds because divisional managers can choose any location for their products at the same cost. If, instead, some locations are costlier to choose than others and incentive schemes have a small slope, equilibrium locations may not differ much in the two organisational forms.

- If for the optimal incentive scheme in the centralized case, the set of customers who desire both products is small, decentralization improves welfare. Indeed, decentralization will induce (slightly) more effort while locations will deviate little from their first best values so that the costs induced by inefficient locations are very small (in the second order).

These intuitive results are made precise and proven in appendix D.

We finally discuss the feasibility of the two organizational forms. Until now we have assumed that top management knows customers' locations. Since in equilibrium it also knows product quality and product location it can infer which product is better for a given customer. If bargaining between the divisions is costly,[23] top management would strictly prefer to allocate customers to divisions without letting them bargain. We thus see that decentralization may not be feasible.

Decentralization may become feasible if top management faces a cost to intervene, for instance if it does not know customers' locations. Bargaining among divisions may then be a more attractive option than mandating a suboptimal allocation of customers to divisions. We plan to study the feasibility of decentralization when top management is less informed, in a later version of this paper.

[23] We can easily introduce a small cost of bargaining in the model.

2.7 Appendix

2.7.1 Proof of Proposition 3.1.

Writing that the left-derivative of expression 2.1 w.r.t e_1 is ≥ 0 at e_1^* we get:

$$measure(S_1') - Ce_1^* \geq 0 \Rightarrow e_1^* \leq \frac{1}{C} min(1, \frac{v_0 + e_1^*}{t}) \leq \frac{1}{C}$$

where:

$$S_1' = S_1 \cap \{x \in [0,1] : v_1^* - t|x - l_1^*| \geq 0\}$$

and S_1 is the set defined in section 4 (for e_1^*, e_2^*, l_1^*, l_2^*).

Therefore:

$$v_1^* = v_0 + e_1^* \leq v_0 + \frac{1}{C} < \frac{t}{2}$$

(and similarly for v_2^*) by assumption A, and one product even if located at 1/2 cannot cover the whole market. It is thus not optimal to offer two products such that all customers (weakly) prefer one to the other.

We will now fix v_1 and v_2 (smaller that $t/2$) and determine the optimal locations, l_1 and l_2.[24] We will then determine the optimal effort levels. Our previous analysis implies that l_1 and l_2 are such that each product is strictly better than the other for some set of customers

Suppose that products do not overlap i.e.

$$l_1 + \frac{v_1}{t} \leq l_2 - \frac{v_2}{t}. \tag{2.1}$$

Suppose then that $l_1 - v_1/t < 0$. It is then easy to show that the right-derivative

[24]These optimal locations exist since we maximise a continuous function over the compact set $[0, 1/2] \times [1/2, 1]$.

of expression 2.1 w.r.t l_1 ($l_1 < 1/2$ since $l_1 - v_1/t < 0$) is $-tl_1 + v_1 > 0$.[25] Therefore:

$$l_1 - \frac{v_1}{t} \geq 0 \tag{2.2}$$

and similarly:

$$l_2 + \frac{v_2}{t} \leq 1. \tag{2.3}$$

Inequalities 2.1, 2.2 and 2.3 imply that $(v_1 + v_2)/t \leq 1/2$. Expression 2.1 is then equal to:

$$\frac{v_1^2}{t} + \frac{v_2^2}{t} - \frac{1}{2}C(e_1^2 + e_2^2). \tag{2.4}$$

Defining $l_1 = 1/4 + (v_1 - v_2)/2t$ and $l_2 = 1/2 + l_1$[26] and using $(v_1 + v_2)/t \leq 1/2$, inequalities 2.1, 2.2 and 2.3 are satisfied and payoff is the same as when locating products at l_1 and l_2.

Suppose now that products overlap i.e.

$$l_1 + \frac{v_1}{t} > l_2 - \frac{v_2}{t}. \tag{2.5}$$

We denote by x_s the location of the customer who is indifferent between the two products i.e. $x_s \equiv (l_1 + l_2)/2 + (v_1 - v_2)/2t$. ($l_1 < x_s < l_2$)

Suppose then that $l_1 - v_1/t \geq 0$. It is then easy to show that the left-derivative of expression 2.1 w.r.t l_1 ($l_1 > 0$ since $l_1 - v_1/t \geq 0$) is $-v_1 + (x_s - l_1) > 0$.[27] Therefore:

$$l_1 - \frac{v_1}{t} < 0 \tag{2.6}$$

and similarly:

[25] We have to distinguish two cases: $l_1 + v_1/t < l_2 - v_2/t$ and $l_1 + v_1/t = l_2 - v_2/t$.
[26] $l_1 \in [0, 1/2]$ since $|v_1 - v_2|/t < 1/2$.
[27] We again have to distinguish two cases: $l_1 - v_1/t > 0$ and $l_1 - v_1/t = 0$.

$$l_2 + \frac{v_2}{t} > 1. \tag{2.7}$$

Inequalities 2.5, 2.6 and 2.7 imply that $(v_1 + v_2)/t > 1/2$.

The derivative of expression 2.1 w.r.t l_1 is $-tl_1 + t(x_s - l_1)$ (therefore $l_1 > 0$ and $l_1 < 1/2$ and this derivation is meaningful). Similarly the derivative w.r.t l_2 is $-t(l_2 - x_s) + t(1 - l_2)$. Setting these derivatives equal to zero we find $l_1 = 1/4 + (v_1 - v_2)/2t$ ($l_1 \in [0, 1/2]$) and $l_2 = 1/2 + l_1$.

Expression 2.1 is then equal to:

$$\frac{v_1 + v_2}{2} - \frac{t}{8} + \frac{(v_1 - v_2)^2}{2t} - \frac{1}{2}C(e_1^2 + e_2^2). \tag{2.8}$$

To summarize, our above analysis implies that $l_1 = 1/4 + (v_1 - v_2)/2t$ and $l_2 = 1/2 + l_1$ are always optimal locations. In addition if $(v_1 + v_2)/t \leq 1/2$ payoff is given by expression 2.4 while if $(v_1 + v_2)/t > 1/2$ payoff is given by expression 2.8.

Assumption A implies $1/Ct < 1/2$ which in turn implies the concavity of expressions 2.4 and 2.8 w.r.t e_1 and e_2. Since these expressions are symmetric in these variables, the optimal e_1 and e_2 are equal and straightforward maximization gives us proposition 3.1.

2.7.2 Proof of Proposition 4.1.

We will first characterize pure-strategy equilibria and then show that they exist.

Characterization.

Taking the left-derivative of expression 2.3 w.r.t e_1 and writing that the derivative is ≥ 0 for the equilibrium value of e_1 we get:

$$A(measure(S_1') + \frac{1}{2}measure(S_2' \cup S_3')) - Ce_1 \geq 0$$

$$\Rightarrow e_1 \leq \frac{A}{C} min(1, \frac{v_0 + e_1}{t}) \leq \frac{A}{C} \tag{2.1}$$

where:

$$S_k' = S_k \cap \{x \in [0,1] : v_1 - t|x - l_1| \geq 0\}$$

and the S_k's are the sets defined in sections 4 and 5 (for the equilibrium values of e_1, e_2, l_1, l_2).

Therefore one product even if located at 1/2 cannot cover the whole market and it is easy to check that no equilibrium in which all customers (weakly) prefer one product to the other, exists. (Otherwise the manager of the inferior product would change its location so that customers strictly prefer his product.)

Suppose that in equilibrium products do not overlap i.e. inequality 2.1 holds. A similar argument to that given in the proof of proposition 3.1. implies that inequalities 2.2 and 2.3 hold.

The derivative w.r.t e_1 is $A(v_0 + e_1)/t - Ce_1 = 0$.[28] Therefore e_1 is given by its expression in proposition 4.1, case I (as well as e_2). Moreover, combining inequalities 2.1, 2.2 and 2.3 and using the expressions for e_1 and e_2, we find that an equilibrium with non-overlapping products exists only if:

$$\frac{2v_0}{t} + \frac{A}{Ct} \leq 1/2 \tag{2.2}$$

Suppose now that products overlap i.e. inequality 2.5 holds. Defining x_s as before we can show that inequalities 2.6 and 2.7 hold.

The derivative of expression 2.3 w.r.t l_1 is $-tl_1 + t(x_s - l_1) + 1/2(tl_1 + v_1 - tx_s)$ (therefore $l_1 > 0$ and $l_1 < 1/2$ and this derivation is meaningful). Similarly the derivative w.r.t l_2 is $-1/2(tx_s - (tl_2 - v_2)) - t(l_2 - x_s) + t(1 - l_2)$. The derivatives w.r.t e_1 and e_2 are $x_s + 1/2(l_1 + v_1/t - x_s)$ and $1/2(x_s - (l_2 - v_2/t)) + (1 - x_s)$. Setting

[28] We again have to distinguish two cases according to whether $l_1 + v_1/t < l_2 - v_2/t$ or $l_1 + v_1/t = l_2 - v_2/t$.

these derivatives equal to zero we get the expressions in proposition 4.1 (case II).

Inequalities 2.5, 2.6 and 2.7 together with the expressions for e_1 and e_2 imply that an equilibrium with overlapping products exists only if:

$$\frac{2v_0}{t} + \frac{A}{Ct} > 1/2 \tag{2.3}$$

Existence.

Suppose that inequality 2.2 holds so that the only possible equilibria involve non-overlapping products located in the middle of their market segments. We first note that if we choose $l_1 = 1/4$ and $l_2 = 3/4$, inequalities 2.1, 2.2 and 2.3 are satisfied. We now show that a manager, say manager 1, does not get anything by deviating. Consider a deviation $(\tilde{e_1}, \tilde{l_1})$. Equation 2.1 implies that $\tilde{e_1}$ is always smaller than its equilibrium value. Since a manager always prefers to offer a product which does not overlap with the other product (provided that the product is always located in the middle of its market segment and that product quality is kept constant) the choice $(\tilde{e_1}, \tilde{l_1})$ is dominated by $(\tilde{e_1}, 1/4)$ which is obviously dominated by $(e_1, 1/4)$.

Suppose now that inequality 2.3 holds so that the only possible equilibrium is the one described in proposition 4.1 (case II). Let us consider a deviation $(\tilde{e_1}, \tilde{l_1})$ by manager 1. If products do not overlap, using inequalities 2.1, 2.2 and 2.3 we find that $\tilde{v_1}/t \leq 1/4$. This last inequality is not compatible with the first-order condition for e_1 and inequality 2.3. Moreover, it is clearly not optimal for manager 1 to offer a product which all customers prefer to product 2 (since it would then have to cover the whole market), nor to offer a product which all customers find inferior to product 2. Finally, it is easy to check that manager 1's payoff is concave in (e_1, l_1) in the domain of overlap, so that the first-order conditions are sufficient for optimality.

Incentive Schemes.

If inequality 2.2 holds, the derivative of expression 2.1 (where e_1, e_2, l_1, l_2 are given by proposition 4.1, case I), w.r.t A is

$$4(1-A)\frac{\frac{v_0^2}{t}\left(\frac{Ct}{2}\right)^2}{\left(\frac{Ct}{2}-A\right)^3}-2\alpha\sigma_u^2 A \tag{2.4}$$

By contrast if inequality 2.3 holds the derivative is:

$$2\frac{\left(\frac{t}{2}+v_0\right)\left(\frac{Ct}{2}\right)^2}{\left(\frac{3Ct}{2}-A\right)^3}\left[3\left(\frac{3}{2}(1-A)-A(\frac{2v_0}{t}+\frac{1}{Ct}-\frac{1}{2})\right)-(\frac{2v_0}{t}+\frac{A}{Ct}-\frac{1}{2})\right]$$

$$-2\alpha\sigma_u^2 A \tag{2.5}$$

The derivative is not always decreasing in A. (It is easy to see that a sufficient condition for the derivative to be decreasing in A is that $1/Ct < 1/6$.) However, the derivative is decreasing in $\alpha\sigma_u^2$ and Topkis's [11] monotonicity theorem implies that the optimal A is decreasing in $\alpha\sigma_u^2$.

2.7.3 Proof of Proposition 5.1.

Equilibrium.

Suppose that in equilibrium all customers (weakly) prefer the product of one division, say division 1, to the product of the other division. Suppose first that the two products are not the same. Then $\bar{x} \geq l_2$. If $\bar{x} < 1$, and since product 1 does not cover the whole market (by assumption A), the manager of division 2 would have an incentive to increase l_2. Therefore $\bar{x} = 1$. If both products are the same, we can assume that $\bar{x} \geq 1/2$ and the same conclusion holds. Ruling out these equilibria where top management allocates all customers to one division, consider an equilibrium where products do not overlap. We then have:

$$l_1 + \frac{v_1}{t} \leq \bar{x} \leq l_2 - \frac{v_2}{t}. \tag{2.1}$$

Following similar steps as in the proof of proposition 4.1, we can show that equations 2.2 and 2.3 hold and that e_1 and e_2 are given by their expressions in proposition 4.1

(case I). Combining inequalities 2.1, 2.2 and 2.3 and using the expressions for e_1 and e_2, we find that an equilibrium with non-overlapping products exists only if inequality 2.2 holds. Clearly, any values l_1, l_2 and $\bar{x}$ that satisfy inequalities 2.1, 2.2 and

2.3 are equilibrium values, and it is easy to check that setting $l_1 = 1/4$, $l_2 = 3/4$ and $\bar{x} = 1/2$ inequalities 2.1, 2.2 and 2.3 hold.

Suppose now that products overlap i.e. inequality 2.5 holds. Defining x_s as before, we have $x_s = \bar{x}$ and we can show that inequalities 2.6 and 2.7 hold. The derivatives of agents' payoffs w.r.t l_1 and l_2 are $-tl_1 + t(x_s - l_1)$ and $-t(l_2 - x_s) + t(1 - l_2)$ and w.r.t e_1 and e_2 are $Ax_s/t - Ce_1$ and $A(1 - x_s)/t - Ce_2$. Setting these derivatives equal to zero we get the expressions in proposition 5.1 (case II). As before, inequalities 2.5, 2.6 and 2.7 together with the expressions for e_1 and e_2 imply that an equilibrium with overlapping products exists only if inequality 2.3 holds. It is clear that managers cannot do better by deviating from equilibrium.

Incentive schemes.

If inequality 2.2 holds, the derivative of expression 2.1 (where e_1, e_2, l_1, l_2 are given by proposition 5.1, case I), w.r.t A is the same as in expression 2.4.

By contrast if inequality 2.3 holds the derivative is:

$$\frac{1}{2C}(1 - A) - 2\alpha\sigma_u^2 A \tag{2.2}$$

As before, the derivative is decreasing in A if $1/Ct < 1/6$, and the optimal A is always decreasing in $\alpha\sigma_u^2$.

2.7.4 Comparison of Organizational Forms.

We will prove the following proposition:

Proposition 6.1.

Part 1: *Suppose that $v_0/t > 1/4$. Then if $\alpha\sigma_u^2$ is sufficiently large, welfare is higher under centralization.*

Part 2: *Suppose that $2v_0/t + 1/Ct > 1/2$. Then if $\alpha\sigma_u^2$ is sufficiently small, welfare is higher under centralization.*

Part 3: *Suppose that the optimal A under centralization verifies $2v_0/t + A/Ct = 1/2$. Then, welfare is higher under decentralization and under some cases is strictly higher.*

Proof:

Part 1.

Since $v_0/t > 1/4$, inequality 2.3 holds for any value of A. Expression 2.5 shows then that the optimal A can be made arbitrarily small if $\alpha\sigma_u^2$ is sufficiently large. Proposition 4.1 implies then that e_1 and e_2 can be made arbitrarily close to zero and l_1 and $1 - l_2$ arbitrarily close to $1/6 + v_0/3t$ (which is strictly larger than 1/4 since $v_0/t > 1/4$). Since in the centralized firm effort levels are positive, welfare is higher under centralization for $\alpha\sigma_u^2$ sufficiently large.

Part 2.

Welfare under centralization is larger than its value for $A = 1$, which in turn can be made arbitrarily close to welfare under the first best for $\alpha\sigma_u^2$ sufficiently small. By contrast, welfare under decentralization cannot be made arbitrarily close to welfare under the first best for $\alpha\sigma_u^2$ sufficiently small. Indeed, since $2v_0/t + 1/Ct > 1/2$, effort levels are strictly smaller than first best effort levels for A such that $2v_0/t + A/Ct \leq 1/2$, and by continuity for A slightly larger. For A even larger, effort levels may approach first best effort levels but the welfare loss due to inefficient locations is bounded away from zero.

Part 3.

If the optimal A under centralization, say A_c, verifies $2v_0/t + A/Ct = 1/2$, decentralization can only improve welfare since welfare under decentralization for A_c is

equal to the maximum welfare under centralization. To provide parameter values such that welfare is strictly higher in the decentralized firm assume that $1/Ct < 1/6$ so that welfare in the centralized firm is concave in A and assume that the right-derivative of welfare in the centralized firm w.r.t A, i.e. expression 2.2, is zero for A_c.[29] Simple algebra shows that expression 2.5 is always strictly larger than expression 2.2 for values of A such that $2v_0/t + A/Ct = 1/2$ (and for A_c in particular). Therefore for A slightly higher than A_c welfare under decentralization is strictly higher than the maximum welfare under centralization. (Obviously this result is not "knife-edge" and holds even if the optimal A is close to an A that verifies $2v_0/t + A/Ct = 1/2$.)

[29]Welfare in both the centralized and the decentralized firm is not differentiable for values of A such that $2v_0/t + A/Ct = 1/2$ so if such a value is optimal, the right-derivative is not necessarily zero.

Bibliography

[1] Ferguson, Charles H. and C. R. Morris (1993) *"Computer Wars"* Times Books.

[2] Hart, Oliver D. (1983) *"The Market Mechanism as an Incentive Scheme"*, Bell Journal of Economics, Autumn, 14, 2: 366-382.

[3] Holmstrom, Bengt R. and Jean Tirole (1989) *"The Theory of the Firm"*, (in R. Schmalensee and Willig, R., eds.), *Handbook of Industrial Organization, Ch. 2*

[4] Holmstrom, Bengt R. and Jean Tirole (1991) *"Transfer Pricing and Organizational Form"*, Journal of Law, Economics and Organization, 7,2:201-228.

[5] Rey, Patrick and Jean Tirole (1986) *"The Logic of Vertical Restraints"*, American Economic Review, 76:921-939.

[6] Sharfstein, David (1987) *"Product Market Competition and Managerial Slack"*, Rand Journal of Economics, 19,1:147-155.

[7] Shleifer, Andrei (1985) *"A Theory of Yardstick Competition"*, Rand Journal of Economics, 16:319-327.

[8] Sloan, Alfred (1963), *"My Years at General Motors"*, Doubleday.

[9] Smith, Vicky (1992), *"Managing in the Corporate Interest"*, University of California Press.

[10] Tirole, Jean (1988), *"The theory of Industrial Organization"*, MIT press.

[11] Topkis, Donald M. (1978) *"Minimizing a Submodular Function on a Lattice"*, Operations Research 26:305-21.

Chapter 3

Semiparametric Measurement of Environmental Effects.

Abstract

We discuss the application of semiparametric methods to characterize the effect of pollution on housing prices. Semiparametric methods are likely to be superior to parametric and nonparametric alternatives, in hedonic price applications, because they combine flexibility in functional form of moments with simple interpretability of results. The ideas are applied to the Boston Housing Data of Harrison and Rubinfeld, using the semiparametric framework in Rodriguez and Stoker [13]. It appears that previous findings from this data are modified when the partial index model framework is used. The partial index structure allows us to identify the features in the data responsible for misspecification of the previous (parametric) studies that used this data base.

3.1 Introduction.

Policy measures designed to affect the environment involve complex interactions between production and consumption technology and economic behavior. Policy evaluation requires benefit-cost comparisons that are somewhat elusive, because of difficulties in measuring overall benefit or cost components. An economic approach to evaluating environmental conditions makes use of implicit valuations revealed by individual actions, or explicit valuations when variations in condition levels are adequately priced in a market. Such an approach is more suitable when the variation in environmental conditions is observed directly, such as a comparison of housing prices before the discovery of a nearby toxic waste dump, or of the cost structure of a plant before and after the use of an environmentally cleaner technology[1].

The purpose of this paper is to present the results of a semiparametric analysis of air pollution effects on housing prices. The semiparametric analysis involves flexible econometric methods, that are designed to give a depiction as realistic as possible of pollution effects. The focus of our analysis is the Boston Housing data of Harrison and Rubinfeld [9]. This is a well known data set used in a seminal study of the economic evaluation of auto emission abatement policies.

The exposition begins with a general discussion of the economic approach to evaluating environmental effects, including the hedonic approach to explaining differences in housing prices. We then discuss issues in the specification of the equations used to measure pollution effects, introducing our semiparametric method[2]. This sets the stage for the empirical depiction of pollution effects from the Boston Housing Data.

[1]Environmental impacts that have not been observed can also be valued with an economic approach, to the extent that the relevant futures markets accurately price differing environmental situations. Such applications are necessarily somewhat problematic, as many aspects of future situations need to be given explicit attention. A good discussion of the issues of measuring the effects of future global warming is given in the *Economic Report of the President, 1991* [6].

[2]Stock [16] uses another kind of semiparametric approach to measure the average cost associated with toxic waste sites. Palmquist [11] discusses the use of nonparametric price estimates in environmental analysis.

Our results indicate that traditional methods missed substantial nonlinearity in the pollution effect. While a systematic pollution effect can be found for lower status communities, no such effect occurs for lower status communities. The overall impact of changes in pollution levels is considerably smaller than that measured in earlier studies. We give graphical depictions of these features, as well as some interpretations of the findings.

3.2 Issues in the Economic Analysis of Pollution Effects.

To fix ideas, consider evaluating air pollution differences through their impact on housing prices. Suppose that a given house is priced at $p(a)$ when the air pollution level is a. Suppose further that the pollution level is increased by Δa to $a + \Delta a$. Since the house is now less desirable to live in, its price falls to $p(a + \Delta a)$. The economic value (to homeowners) of the pollution change Δa is the change in the housing price:

$$p(a + \Delta a) - p(a) = p_a \Delta a \tag{3.1}$$

where p_a is the per unit value (or cost, since it is negative) of the change in the pollution level:

$$p_a \equiv \frac{p(a + \Delta a) - p(a)}{\Delta a} \tag{3.2}$$

A measure of the "pollution price" p_a, therefore, gives the economic value of the environmental change Δa in air pollution level. To the extent that the change in the level of pollution occurs for many houses in the area, the total impact is the total change in all house prices. The total impact can then be summarized through the values of the pollution price p_a measured for each house in the affected area[3].

[3]One could take the evaluation a further step, relating the change in house value to "constant utility" income differences (as was carried out in the original study of Harrison and Rubinfeld [9]).

While simple in concept, the empirical measurement of pollution prices involves at least two well known blocks of issues. The first one concerns the comparability of "before" and "after" pollution situations as discussed above. In particular, to attribute the entire change in observed house prices to pollution neglects other market features that impinge on demand and supply, an omission that becomes more problematic as the time interval between "before" and "after" increases. For instance, suppose the group of home buyers increases in size, or that tastes change to increase the overall attractiveness of the location of the particular house under consideration. In that case the observed price difference has two components, one based on the pollution difference and the other based on the increase in demand. Alternatively, the supply of housing could have increased in a fashion to depress overall house prices, so that the observed price decrease is again comprised of the pollution effect and the supply effect. Once supply differences are taken into consideration, many outcomes are possible, ranging from the perfectly competitive situation where observed house prices reflect construction cost plus capitalized value of the land (which would reflect the pollution difference), to situations of over or under adjustment over time, due to lags in construction. At any rate, many outcomes are possible, and to properly isolate the pollution effect involves the explicit modeling of the supply response. Unfortunately, no well accepted econometric methods have been developed for this, and so it is standard to measure pollution effects under the assumption of no supply response in the stock of housing, and we do so here[4].

The second block of issues, which are more germane to our study, relate to the case when "before" and "after" assessments are inferred from the comparison of different houses at a given point in time. In particular, suppose one could view two houses as entirely identical (including locational aspects), except that one exists in an area with pollution level a and the other in an area with pollution level $a + \Delta a$. In this context,

Here we just focus on the house price equations, in part because of difficulties in obtaining the original income data required for that latter step.

[4]For an interesting analysis of how the price mechanism matches consumer tastes and product attributes, see Tinbergen [20], and for a useful discussion of competition in supply with differentiated products see Rosen [14].

the price difference between these houses could be attributed to the only thing that differed between them, namely the pollution difference Δa, and the pollution price p_a could be measured as above from the price difference.

This type of comparison is the most widely used method of assessing pollution differences, and is the only available method when the data is on prices of different houses in a single time period. The overall approach of regarding a house as a bundle of attributes, with the pollution level as one, is known as the hedonic price approach[5]. The attributes that rationalize housing prices include number of rooms and other characteristics of the structure, as well as aspects of location, such as lot size and proximity to schools, etc. If we denote all relevant (non-pollution) attributes of a given house as h, then the housing price is determined as $p(a, h)$, and the pollution price is determined as:

$$p_a \equiv \frac{p(a + \Delta a, h) - p(a, h)}{\Delta a} \tag{3.3}$$

This formulation reflects how the actual "house" is held constant through the vector of characteristics (other than air quality) h, with the pollution price reflecting the change in value associated with the pollution change Δa.

The main ingredient to a "hedonic" study of pollution effects is a statistical characterization of the house price function $p = p(a, h)$. Differences in the distributions of housing prices under alternative pollution scenarios can be calculated from this function, as well as the distribution of pollution prices across different kinds of houses. For small changes in pollution levels, pollution prices are given as the partial derivative:

$$p_a = \frac{\partial p(a, h)}{\partial a} \tag{3.4}$$

where again, other housing attributes h are held constant.

[5]Palmquist [12] surveys the use of hedonic methods in studies of environmental policy effects.

3.3 Specification of the Hedonic House Price Equation.

We have spelled out the basic issues above to retain focus on what is being measured in empirical studies of pollution effects. The results depend intrinsically on the hedonic equation representing housing prices, which is characterized statistically. In this section we discuss standard issues in specifying the hedonic equation, in order to introduce nonparametric and semiparametric methods. In the following sections we give the results of our analysis the Boston Housing Data.

The logic of the hedonic approach - house prices determined by housing characteristics - is hardly controversial, provided that all relevant (valued) housing characteristics could be included in the analysis. In practice, however, the observed characteristics are but a fraction of what one could sensibly regard as an exhaustive list. As such, unobserved characteristics are modeled as stochastic elements, and the treatment of such attributes is an integral part of a housing pricing model. Including the pollution level, we denote the observed attributes of a given house by $x = (a, x_0)$, and a stochastic disturbance that represents unobserved attributes by ϵ, so that the full list of housing characteristics is $(a, h) = (a, x_0, \epsilon) = (x, \epsilon)$. The overall object of the statistical analysis is the estimation of the connection between housing price p and observed attributes x, as well as the hedonic price of pollution p_a.

The simplest form of hedonic price equation $p(a, h) = p(x, \epsilon)$ is the standard linear model in price levels:

$$p = x^T\beta + \epsilon = \beta_a a + x_0^T\beta_0 + \epsilon \tag{3.5}$$

where the unobserved characteristics have mean $\alpha = E(\epsilon|x)$. Here, the coefficients (β_a, β_0) are directly interpreted as hedonic prices, with the pollution price above given as $p_a = \beta_a$, and they are constant regardless of the level of pollution or other housing characteristics. The relative importance to housing price of any two house

characteristics is given by the ratio of their respective coefficients. The linear model is dictated if arbitrage exists under perfect competition, when houses can be easily repackaged (or their characteristics unbundled, with an effective market for each characteristic). To the extent that a house representing any "bundle" of housing attributes can be purchased, competition will result in each attribute being priced independently of the others. No arbitrage opportunities in equilibrium necessarily implies the linear model above.

While giving easily interpretable results, it is well known that the linear model does not perform well statistically (in terms of goodness-of-fit) in hedonic price applications, including studies of housing prices. Moreover, the kind of unbundling that dictates a linear model is especially unreasonable for housing (for any small number of characteristics), even approximately. Because different locations of houses are associated with persistent differences in prices, it is natural to question whether a house could be "unbundled" from the location on which it sits. On statistical grounds, the linear model takes all unobserved house price differences to be additive, which implies that unobserved differences in comparable house prices should be the same as one moves from one locale to another, which does not even hold up under casual observation. The implausibility of the (linear) model of perfect competition and the inherent analytical difficulties of modeling a realistic matching process (see Tinbergen [20]) have had as a consequence the fact that the characterization of hedonic price equations has become in practice a purely statistical problem, with the proper form to be decided on the basis of goodness of fit to the data.

The most commonly used hedonic price equation is the log-linear model, that relates log-price variations to characteristics:

$$y \equiv \ln p = x^T\beta + \epsilon \tag{3.6}$$

where again, $E(\epsilon|x) = \alpha$. Here, the coefficients are interpreted as the proportional changes in prices associated to changes in characteristics, holding specific location

features constant. The proportional impact of x_j relative to x_k is summarized compactly as β_j/β_k. As in (3.6), this model dictates that when the location value ϵ is changed, the same proportional configuration of housing prices exists along the lines of the observed characteristics.

It should be noted that because arbitrage is not relied on in the specification of the log-linear model, the characteristic vector "x" can include transformations of basic attribute values: for example, the model does not rely on a unique "price per room", so that polynomial terms in "number of rooms" could be included as part of the statistical investigation. This flexibility makes the log-linear model a considerably richer framework for empirical analysis than the linear model based on competition.

Despite these advantages, the log-linear model has also been recently questioned as to its statistical adequacy in various hedonic price contexts. For instance, it is rejected for several data bases in Berndt, Showalter and Wooldridge [2][6]. The most common parametric method of assessing the log linear model is to estimate a model based on a more general transformation of prices than the logarithm, and check if the logarithm is suggested by the results. In particular, one estimates:

$$y^{(\lambda)} = x^T\beta + \epsilon \tag{3.7}$$

where $y^{(\lambda)}$ is the "Box-Cox" transformation of prices:

$$y^{(\lambda)} = \begin{cases} \frac{p^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln p & \text{if } \lambda = 0 \end{cases} \tag{3.8}$$

and checks whether λ is statistically different from zero. While this estimation involves some delicate econometric issues, the basic point is that a value of λ different from zero rejects the log linear model, giving $y^{(\lambda)}$ as the transformation of prices that is suggested by the data. In this context, the interpretation of the values of the different

[6]The recent surveys by Case, Pollakowsky and Wachter [4], and Smith and Huang [15] each stress how the functional form of hedonic equations remains a major issue in environmental and other types of studies.

coefficients β becomes somewhat obscure, as they translate changes in x to changes in $y^{(\lambda)}$; i.e. the only situations of easy interpretation of β are for $\lambda = 0$ (proportional changes) or $\lambda = 1$ (level changes).

The linear, log-linear and ("Box-Cox") transformation models discussed above constitute the primary parametric models of hedonic prices, such as housing prices, where "parametric" refers to the fact the model is determined by the "few" (finite number of) parameter values, namely β, α and λ. Because they focus attention on a few parameters, each of these approaches can potentially miss important variations in the price data, and therefore give mismeasurements of hedonic prices for attributes, or in our case, for pollution levels. For example, it is easy to verify that the "Box-Cox" transformation model always implies that prices are monotonically increasing in $x^T\beta$ (or characteristics with positive coefficients). Moreover, the connection between price and $x^T\beta$ is globally convex if $\lambda < 1$, i.e. all characteristics x have larger impacts (hedonic prices) on expensive houses than less expensive houses. When $\lambda > 1$, the opposite is true.

To the extent that the empirical investigator is lucky, in that the a priori impositions of the parametric model used adequately hold in the data, there will not be any systematic mismeasurement of the price equation or mismeasurement of the hedonic prices. The only way to be certain of this, however, is to use methods that are more flexible still, i.e. that permit quite general patterns of interaction between house prices and characteristics. Nonparametric and semiparametric methods are designed to permit this kind of "imposition free" determination of the housing equation.

The nonparametric approach to the estimation of the hedonic price equation is based on an unstructured model connecting prices to characteristics. In log form, one could write this "model" as:

$$y = \ln p(x, \epsilon)$$

Statistical analysis can be based on characterizing the mean of log-price for different characteristic values, or the regression function:

$$E(y|x) = E\{p(x,\epsilon)|x\} \equiv m(x) \tag{3.9}$$

Nonparametric methods measure the function $m(x)$ directly, without restricting it to take on a linear or log-linear form as above. The main requirement is that the function $m(x)$ is suitably smooth, so that *small* changes in characteristics x are associated with *small* (continuous, possibly differentiable) changes in mean log-price $m(x)$.

Regression functions can be estimated with any method that permits arbitrarily approximation with a large amount of data. For instance, $m(x)$ could be measured with series approximation, such as polynomials or Fourier series, or with local averages, such as the so-called "nearest neighbour" or kernel regression estimators[7]. In the next section we present kernel estimators, so here we explain the local average the local average method of measuring the log-price equation. This discussion reveals the advantages and the drawbacks to fully nonparametric regression, as well as the role of semiparametric methods.

Local averages measure the mean log-price $m(x)$ by averaging log-price values of houses that have characteristics close to the value x. A useful way to approach these statistical tools is to consider the familiar method of "market analysis" used by realtors for appraising homes. When putting one's home up for sale, the realtor will find "comparable" homes in the locale, look up their selling price, and estimate the "market value" of the home using an average selling price of the comparables. Local averages (such as kernel estimators) implement this idea with data - namely, "comparables" are defined by similar "x" values, and observations with virtually identical characteristic value are given higher weight in the estimation than observations with similar, but not as close, characteristic values.

[7]Nonparametric regression methods are surveyed by Härdle [8].

With a bit more formality, suppose that one has observations for N homes, with the natural log of the price and the observed attributes for the ith house denoted as (y_i, x_i), for $i = 1, ..., N$. A local average estimator of the mean of log-price y given the values of attributes x takes the form

$$\hat{m}(x) = N^{-1} \sum_{i=1}^{n} w_i(x) y_i$$

where $w_i(x)$ denotes the local weight applied to y_i, which is larger for observations with x_i close to the evaluation point x, and smaller (possibly zero) for observations with x_i far from x. The results presented below are based on kernel regression estimators, where the local weight $w_i(x)$ is specified as

$$w_i(x) = \frac{\mathcal{K}\{(x - x_i)/h\}}{\sum_{j=1}^{N} \mathcal{K}\{(x - x_j)/h\}}$$

where $\mathcal{K}(.)$ is a prespecified density function giving the shape of the local weights, and h is the bandwidth parameter that gauges the "proximity" of x_i to x[8]. Figure 1 illustrates local average estimation, depicting it as an enhanced method of curve or surface fitting; or as a flexible method of indicating the structure of the basic data. This amounts to "market analysis" with a systematic method of determining "comparables", given by the formulation of local weights.

The advantage of local averages are obvious - estimating house values using similar houses is a quite natural method of evaluation. The statistical drawbacks are also fairly clear. In particular, if only a few approximately comparable houses can be found, the resulting estimate of the house price will be quite imprecise. This kind of problem is exacerbated when there are many characteristics to take into account in the analysis, wherein it becomes increasingly difficult to find comparable matches to all observations. In a different light, this issue says that a fully general nonparametric

[8]Our empirical analysis used product kernels $\mathcal{K}(u) = \prod_{j=1}^{N} \kappa(u_j)$, with $\kappa(u_j) = 15/16(1 - u_j^2)^2$. Bandwidths are set via generalized cross-validation (c.f. Rodriguez and Stoker [13], and Stoker [17] for details on the estimation).

approach will yield precise log-linear mean values only when there is a great deal of data, so that a fair number of comparable houses can be found for any given observation. This issue is referred to in technical parlance as the "curse of dimensionality" for nonparametric estimation. For our application there are nine characteristics and 506 census tract observations. In this case one faces the additional problem of how to display the results of nonparametric estimation, because the conditional expectation $m(x)$ is a function of nine arguments.

The semiparametric approach combines features of parametric and nonparametric methods, to retain simple interpretability of the results, and to avoid arbitrary mismeasurement by an incorrect parametric formulation of the hedonic price model. In particular, the hedonic price equation has some parametric structure, utilizing parameters to summarize key features of attribute-price connections, but also permits other parts of the hedonic relationship to be measured flexibly, with nonparametric estimators. With reference to the log-linear model, one semiparametric generalization is to assume that the mean log price y is determined by an "index" $x^T\beta$ of characteristics, but that its connection might be an arbitrary nonlinear function G_0:

$$E(y|x) = G_0(x^T\beta) \tag{3.10}$$

This specification is the so-called "single index" model. Here the coefficients β can be estimated directly, and then the (univariate) function G_0 estimated with a kernel regression estimator. With regard to our motivation above, the problem of finding comparable houses is substantially reduced (via "dimension reduction"), because comparability here means similarity in the value of the *index* $x^T\beta$, and the univariate function G_0 can generally be estimated with greater precision than a relation with more than one argument. This specification gives nonparametric treatment to a similar kind of transformation as the "Box-Cox" model above.

While a substantial generalization over the log-linear model, the single index model is still a restrictive specification relative to a fully general regression model, such as

estimated by $\hat{m}(x)$ above. If the single index model is statistically rejected against a general regression, one needs to consider more general semiparametric specifications, and many can be devised. For instance, partial index models employ an "index" specification for certain characteristics, but a fully general structure for other variables. For instance, suppose that the impact of pollution level was not adequately represented by the "index" formulation (3.10), then a partial index model could be constructed as:

$$E(y|x) = G_1(a, x_0^T\beta_0) \tag{3.11}$$

In this model, the coefficients β_0 are estimated, and the two-dimensional function G_1 is estimated nonparametrically. "Comparables" are determined via close values of a and $x_0^T\beta_0$, and the impact of pollution a on mean log house price is not necessarily connected to the impact of any of the other characteristics.

In a related paper (Rodriguez and Stoker [13]), we devise a testing procedure for assessing what degree of functional structure is called for in data sets on log prices and characteristics, along the lines above. For the Boston Housing Data, we found that the log-linear and single-index models gave statistically equivalent depictions of the data, but both were rejected against a general kernel regression estimator. We concluded that a certain partial index model gave a statistically adequate description, in that it was not rejected against nonparametric regression. We now turn to a discussion of the specifics of the data and model, and a description of the basic results on log housing prices and the hedonic price of pollution. The reader is referred to the above reference for the details of our testing procedure. Our aim in this paper is to illustrate how one can interpret our final model, in terms of house prices - attribute relationships as well as the associate hedonic price of pollution.

3.4 Environmental Effects in the Boston Housing Data.

3.4.1 Some Preliminaries.

The Boston Housing Data consists of 506 observations on the average value of housing in census tracts in the Boston area in 1970. This data was first used in the analysis of pollution abatement policy by Harrison and Rubinfeld [9] (hereafter HR), and [10], and the structure of their hedonic house price equation has received considerable attention since then. For instance, Belsley, Kuh and Welsch [1] (BKW hereafter) analyze the basic log-linear equation for robustness, including the identification of various influential observations and other outliers. The variables in the data set are listed in Table 1, where we have focused our analysis on the nine predictors found significant in the Harrison-Rubinfeld and Belsley-Kuh-Welsch work[9]. Likewise, we have retained the transformations of the basic observed variables as used by these authors. The ordinary lest squares (OLS) estimates of the coefficients of the log-linear model are given in Table 2.

The coefficient estimates for the single index model (3.10) are also given in Table 2 under the heading of "average derivatives". These coefficients are measures of the average of the effects (slopes) of the attributes on log-prices, treating the regression as arbitrarily general. If $m(x) \equiv E(y|x)$ denotes the true statistical relationship, then the average derivatives are $\delta \equiv E(\partial m/\partial x)$, and these parameters can be measured more precisely than the attribute specific effects $\partial m/\partial x$ for any given value x of attributes. Stoker [17] gives a lengthy treatment of average derivatives, their estimation,

[9]The variable $LSTAT$ is very important in the analysis that follows. It is the "proportion of residents of lower status", which is defined as the percentage of adults who are laborers and do not have a high school education.

Table 3.1: VARIABLE SPECIFICATION

$y = \ln p$	LMV	log of home value
x_1	NOXSQ	nitrogen oxide concentration
x_2	CRIM	crime rate
x_3	RMSQ	number of rooms squared
x_4	DIS	distance to employment centers
x_5	RAD	accesibility to radial highways
x_6	TAX	tax rate
x_7	PTRATIO	pupil teacher ratio
x_8	B	$(B_k - .63)^2$, where B_k is proportion of black residents in neighborhood
x_9	LSTAT	log of proportion of residents of lower status

and how they provide consistent estimates of the coefficient parameters of single and partial index models[10]. If the log-linear model were, in fact, statistically adequate, the average derivatives would measure the same values as the OLS coefficients.

The average derivative estimates have the same qualitative pattern as the OLS coefficients, except for the race effect, that has a negligible average derivative estimate[11]. The average derivatives indicate weaker effects than OLS for pollution, distance to employment and access to radial highways, pupil-teacher ratio and lower status, and stronger effects for crime rate, number of rooms, and taxes. Despite these differences in magnitude, it is not possible to reject the hypothesis that the OLS coefficients and the average derivatives are measuring the same values, as indicated by the Wald statistic at the bottom of Table 2. In other words, from just looking at coefficient estimates, there are no grounds for rejecting the basic log-linear model.

[10]We have used IV estimators of the average derivatives, that use score estimates as instruments for linear coefficients of y on x.

[11]The curious transformation used by HR and BKW for the race percentage means that a positive coefficient indicates a negative impact of the presence of minorities; $B = (B_k - .63)^2 = B_k^2 - 1.26B_k + (.63)^2$, so that for small proportions of black residents B_k, B will vary as $-1.26B_k$. Consequently, the OLS coefficient indicates a substantial negative race effect, and the average derivative indicates no effect of race.

Table 3.2: COEFFICIENT ESTIMATES

$Dependent Variable, y = LMV(\ln p)$			
		Average Derivatives $\hat{\delta}$	OLS $\hat{\beta}$
x_1	$NOXSQ$	-.00340	-.00600
		(.0035)	(.0011)
x_2	$CRIM$	-.02560	-.01200
		(.0056)	(.0012)
x_3	$RMSQ$	.01060	.00680
		(.0025)	(.0012)
x_4	DIS	-.07460	-.19950
		(.0504)	(.0265)
x_5	RAD	.06690	.09770
		(.0468)	(.0183)
x_6	TAX	-.00090	-.00450
		(.0003)	(.0001)
x_7	$PTRATIO$	-.01750	-.03200
		(.0152)	(.0047)
x_8	B	-.05260	.37700
		(7.5140)	(.1033)
x_9	$LSTAT$	-.25830	-.36500
		(.0370)	(.0225)
(Standard errors in parenthesis)			

Wald Test of $\delta = \beta$ W = 13.44	$Prob\{\chi^2(9) > 13.44\}$ = .143

However, Rodriguez and Stoker [13] note that the log-linear model is indeed rejected against a general regression model, so that apart from the coefficients, there are systematic departures from the log-linear model in the data. The same is true for the strict single index model (with average derivative estimates as coefficients). So that using a single index to summarize the effects of all attributes likewise misses some nonlinear structure. Our testing procedure concluded with a partial index model that omitted the pollution variable, and the lower status variable, from the index summarizing the remaining attributes. In particular, our procedure failed to reject such a partial index model against general regression, so that the non-linearity not accounted for by the log-linear model arises from the treatment of pollution and lower status effects. Since the pollution impact is a primary concern, we focus on the hedonic price of pollution after describing the basic model below.

The partial index model that satisfies our criteria gives the mean log housing price as:

$$E(y|x) = \hat{G}_2(NOXSQ, LSTAT, INDEX) \tag{3.1}$$

where the index variable is constructed from the remaining seven predictors as:

$$INDEX = (-.0256\, CRIM) + (.0106\, RMSQ) + (-.0746\, DIS) +$$

$$(.0669\, RAD) + (-.0009\, TAX) + (-.0175\, PTRATIO) + (-.0526\, B)$$

The $INDEX$ variable can be regarded as a linear representation of log house prices, omitting the influence of air quality ($NOXSQ$) and the lower status position of the communities. The function $\hat{G}_2(.)$ gives a nonparametric description of the impact of pollution and lower status, with remaining housing attributes controlled for via the index representation. We now give a graphical description of the pollution impacts.

3.4.2 Graphical Analysis of the Boston Housing Data.

The estimated log-price function $\hat{G}_2(.)$ is a function of three variables, so we cannot depict it entirely on one diagram. Since our focus is on the pollution effect, we first give three-dimensional diagrams of mean log-price $\hat{G}_2(.)$ over values of pollution and lower status, for various values of the index of the other housing attributes. We then give diagrams of $\hat{G}_2(.)$ over values of pollution and the index, for various values of lower status. For interpretability, we detransform the pollution-squared and log-lower status variables as used in estimation, plotting $\hat{G}_2(.)$ over values of pollution and lower status[12]. As discussed above, $\hat{G}_2(.)$ represents the mean of log house prices for given values of its arguments: twists, bumps and other kinds of curvature indicate nonlinearity in the hedonic house price relationship[13].

Figure 2 graphs the relation between log house prices (in the vertical axis) and pollution and lower status, with $INDEX$ set to its mean value. For lower status communities we see that the pollution effect is negative as expected, or that lower house prices are associated with higher pollution levels. However, for higher status (low $LSTAT$) communities, the opposite is true, namely that higher pollution levels are associated with higher house prices. This counterintuitive finding indicates immediately how the log-linear model fails to adequately account for the empirical patterns of higher status communities.

Before interpreting this feature, we should consider the possibility that it might be explained by the presence there of just a few high price - high pollution - high status communities observed at the mean index level; in other words, there may

[12]It is worthwhile noting that since $\hat{G}_2(.)$ is a nonparametric estimate with $NOXSQ$ and $LSTAT$ as free arguments. The (invertible) transformations used in their construction do not affect the estimate of mean log house price. In other words, $\hat{G}_2(.)$ is sufficiently flexible to "undo" the transformations if that was dictated by the data.

[13]For comparison, it should be noted that the log-linear model would be depicted as smooth planar surfaces in the diagrams that follow. It would not result in flat planes per se, as our detransformation would add some minimal curvature due to removing the "square" from $NOXSQ$ and the "log" from $LSTAT$. This comparison is also not exact, because of the difference between $INDEX$ and the weighted sum of the other attributes (OLS coefficients) of the log-linear model.

be a few outliers that arise from an unfortunate choice of plotting at the average value for $INDEX$. To consider this, Figure 3 gives an analogous plot with the level of index set to a higher value (mean plus 1.5 standard deviation) and Figure 4 gives an analogous plot with the index level set to a lower value (mean minus 1.5 standard deviation). These figures illustrate the same phenomena, namely a negative pollution effect for lower status communities and a somewhat positive effect for higher status communities. As such, the misspecification of pollution effects for higher status communities is a robust feature regardless of the overall level of house prices (omitting pollution and status effects).

There are some subtle differences in Figures 2, 3 and 4 worth noting. Figure 4 depicts a relatively smaller upturn in prices for pollution values in high status communities than Figures 2 and 3. This is quite natural as the level of house prices can be regarded as an alternative measure of a "status" phenomenon, so that the counter-intuitive pollution effect with high status would be weaker in low price communities. These subtle differences give further indication that the generality of model (3.1) is important for an adequate data description.

An alternative depiction of these results is given from the implied hedonic price of pollution for the model 3.1 versus the log-linear model. For a given value of the observed attributes x, the hedonic price of pollution is obtained by differentiating 3.1, giving:

$$\hat{p}_{NOX} = \hat{G}_2(x)\frac{\partial \hat{G}_2(x)}{\partial NOXSQ}\,\hat{\delta}_{NSQ}\,NOXSQ \tag{3.2}$$

where $\hat{\delta}_{NSQ}$ is the estimated average derivative for $NOXSQ$ from Table 2. For the log-linear model, the implied hedonic price of pollution is:

$$\tilde{p}_{NOX} = \hat{L}(x)\,\hat{\beta}_{NSQ}NOXSQ$$

where $\hat{L}(x)$ is the fitted value from the log-linear model, and $\hat{\beta}_{NSQ}$ is the OLS coefficient of $NOXSQ$.

Figure 5 depicts the hedonic price as implied by model (3.1) (for $INDEX$ set to its mean value). It shows a smoothly increasing cost (negative price) of higher pollution levels in lower status communities. As the status of the community is increased, the prices become positive, as we would expect from the relations in Figures 2, 3 and 4. For further illustration of this, Figures 6, 7 and 8 illustrate the pollution price as a function of pollution level, and include the pollution price that is implied by the log-linear model. Here $INDEX$ is set to its mean value, and Figures 6, 7, 8 depict the pollution price for high, middle and low status communities respectively[14]. These figures illustrate more clearly the sensitivity of the pollution effect to the status of the community, as well as the substantial differences between semiparametric and log-linear estimates of the hedonic price of pollution[15]. Reliance on a log linear model amounts to asserting serious impositions on the pollution price that are in conflict with the observed data patterns. The hedonic price of pollution has a considerably more complicated structure than that given by the standard log linear model.

The same basic lesson arises when we consider the interrelationship between log house price, pollution level and the index of the other housing attributes. Figure 9 depicts this structure at the mean value of lower status. While somewhat accentuated by the orientation of the graph, the increase in housing prices at quite high pollution levels is evident, especially for high values of the index.

A more vivid depiction of the intrinsic nonlinearity of the basic hedonic relationship is found by considering the hedonic price of pollution as compared to the index, at different levels of community status. Figures 10, 11 and 12 depict the hedonic price as a function of pollution and the index, with the level of $STATUS$ set at high, middle

[14] "Middle status" is defined by setting $LSTAT$ equal to its mean value , "High Status" by setting $LSTAT$ to its mean value minus 1.5 its standard deviation, and "Low Status" by setting $LSTAT$ to its mean value plus 1.5 its standard deviation.

[15] Stoker [19] argues how kernel regression estimators can under estimate derivatives. We have not studied this phenomenon here, but it would not affect the qualitative features of these diagrams. It could nevertheless imply an exacerbation of the basic differences noted. This issue of derivative bias is not relevant for our method of estimating the average derivatives (Stoker [18]), and therefore it is not responsible for the difference between the average derivative of $NOXSQ$ and its OLS regression coefficient.

and low status levels respectively. Most evident is how the lower status communities show a strongly negative effect of increases in pollution, on housing prices, especially for high pollution ranges. This effect is mollified for lower status communities, and apparently reversed for middle status communities. While yet just another method of isolating the substantial nonlinearity in the log house price - pollution relationship, the differences between the estimated prices over status ranges are quite striking. Finally, for comparison, Figure 13 presents the hedonic prices implied by the log-linear model at the mean status level. Noting differences in the plotting range of hedonic prices, one can see how the log linear model predicts smooth pollution impacts in a fairly narrow range as compared to those seen in Figures 10, 11 and 12.

3.4.3 Interpretation.

One of the more powerful features of the semiparametric methods we have used is the ability to clearly depict the structure of the data, giving a visual indication as to where a standard model can fail. In this regard it is worth noting some findings from previous analysis of this data which are consistent with our findings above.

First, the original study of Harrison and Rubinfeld [9] found that the estimates of pollution effects under their parametric log-linear specification where not robust to perturbations of the basic model. The estimates of aggregate benefits of air quality improvements implied by their model could be decreased by 60%, when alternative specifications were considered (HR [9] p. 78). As shown in Table 2, the average derivative estimate for $NOXSQ$ is considerably below the OLS estimate, and considerably less precise , reflecting this feature. The substantial nonlinearities evidenced from our graphical depiction of model (3.1) would lead to very sensitive estimates based on varying specifications of a log-linear model. In essence, different log-linear specifications amount to different weighting of the observed pollution effect over different ranges of the data. Therefore, the substantial differences in the hedonic price of pollution noted above would lead to wide variation in summary estimates such as OLS coefficients from differing specifications.

The analysis of Belsley, Kuh and Welsch [1] carried out careful diagnostic analysis on the residuals of the log-linear model, and found strong evidence on the presence of outliers, or strongly influential observations. In particular they note a correlation between the size of the residuals and the census tract to which they belong, concluding that misspecification has occurred with respect to a factor that is correlated with the geographic district assignments of the data. More specifically, they note

> The influential data tend to be quite heavily concentrated in a few neighborhoods and these are, for the most part, the central city of Boston, which leads us to believe that the housing price equation is not as well specified as it might be (BKW [1] p.243).

These authors did not pursue model variations that isolate these influential observations.

Our results find systematic departures from the model in terms of nonlinearity of the pollution - status relationship. While we have not included Boston as an explanatory factor, we have noted how the pollution-status effect is basically robust to the levels of the other variables in the data. Since there are high status census tracts other than those in Boston, we consider what further interpretation can be given to the inadequacy of the parametric model beyond the fact that "Boston may be special".

The covariation of pollution and housing prices could arise because of genuine disutility associated with breathing poorer air, but it could also be the outcome of various locational features that are associated with air quality. For example, the air quality in a community could be poor because of the presence of high density traffic areas. Another example concerns proximity to factories or other pollution-emitting sources. Moreover the possibility that pollution is a proxy for other locational effects is consistent with our semiparametric findings, The fact is that lower status communities are likely to be those in which high density traffic areas or industrial sources are to be found. Consequently, the lower status communities would display a systematic

negative pollution effect as we have found. Alternatively, if the higher status communities were relatively free of traffic and other problems, the measured pollution level may not be systematically associated with lower prices, also as we have found.

The possibility of pollution acting as a proxy for other kinds of locational effects also places a proviso on the applications of hedonic results to the assessment of pollution abatement policies. Suppose, for instance, that laws were passed that severely limited automobile emissions. Over a period of time, housing prices in previously polluted areas would adjust upwards to the extent that improved air quality improved living conditions. This impact would likely be most relevant for communities where the pollution effect were highly associated with traffic and congested downtown areas. But such an effect would not be so pronounced for areas where pollution arose from an industrial site. Moreover, even in areas where the major polluters are automobiles, the emission laws would not mollify other factors associated with traffic density that impact with housing prices. For instance, the noise level of close busy roadways, or issues of child safety near busy roadways would go unchanged, and their implicit valuation in observed housing prices would remain. As in many contexts where there is the potential for important omitted variables, the measured pollution effect may overstate the true benefits of environmental policies.

3.5 Conclusion.

The purpose of this paper is to introduce many of the ideas of semiparametric modeling in the context of the measurement of the implicit value of changes in air quality. We have illustrated these ideas with an application to the Boston Housing Data, where we have graphically depicted the inadequacy of a standard log-linear model of housing prices. While our results are reflective of findings of previous analysis of this data, we have isolated more systematic features of where the pollution effect is stable and where it is not. In particular a quite different pollution - house price structure exists for low status housing as for high status housing areas.

Our application of flexible nonparametric methods involves more complicated issues and computational methods than standard log-linear modelling. Moreover, since our application involved nine predictor variables, there was the possibility that no simplified semiparametric model would give a statistically adequate fit. In such a case, we would have been hard pressed to give a very conclusive description of the true empirical relationship, because of the difficulties describing a nine dimensional function.

Our discovery of a statistically adequate partial index model permitted graphical analysis and interpretation, but even our basic model 3.1 involved estimation of a function of three arguments ($NOXSQ, LSTAT$ and $INDEX$). The analysis of our final model was considerably more complicated than familiar analysis of log-linear regression estimates. With this in mind, one might think that it is in some ways better to use a log-linear model, relying on the estimated OLS coefficient of a pollution variable as an "average effect" applicable across the entire data sample.

We would argue strongly against such an approach, which amounts to ignoring model specification issues on the grounds of expediency. It is clear that "back of the envelope" policy evaluation is aided greatly by the use of a single summary effect, but there is not systematic reason to believe that OLS estiamtion of a log-linear model will give anything like an *adequate* summary effect. For instance, in Table 2 the OLS coefficient of $NOXSQ$ was almost twice as large as our measure of the average derivative, or average effect of $NOXSQ$. Moreover, the apparent precision of the OLS coefficient masks the fact that the empirical effect of pollution varies quite widely across the data, with a systematic negative effect for low status communities combined with a rather unstable measured effect for high status communities. If one is to use a summary statistic of the basic effect, it is important to utilize statistics whose connections to the empirical effects are well understood. One possibility, among many, is the average derivative, or average effect across the sample. However, unless the data relationship is demonstrably consistent with a log-linear model, there is no clear way to trace the connection between OLS coefficient estimates and the underlying

empirical effects.

As such, a major conclusion of our empirical analysis is that reliance on a log-linear framework, even when augmented with the standard tools of inference for linear models, can miss systematic nonlinear structure in the basic data. The only way to allow for a more complete range of possibilities is to adopt a framework that is sufficiently flexible to account for such a *complete* range, and that provides methods of interpretation of the basic findings. Our use of kernel estimators permitted examination of the nonlinear house price - pollution relation on diagrams, which make clear which features are not captured by the log-linear model. While our analysis is definitely more complicated than what is available from familiar, more standard methods, the result is a much richer understanding of the house price - pollution relationship in the Boston Housing Data.

Bibliography

[1] Belsley, D.A., E. Kuh and R.E. Welsch. (1980) "Regression Diagnostics: Identifying Influential Data and Sources of Collinearity." *New York, Wiley.*

[2] Berndt, E. R., M.H. Showalter and J.M. Wooldridge. "A Theoretical and Empirical Investigation of the Box-Cox Model and a Nonlinear, Least Squares Alternative." *Sloan School of Management,* Working Paper # 3187-90.

[3] Brown, J.N. and H. Rosen. (1982) "On the estimation of Structural Hedonic Price Models." *Econometrica*, 50, 3:765-768.

[4] Case, B., H.O. Pollakowsky and S.M.Wachter. (1991) "On Choosing Among House Price Index Methodologies." August, draft, *Joint Center for Housing Studies, Harvard University.*

[5] Cropper, M.L., L.D. Deck and K.E. McConell. (1988) "On the Choice of Functional Form of Hedonic Price Functions." *The Review of Economics and Statistics*, 70:668-675.

[6] "Economic Report of the President". (1991) February, *United States Prining Office*, Washington.

[7] Graves, P., J.C. Murdoch, M.A. Thayer and D. Waldman. (1988) "The Robustness of Hedonic Price Estimation: Urban Air Quality." *Land Economics*, 64, 3:220-233.

[8] Härdle, W. (1990) "Applied Nonparametric Regression", *Cambridge University Press*

[9] Harrison, D. and D.L. Rubinfeld. "Hedonic Housing Prices and the Demand for Clean Air." *Journal of Environmental Economics and Management*, 5:81-102.

[10] Harrison, D. and D.L. Rubinfeld. "The Distribution of Benefits from Improvements in Urban Air Quality." *Journal of Environmental Economics and Management*, 5:313-332.

[11] Palmquist, R. (1988) "Welfare Measurement for Environmental Improvements Using the Hedonic Model: the Case of Nonparametric Marginal Prices." *Journal of Environmental Economics and Management*, 15:297-312.

[12] Palmquist, R. (1988) "Hedonic Methods" in "Measuring the Demand for Environmental Improvements," J.B. Braden C.D. Kolstad, eds. *Institute for Environmental Studies, University of Illinois at Urbana.*

[13] Rodriguez, D. and T.M. Stoker. (1992) "A Regression Test of Semiparametric Index Model Specification." *Center for Energy and Environmental Policy*, Working Paper # 92-006.

[14] Rosen, S. (1974) "Hedonic Prices and Implicit Markets: Product Differentiation in Perfect Competition." *Journal of political Economy*, 82:34-55.

[15] Smith, V.K. and J.C. Huang. (1991) "Can Hedonic Models Value Air Quality?: A Meta Analysis." November, draft, *North Carolina State University.*

[16] Stock, J.H. (1989) "Nonparametric Policy Analysis." *Journal of the American Statistical Association*, 84:567-575.

[17] Stoker, T.M. (1992) "Lectures on Semiparametric Econometrics". *CORE Foundation, Louvain-laNeuve, Belgium.*

[18] Stoker, T.M. (1993) "Smoothing Bias in Density Derivative Estimation." *Sloan School of Management*, Working Paper # 3336-91, forthcoming: *Journal of the American Statistical Association.*

[19] Stoker, T.M. (1993) "Smoothing Bias in the Measurement of Marginal Effects." January, *Sloan School of Business*, Working Paper # 3522-93.

[20] Tinbergen, J. (1956) "On the Theory of Income Distribution." *Weltwirtschaftliche Archiv*, 77:155-173.

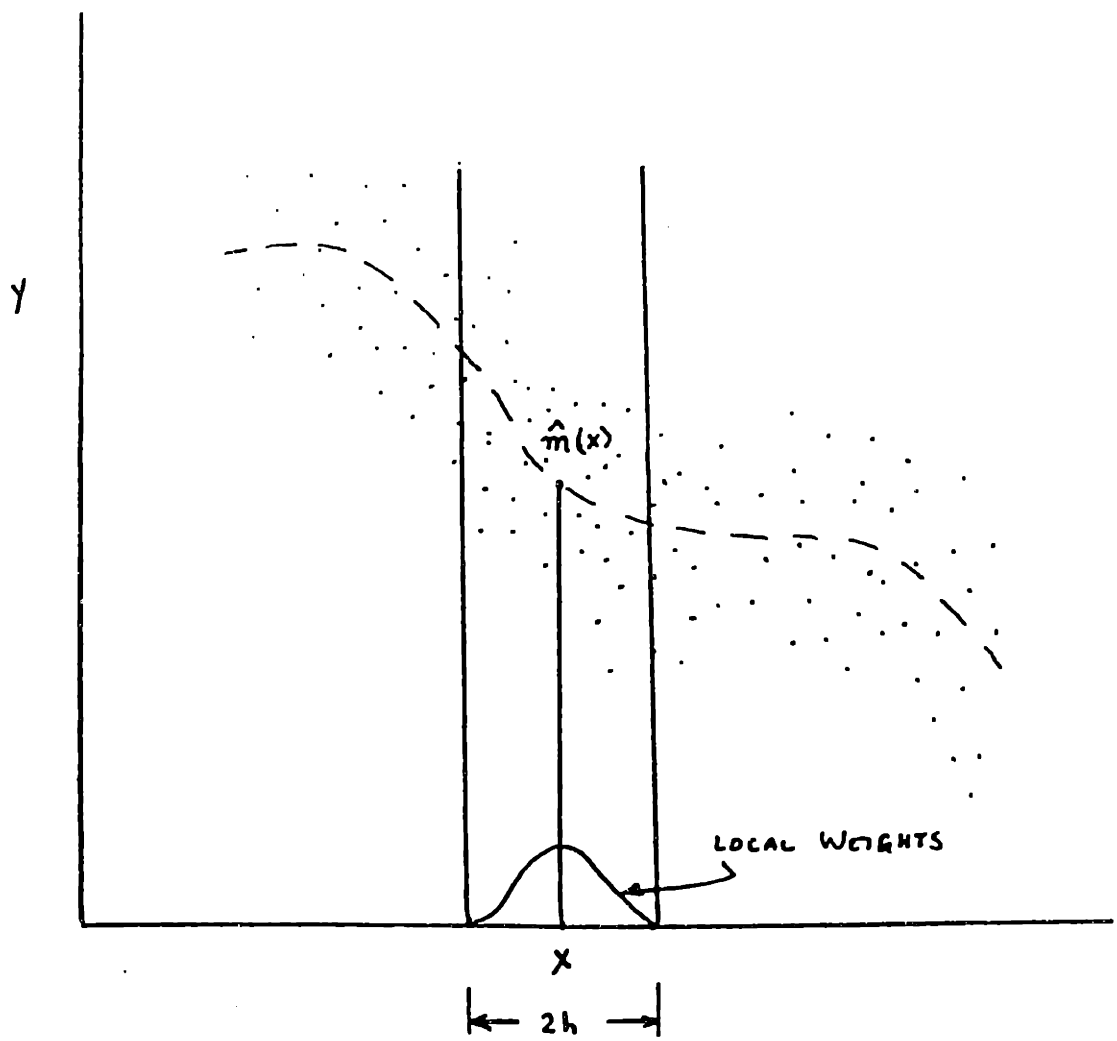

Figure 3-1: Regression Estimation Via Local Averages

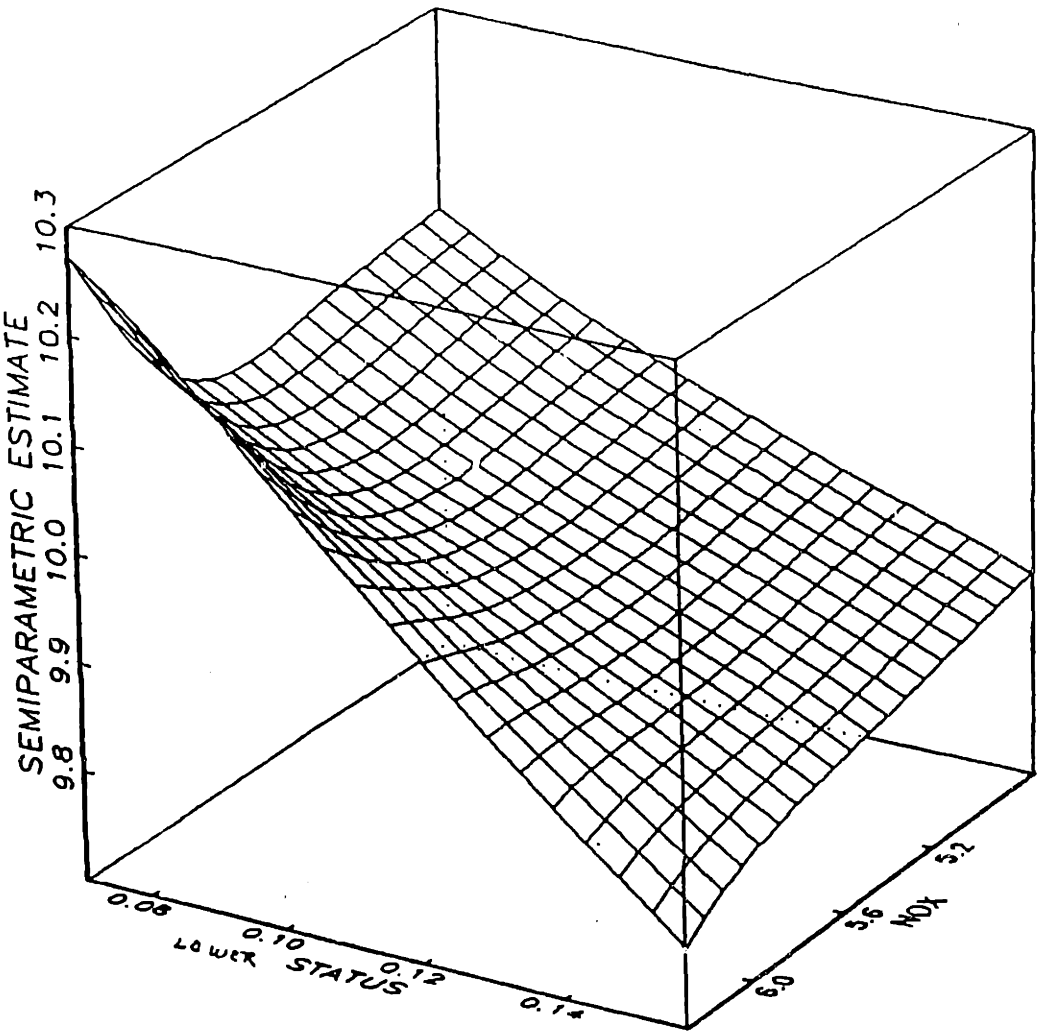

Figure 3-2: Log House Price as Related to Pollution and Lower Status. (Middle INDEX Value).

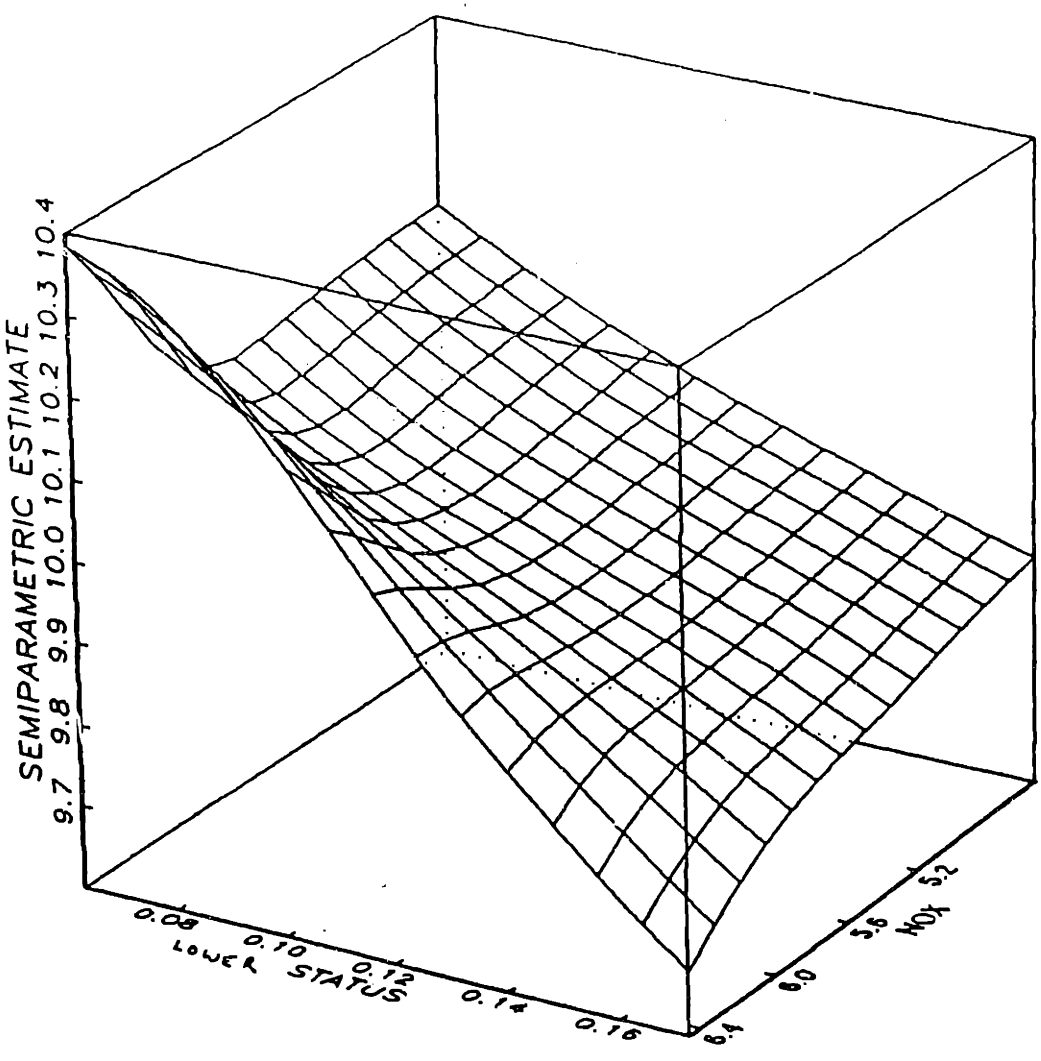

Figure 3-3: Log House Price as Related to Pollution and Lower Status. (High INDEX Value).

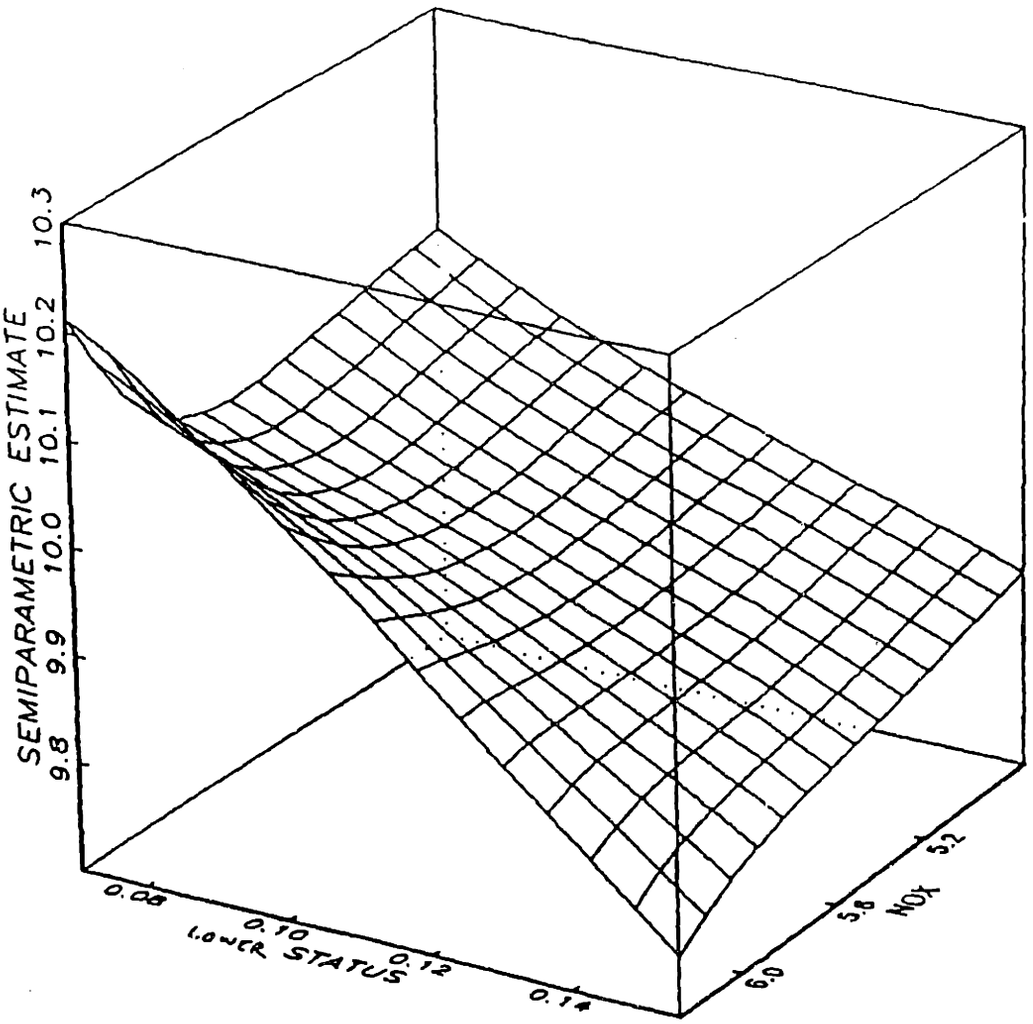

Figure 3-4: Log House Price as Related to Pollution and Lower Status. (Low INDEX Value).

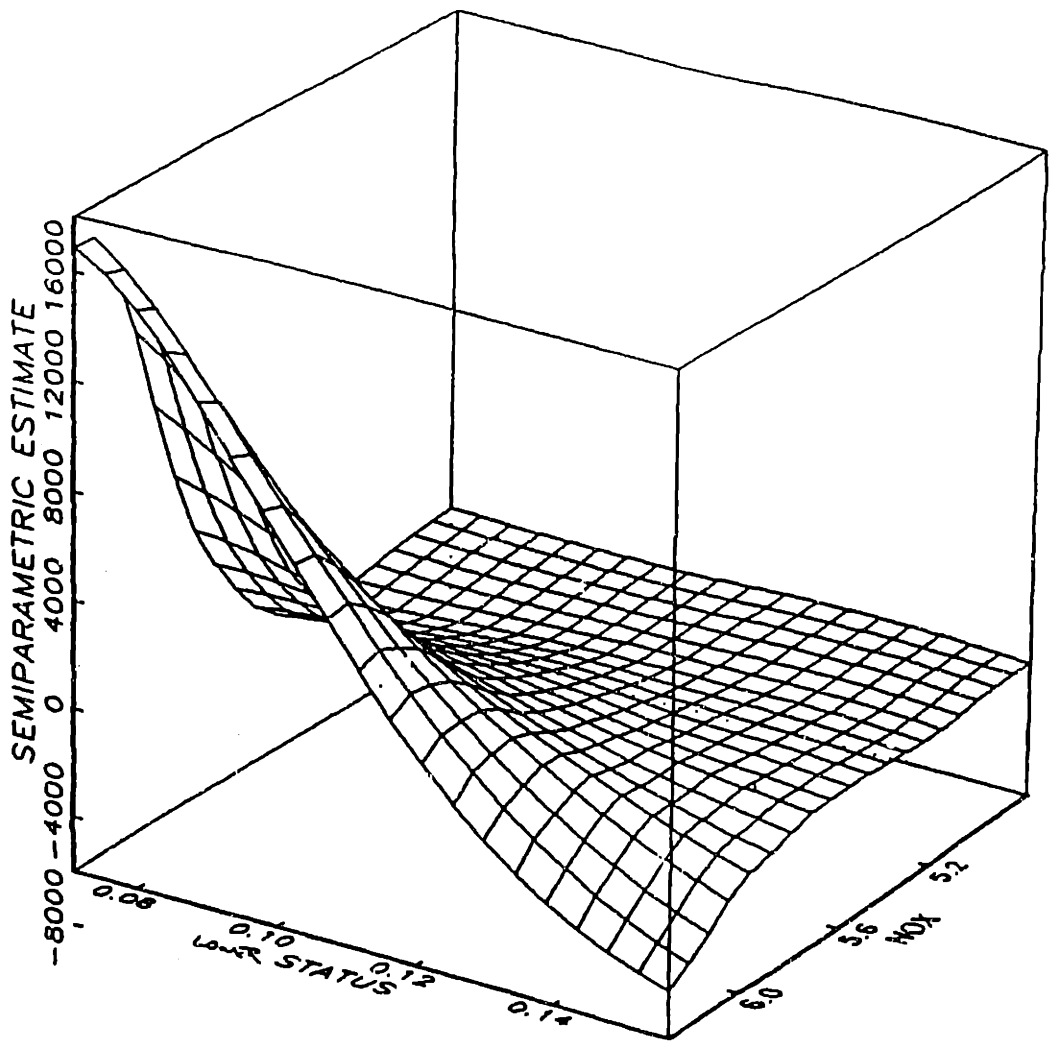

Figure 3-5: Hedonic Price of Pollution. (Middle Index Value).

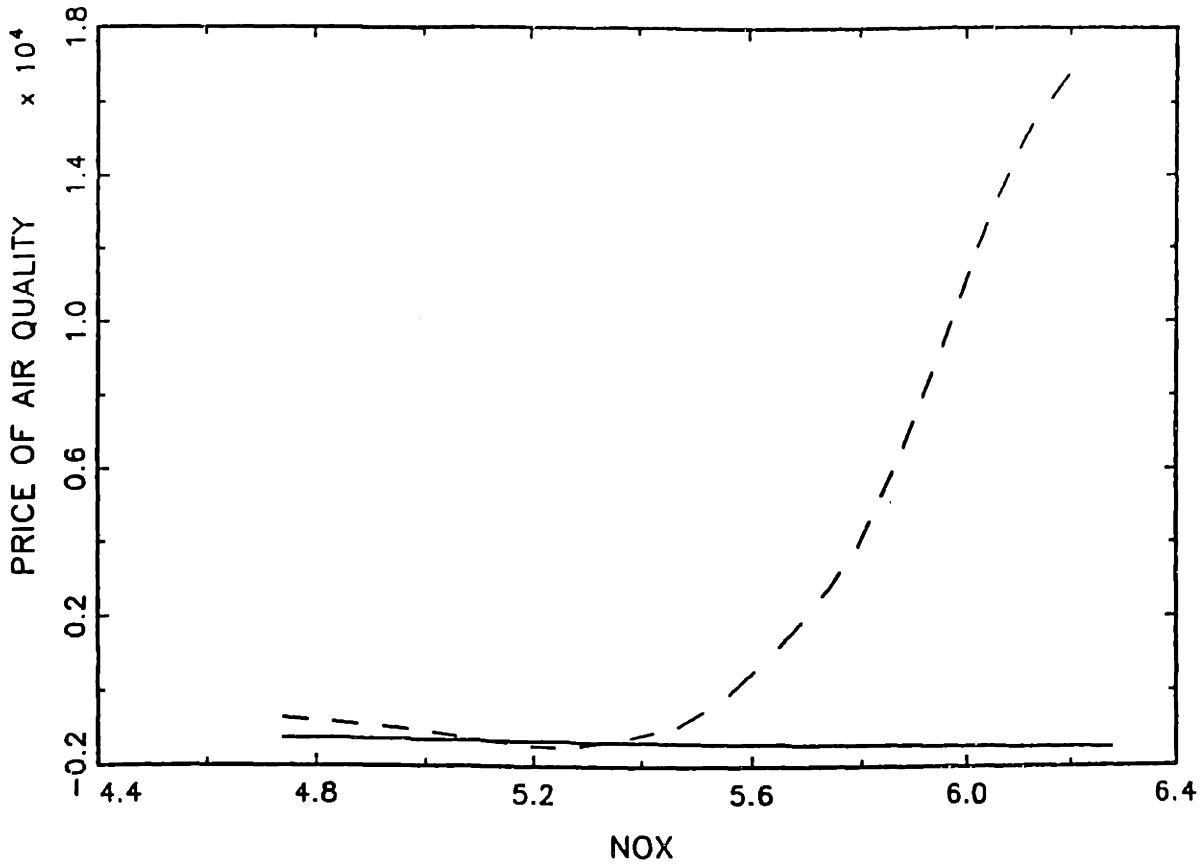

Figure 3-6: Hedonic Price of Pollution: High Status Communities. Middle Index Value; (Continuous line: Log Linear Model; Broken Line: Semiparametric Model)

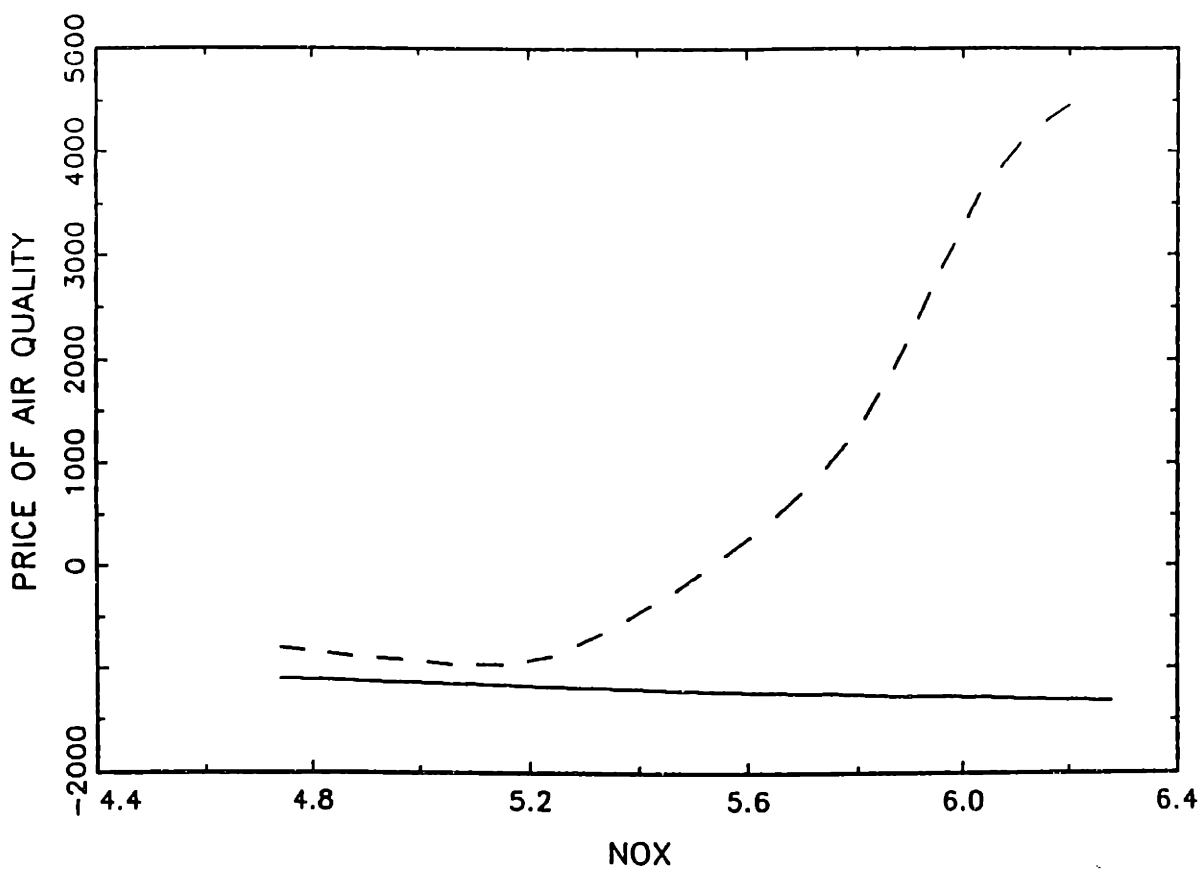

Figure 3-7: Hedonic Price of Pollution: Middle Status Communities. Middle Index Value; (Continuous line: Log Linear Model; Broken Line: Semiparametric Model)

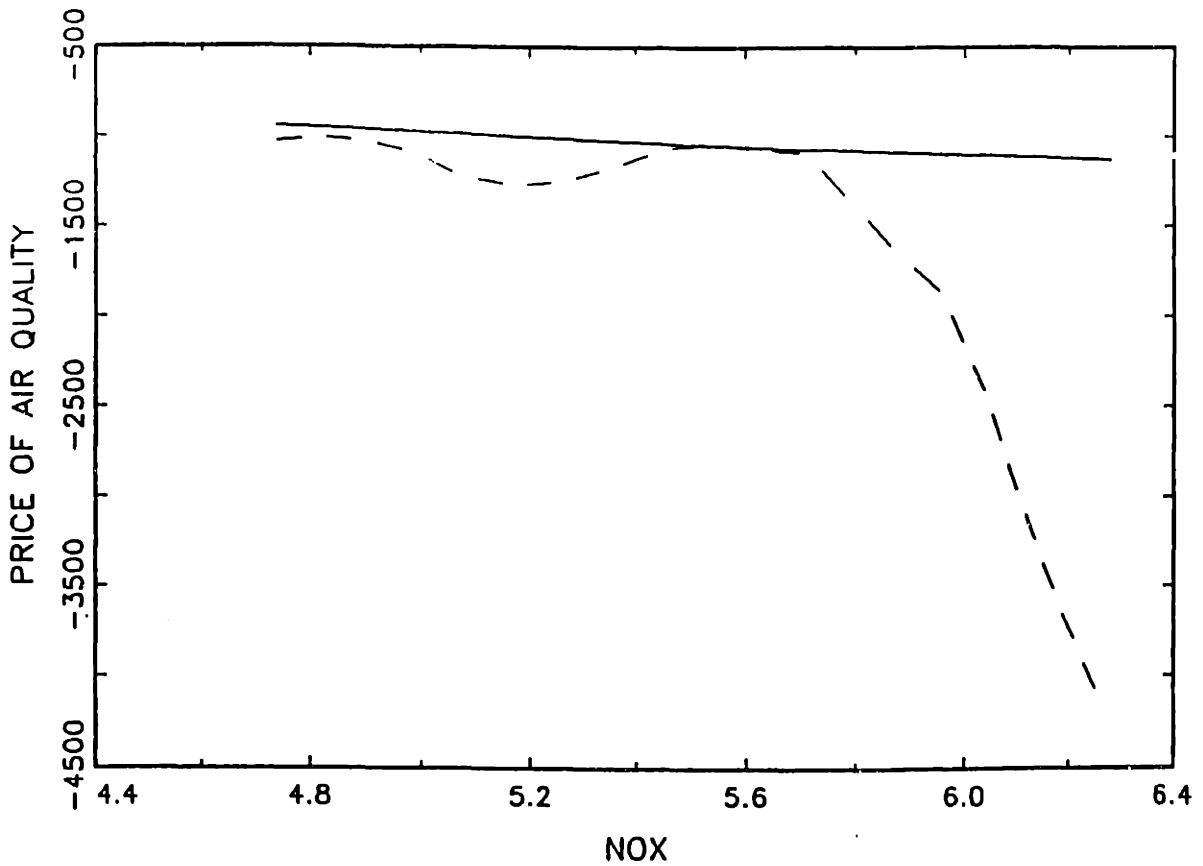

Figure 3-8: Hedonic Price of Pollution: Low Status Communities. Middle Index Value; (Continuous line: Log Linear Model; Broken Line: Semiparametric Model)

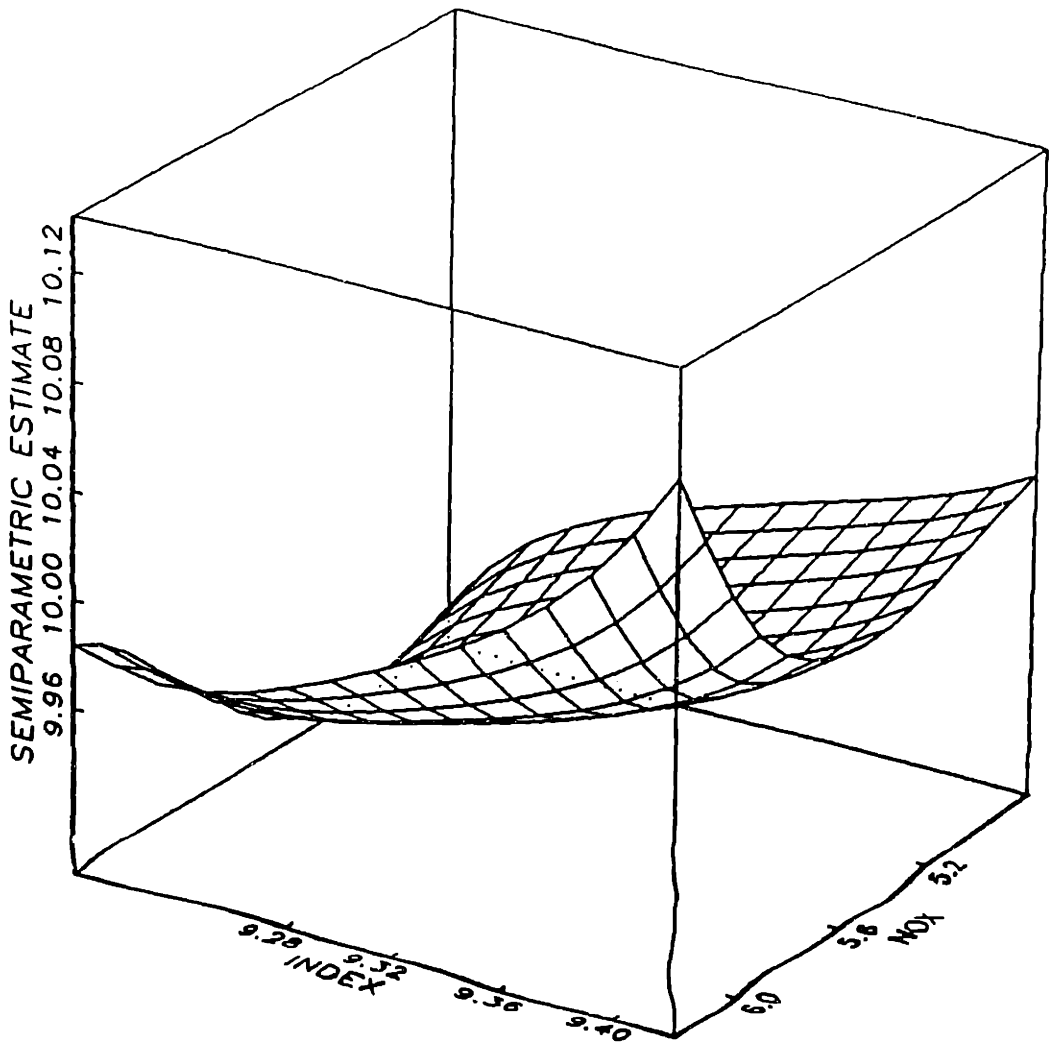

Figure 3-9: Log House Price as Related to Pollution and Index of Other Housing Attributes. (Middle LSTAT Value).

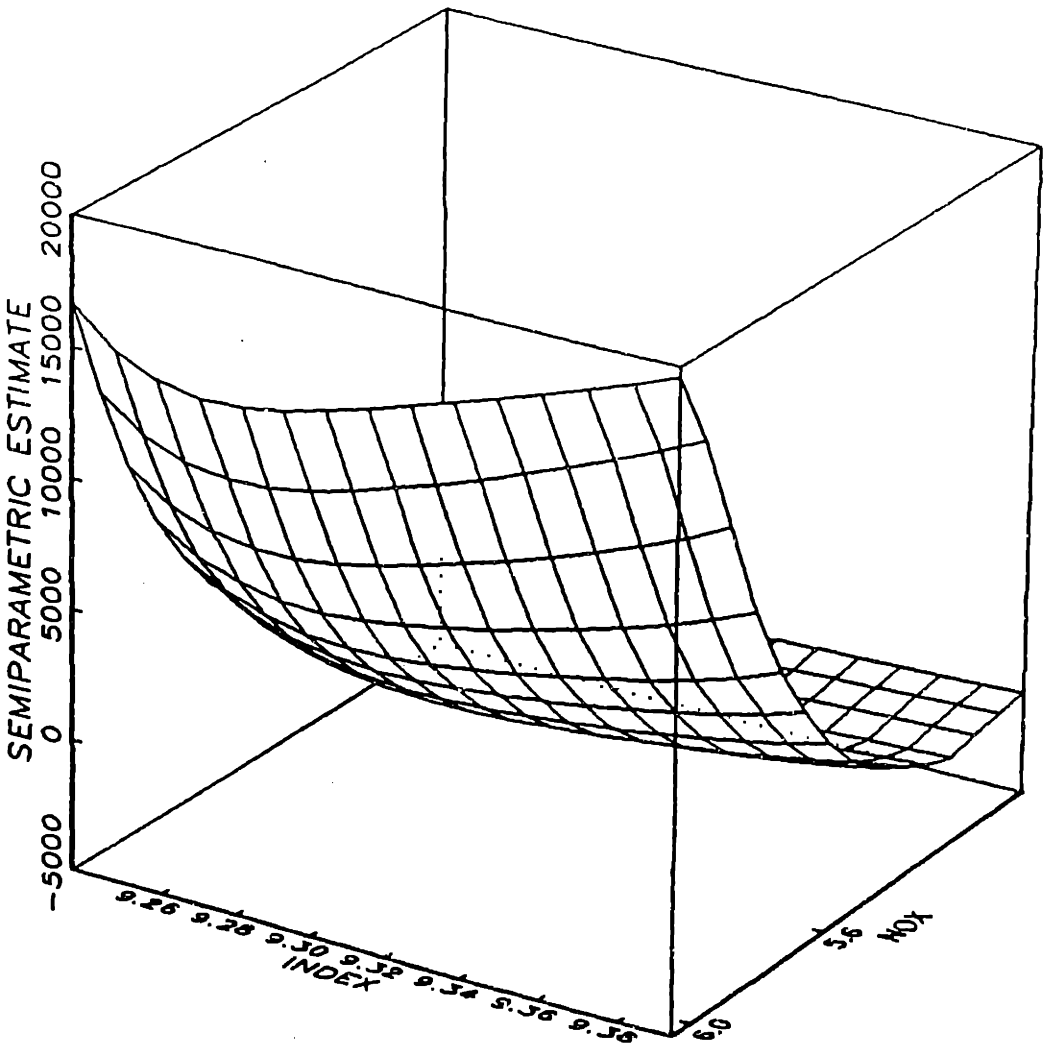

Figure 3-10: Hedonic Price of Pollution: High Status Communities.

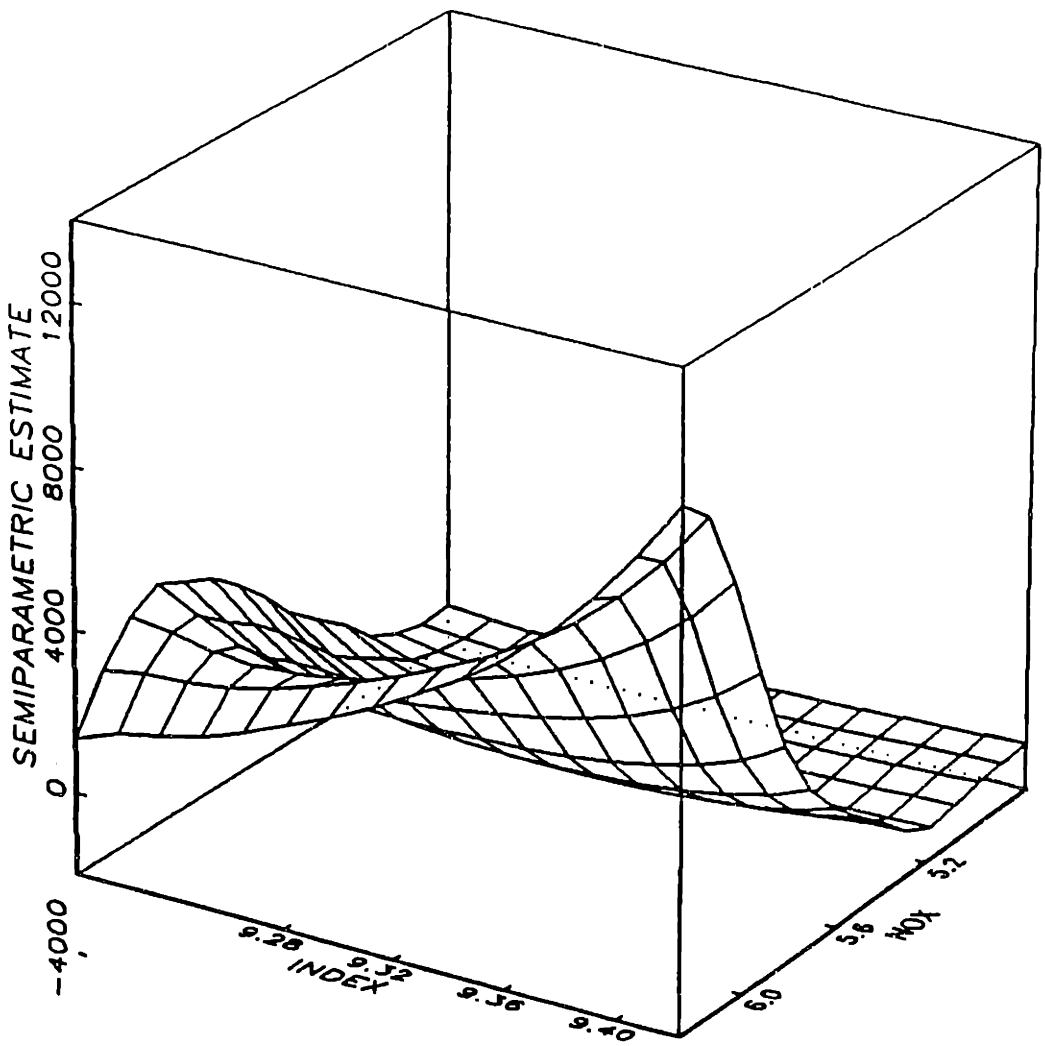

Figure 3-11: Hedonic Price of Pollution: Middle Status Communities.

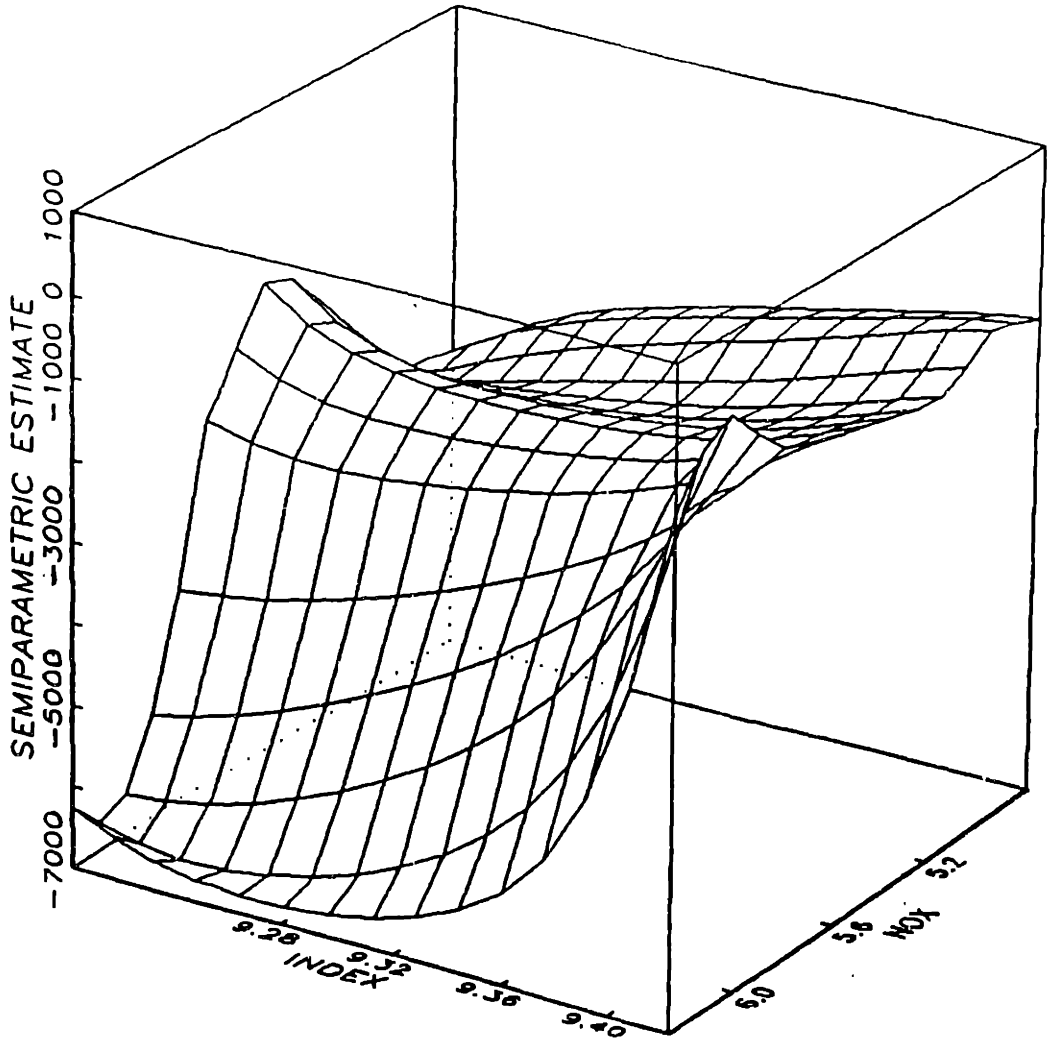

Figure 3-12: Hedonic Price of Pollution: Low Status Communities.

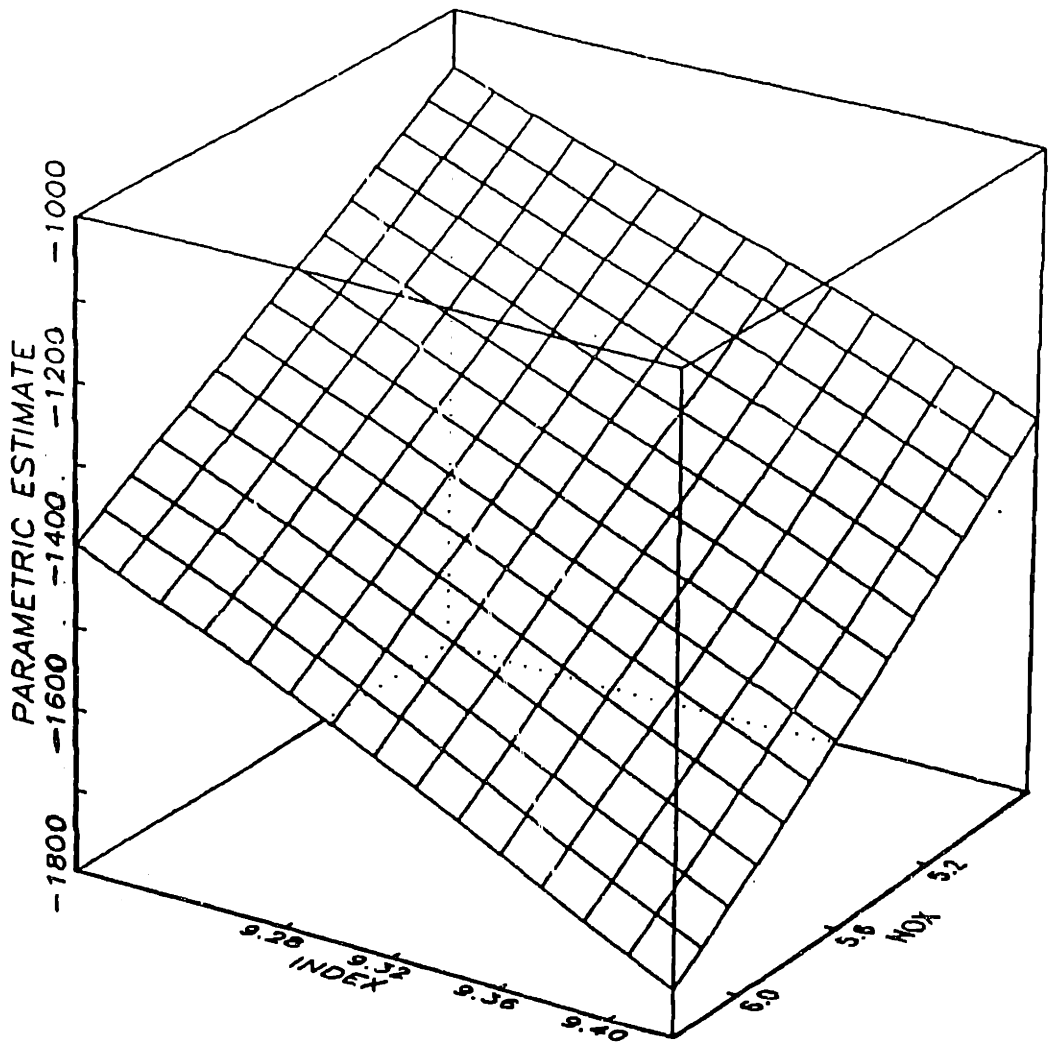

Figure 3-13: Hedonic Price of Pollution: Log Linear Model.