NORTHWESTERN UNIVERSITY

Innovation, Competition and Networks in the Supercomputer Industry

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

João Macieira

EVANSTON, ILLINOIS

December 2007

ABSTRACT

Innovation, Competition and Networks in the Supercomputer Industry

João Macieira

This dissertation develops dynamic models to examine markets with product differentiation where both firm conduct and consumer behavior is jointly influenced by switching costs, network effects and technological innovation. In Chapter 1 I propose a structural model of competition where firms set prices, introduce new products and scrap obsolete models. Network effects and switching costs are assumed to affect product demand, resulting in endogenous network creation dynamics. The model allows for product introduction and destruction without explicit modeling of these actions by instead considering a quality investment decision. I build on recent advances on dynamic games estimation and minimum-distance sieve estimators to suggest an estimator for the model.

Chapter 2 builds upon and extends the model from Chapter 1 by allowing for technological frontier investment. I estimate the model using data from the supercomputer industry. These estimates allow for counterfactual evaluation on how technological progress depends on market structure. The evidence suggests that increased levels of competition are associated with higher rates of innovation on the maximal computing speed available

in the industry. I also argue that increased competition is also associated with increased welfare, but the marginal increase in welfare is decreasing in the number of competitors.

Chapter 3 examines the long-run effects of a merger in the supercomputer industry. The primary methodology is to assume that firm behavior is consistent with Markov-Perfect Nash equilibrium (MPNE) both with and without a merger and that a merger can be described by changing observed industry states. The proposed method accounts for the effects of dimensions of non-price competition between manufacturers, which could lead to misleading conclusions about the merger effects if ignored in the analysis. I evaluate the effects of the merger between Hewlett-Packard and Convex on consumer welfare and technological progress. I argue that this merger led to increases on the maximal computing speed available in the supercomputer industry at the cost of small losses in consumer welfare.

Acknowledgements

First and foremost, I would like to thank my dissertation supervisor, Aviv Nevo, for his guidance. His help, advice, support and patience have been invaluable to me. I am also particularly indebted to Scott Stern for his endless encouragement and advice throughout my dissertation. He, along with Aviv, was an enormous help in thinking clearly about interesting questions in empirical IO and converting them into concrete research. I also thank Igal Hendel for very helpful insights. His sharp comments helped me to focus on important technical aspects of my work that may have otherwise escaped unnoticed. I am also grateful to Robert Porter for his advice and candid suggestions. Overall, regarding my committee of Aviv Nevo, Scott Stern, Igal Hendel and Robert Porter, I can't say enough about the impact of their combined expertise and general approachability.

I am grateful to many others at Northwestern, faculty and students alike. In particular, would also like to thank Shane Greenstein for invaluable discussions leading up to this dissertation. I also thank Andrew Sweeting and Elie Tamer for very insightful comments on early stages of my work. Discussions and comments from Federico Rossi, Luis Vasconcelos, Adeline Delavande and Carlos Madeira were extremely helpful.

I am also indebted to faculty outside Northwestern. I would like to thank Daniel Ackerberg and Mark Rysman for very insightful comments and suggestions. I am especially grateful to John McCallum for extremely helpful references on supplementary data sources to the TOP500 dataset. I would like also to thank Jack Dongarra for helpful clarifications about supercomputing and the TOP500 project. Financial support from Fundação Calouste Gulbenkian and Fundação para a Ciência e Tecnologia (FCT) is also gratefully acknowledged.

Special thanks are also due to Sara Filipe. I can't thank her enough for her everyday unconditional love and support. Finally, I would like to thank my sister Sofia and my mother Celeste for their endless encouragement and support. They were always there to cheer me up. To them I dedicate this dissertation.

Contents

List of Tables

List of Figures

Introduction

Most economic theory predicts that switching costs reduce consumer welfare by making markets less competitive. Recent research, however, has challenged this claim for the case of markets with differentiated products and imperfect lock-in. Dube, Hitsch and Rossi (2006) provide empirical evidence that equilibrium prices may be lower in the presence of switching costs if these are sufficiently low. The intuition for this result is that, under imperfect lock-in, the firm's incentive to attract consumers by charging low prices may dominate its incentive to "harvest" its consumer base. However, if consumers benefit from network effects by purchasing a product, the analysis of markets with switching costs must also account for two facts. First, network sizes affect not only buyer's utility, but also the distribution of switching costs across consumers. Second, firms will account for all existing network sizes when computing optimal prices. To my knowledge, no previous attempts have been made to provide an empirical framework controlling for these issues. In addition, much of the existing literature ignores the endogeneity of technological advances observed in most markets with switching costs and network effects. Policy evaluation results can be misleading if any of these issues is not accounted for in the analysis. All three chapters of my dissertation aim to fill in these gaps.

In Chapter 1 I model competition through markup and quality investment decisions when network effects and switching costs influence product demand in markets with product differentiation. I build on the dynamic games estimation literature to develop a structural model of the benefits and costs of strategic pricing and product introduction and scrappage. The model describes the dynamics of network creation under strategic behavior of firms. It nests the possibility of market tipping practices, which are commonly reported in markets with network effects and switching costs. By building on recent extensions of dynamic models to settings with multi-product firms, the model also accounts for the future benefits of quality increases (or decreases) without modeling product launch and exit explicitly. I build on recent advances on dynamic games estimation and minimum-distance estimators with unknown functions to provide necessary conditions for structural estimation of the model.

In Chapter 2 I propose and estimate a structural model of dynamic competition in the supercomputer industry to evaluate the dependence of technological innovation on market structure. Building on both the model outlined in Chapter 1 and on recent advances in the structural estimation of investment games, I develop a tractable yet dynamic model of the benefits and costs of frontier technology innovation, quality investments and strategic pricing. I account for the dependence of innovation cost on the current technological position of the firm and for the dynamic benefits of technological leadership. This model identifies the innovation incentives for all firms in the industry, and yields an estimable model of the structure of profits, frontier innovation and quality investment costs. Taking advantage of a novel and unique panel data set comprising nearly all supercomputer purchases over a 16-year period, the approach builds upon and extends recent advances in the estimation of dynamic investment games in the presence of strategic interactions. The estimates facilitate counterfactual comparisons of how the evolution of the maximal computing speed supplied in the supercomputer industry differs under different market structures. Consistent with the importance of a technology "selection effect" (Aghion, Harris, Howitt and Vickers (2001)), increased levels of competition are associated with a higher rate of innovation in the supercomputer industry. Increased competition is also associated with increased welfare, but the marginal increase in welfare is decreasing in the number of competitors.

In Chapter 3, I provide a framework for merger analysis in differentiated product industries. Unlike previous literature, I account for the effects of dimensions of non-price competition between manufacturers (e.g., advertising, R&D, product introduction and destruction) in a dynamically competitive industry. The proposed toolkit therefore avoids misleading conclusions about the post-merger welfare and equilibrium prices potentially caused by ignoring these nonprice dimensions. The proposed framework uses an estimator for firm continuation value to simulate mergers with several firms while keeping computational tractability. I use the proposed framework to examine the welfare and markup effects of the merger between Hewlett-Packard (HP) and Convex in 1995. The results suggest this merger led to improvements on the maximal computing speed available in the supercomputer industry while implying very small consumer welfare losses.

The analyses below make progress on our understanding of the effects of network effects, switching costs and technological progress on firm conduct. These analyses are also particularly valuable for policy makers. For example, they allow for the assessment of the optimal market structure both in terms of welfare and technological progress. Of course, other alternativeeapplications could be considered. One could use the model in Chapter 1 to analyze other markets with network effects and switching costs - cell phone plans and computer software may be good candidates. The model developed in Chapter 2 can be used to examine the effects of subsidy schemes to innovation on the best technology available in a given industry. The framework presented in Chapter 3 can be used to measure long-run price effects of mergers, which is a central issue in antitrust analysis.

CHAPTER 1

A dynamic model of switching costs and network creation in a differentiated products' industry

1.1. Introduction

A large body of literature examines the effects of consumer switching costs on firm conduct (see Farrell and Klemperer (2006) for a survey). However, both theoretical and empirical work rarely accounts for the importance of network effects when examining markets with switching costs. Network sizes may affect not only buyer's utility, but also the distribution of switching costs across consumers. Firms will account for network sizes when computing optimal strategies, and therefore competition models with switching costs but without network effects may lead to misleading conclusions on firm conduct. Despite the fact that switching costs and network externalities are inherently linked to dynamic competition, most of the literature has focused on static models (e.g., Greenstein (1993), Rysman (2004)).To my knowledge, no previous attempts have been made to provide an empirical framework that jointly accounts for network effects, switching costs and dynamic competition.

This Chapter considers the problem of estimating dynamic competition models in markets with product differentiation, switching costs and network externalities. I propose a dynamic oligopoly model where firms set prices, introduce new products and scrap obsolete models. In common with recent literature on markets with product differentiation, market demand is derived from a general class of discrete-choice models. Consumer utility depends on product characteristics, network sizes, and the identity of the firm from which the consumer made his last purchase.

My model builds on the frameworks of Ericson and Pakes (1995), Dube, Hitsch and Rossi (2006) and Jenkins, Liu, Matzkin and McFadden (2004). One of the major challenges is how to deal with multi-product firms. I build on the model proposed by Nevo and Rossi (2007), in which a markup-adjusted inclusive value is included as state variable. This allows me to control for product market profits while keeping a computationally manageable state space dimension. A second difficulty is how to deal with product introduction and scrappage. I use the properties of the markup-adjusted inclusive value statistic to define firm quality investment while avoiding the direct modeling of product introduction and destruction. A related difficulty is how to define pricing strategies when firms may have a time-varying number of products. I build on the model proposed by Nevo and Rossi (2007) to define equilibrium in firm-specific markups and quality investments. This allows me to have well-defined pricing strategies regardless of the firms' product portfolio composition. Another difficulty is how to allow for consumer heterogeneity in dynamic oligopoly models. I build on the frameworks of Berry, Levinsohn and Pakes (1995) and Nevo (2001) to control for switching costs differences across consumers.

The model estimation method builds both on the two step method proposed by Bajari, Benkard and Levin (2006) and on the minimum-distance sieves estimator introduced by Ai and Chen (2003). The first step is to recover the profit function parameters, markups and marginal cost series.from demand and supply estimation. In addition, this step estimates policies and state transitions nonparametrically as functions of observable states.

The second step is to estimate the remaining structural parameters of the game by using the sieve minimum distance (SMD) estimator of Ai and Chen (2003). Instead of using the continuation values in the inequality sampling method proposed by Bajari et al. (2006), this step forms a set of moments where value functions are replaced with sieves approximations. I use firm optimality conditions, envelope conditions and Bellman equations as moments in the SMD objective function. This methodology has two advantages. First, unlike the method of Bajari et al (2006), it avoids the computational burden of simulating firm continuation values. Second, it does not require the computation of equilibrium, which is problematic in the context of dynamic oligopoly games.

Understanding the forces driving market outcomes in cases where switching costs and network effects are present requires focus on a specific industry. A recent body of the industrial organization literature pursues this approach using static models of supply and demand. For example, Greenstein (1993) examines the significance of installed based in federal computer procurement decisions. Other examples include Rysman's (2004) study of network effects in the market of Yellow pages and Ho's (2005) quantification of welfare effects of observed hospital networks. Few empirical studies address the effects of either switching costs or network effects by considering dynamic models instead. Sweeting (2007) estimates the costs of format switching in the commercial radio industry. Dube, Hitsch and Rossi (2006) assess the impact of brand switching costs in equilibrium prices of package goods. Jenkins, Liu, Matzkin and McFadden (2004) analyze the browser war between Netscape and Microsoft by explicitly accounting for the network effects in that market. Most of these studies do not model the joint effect of network externalities and switching costs in consumer demand. In addition, the existing research ignores the endogeneity of network dynamics resulting from firm strategic interactions. I introduce a model which accounts for these features.

This paper also extends the endogenous quality choice literature, which has rarely focused on dynamic competition models[1]. Carranza (2006) proposes a dynamic model of product innovation in the digital camera market where the quality of new products is optimally decided by each firm. However, the impact of firms' quality decisions in the state space is assumed to be negligible in Carranza's model. This assumption cannot be considered in industries with both network effects and frequent product launch and destruction, as it defeats the purpose of relating the firms' actions to observed quality and network states. In contrast, my modeling of firms' markups and quality investment explicitly accounts for both firm strategies' dependence on the state variables and its impact on the state space.

The rest of the Chapter is organized as follows. In section 1.2 I present the dynamic oligopoly model, describing firm behavior on pricing and quality investment. The details about model estimation are presented in section 1.3. Section 1.4 concludes with a contributions summary, and discusses extensions for the proposed framework.

1.2. The Dynamic model

In this section I present a model of price and quality competition in markets with both switching costs and network effects. Time is assumed discrete with an infinite horizon, and indexed by $t \in 1, 2, ..., \infty$. At each period, multiproduct firms simultaneously decide on retail prices and quality investments. I denote F as the maximum number of firms that

[1]Some of the most recent research on static endogenous quality choice can be found in Mazzeo(2002) and Seim (2005).

can operate in the market. Only a subset of these firms, F_t^A, are active at time t. Quality investment is defined as percentage increase (or decrease) in a consumer's expected utility measure from buying a firm's product. This variation is a result of the firm's decisions on product introduction and destruction. These decisions are not modeled directly. Instead, I consider their joint impact on the firm's average quality supplied to consumers as the firm's decision variable. Firms can exit the market by scrapping all their existing products and introducing none. The number of inactive firm slots, $F - F_t^A$, are taken by potential entrants. The later are assumed to randomly enter the market, where the probability of entry depends only on observed states. In what follows, I build on the work of Jenkins et al. (2004), Ericson and Pakes (1995), Dube, Hitsch and Rossi (2006) and Nevo and Rossi (2007).

Observable states. I assume that all payoff-relevant features of firms can be encoded into a state vector. All firms observe the number of consumers who never purchased in this market, denoted N_{0t}, and firm-specific states, N_{ft}, ν_{ft} and Ξ_{ft} for all $f = 1, ..., F$ players and period $t \in 1, 2, ..., \infty$. N_{ft} (henceforth *firm network*) represents the number of consumers whose last purchase was from firm f. ν_{ft} denotes a measure of a consumer's expected utility from buying from firm f. It consists of the markup-adjusted inclusive value developed in Nevo and Rossi (2007). The latter is a function of all the firm's products characteristics, which are jointly denoted by Ξ_{ft}. I denote $\tilde{\mathbf{s}}_t \in \tilde{S}$ as the vector of all observable states at time t.

Incumbent firms. Let $\mathbf{p}_{ft}$ denote the vector of prices charged by firm f at time t for its products. In addition, I denote A_{ft} as the quality investment of firm f. At the time of the investment and pricing decisions, each firm observes a private information shock in quality investment costs, denoted ε_{ft}. Therefore, firm f's state space is $\tilde{S} \times \Theta_f$, where Θ_f corresponds to the space of realizations of ε_{ft}.

Before defining each firm's intertemporal optimization problem, I impose the following assumption:

Assumption A1 (Markovian pure strategies): In equilibrium, all players' choices are deterministic functions of payoff-relevant information.

Formally, this corresponds to a map $\tilde{\sigma}_f : \tilde{S} \times \Theta_f \longrightarrow (\mathbf{p}_f, A_f)$, for any firm f. In what follows, I assume $\Theta_f = \mathbb{R}, \forall f = 1, ..., F_t^A$. That is, shocks in quality investment costs can be negative. Innovation may be less costly due to outside factors (e.g. governmental subsidies for certain innovation projects). A firm's intertemporal optimization problem can be written in recursive form, for any profile $\tilde{\sigma} = (\tilde{\sigma}_1, ..., \tilde{\sigma}_{F_t})$ of Markovian strategies. One may question how reasonable is to assume away mixed-strategy equilibria. Fortunately, Jenkins, Liu, Matzkin and McFadden (2004) show that, under a set of mild assumptions satisfied in my setup, there are only pure-strategy MPNE in the game with probability one.

The flow payoff of firm f at time t is defined as

$$\pi_f(\mathbf{p}_t, A_{ft}, \tilde{\mathbf{s}}_t, \varepsilon_{ft}) \equiv \Pi_f(\mathbf{p}_t, \tilde{\mathbf{s}}_t) - C_A(A_{ft}, \tilde{\mathbf{s}}_t, \varepsilon_{ft})$$

where $\tilde{\mathbf{s}}_t \equiv (N_{0t}, N_{ft}, \nu_{ft}, \Xi_{ft})_{f=1}^{F}$, $\mathbf{p}_t \equiv \{\mathbf{p}_{ft}\}_{f=1}^{F}$ is a tuple of prices charged by the F_t^A firms, $\Pi_f(\mathbf{p}_t, \tilde{\mathbf{s}}_t)$ corresponds to firm f's flow market profits. $C_A(A_{ft}, \tilde{\mathbf{s}}_t, \varepsilon_{ft})$ represents quality investment costs. The firm is assumed to decide quality investment and retail prices in order to maximize the expected sum of discounted payoffs. Firms do not observe the private shocks of their rivals. Therefore, the Bellman equation of the firm is

$$\begin{aligned} V_f(\tilde{\mathbf{s}}_t, \varepsilon_f) \;=\; & \int_{\varepsilon_{-f}} \underset{\tilde{\sigma}_f(\mathbf{s},\varepsilon_f)}{Max} \{\pi_f(\tilde{\sigma}_f(\tilde{\mathbf{s}},\varepsilon_f), \tilde{\mathbf{s}}, \varepsilon_f) \\ & +\beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}} \left[V_f(\tilde{\mathbf{s}}', \varepsilon_f') | \tilde{\mathbf{s}}, \tilde{\sigma}_f(\tilde{\mathbf{s}},\varepsilon_f), \tilde{\sigma}_{-f}(\tilde{\mathbf{s}},\varepsilon_{-f})\right] \} \, dF(\varepsilon_{-f}) \end{aligned}$$

where $\tilde{\sigma}_f(\tilde{\mathbf{s}}, \varepsilon_f) \equiv (p_f(\tilde{\mathbf{s}}, \varepsilon_f), A_f(\tilde{\mathbf{s}}, \varepsilon_f))$ represents a Markovian strategy for firm f, and $E_{\mathbf{s},\sigma_f,\sigma_{-f}}[.]$ denotes firm f's expectations conditional on all firms choosing Markovian strategies, and on observable states.

Product portfolio changes induced by product introduction and scrappage may change the number of prices that each firm decides on each period. Consequently, the model just outlined may fail to have a well-defined set of actions for every period. Even if firms do not change their product portfolio over time, the problem of intertemporal choice of several products' prices implies an unmanageable state space dimension. The reason is that player's pricing strategies would be conditional on each product's characteristics, implying a computationally prohibitive state space. I deal with these difficulties by restricting the firm's pricing strategy set as follows. Let Mkp_{ft} be the equilibrium absolute markup that firm charges for each of its products at time t. In addition, define $\mathcal{F}_{ft}$ as the set of product firm f commercializes at time t, and denote mc_{jt} the marginal cost of producing product j at time t. Under the conditions on product demand proposed by Nevo and Rossi (2007) the following assumption holds:

Assumption A2 (Constant absolute markup per firm): In equilibrium, each firm f restricts attention to constant markup strategies Mkp_{ft}. That is,

$$p_{jt} - mc_{jt} = p_{lt} - mc_{lt}, \forall j, l \in \mathcal{F}_{ft}, \forall f = 1, ..., F_t^A$$

Although seemingly arbitrary, assumption A2 applies to several settings of practical interest. For example, if product demand is multinomial logit, the solution to the firm's first order conditions for each product j in static price competition implies a constant absolute markup per firm (see Nevo and Rossi (2007) and Anderson, de Palma and Thisse (1992), pp 251-252). Despite its analytic convenience, one may question whether there exists equilibrium in markup strategies. As shown later in this paper, there exists a pure-strategy dynamic equilibrium in firm markups and quality investment. One may also question whether assumption A2 is realistic in some industries, and to what extent it can be relaxed without compromising model tractability. Fortunately, assumption A2 can be extended to accommodate cases where constant markup strategies are not plausible. For example, one can instead assume that firms charge a constant absolute markup for every product that lies within a given market "segment". Under this assumption, firms would choose M markups to maximize the discounted sum of payoffs, where M is the number of segments. The more intuitive scenario of each firm choosing each product's price can be viewed as the limiting case where every product is considered a market segment. However, expanding the number of constant absolute markups that a firm may choose comes at the cost of an exponential increase in the state space's dimension. In particular, one would need M markup-adjusted inclusive values per firm in the state space. For ease of exposition and without loss of generality, I will proceed under assumption A2.

Assumption A2 implies that one can replace prices by absolute constant markup in the firm's problem. This in turn indicates that firm markup decisions are not guided by the specific attributes of each product in the market, but instead by some metric that describes each firm's average quality. Hence, I further assume that firm's strategies depend only on a subset of observable states $\mathbf{s}_t \equiv (N_{0t}, N_{ft}, \nu_{ft})_{f=1}^{F}$, and private information ε_{ft}. Under this assumption, the set of all firms' product characteristics, Ξ_{ft}, can be removed from the set of observed states. Therefore, the firm's Bellman equation becomes

$$\begin{aligned} V_f(\mathbf{s}, \varepsilon_f) \;\; = \;\; & \int_{\varepsilon_{-f}} \underset{\sigma_f(\mathbf{s},\varepsilon_f)}{Max} \{ \pi_f(\sigma_f(\mathbf{s}, \varepsilon_f), \mathbf{s}, \varepsilon_f) \\ & + \beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}} \left[V_f(\mathbf{s}', \varepsilon_f') | \mathbf{s}, \sigma_f(\mathbf{s}, \varepsilon_f), \sigma_{-f}(\mathbf{s}, \varepsilon_{-f}) \right] \} \, dF(\varepsilon_{-f}) \end{aligned}$$

where $\sigma_f(\mathbf{s}, \varepsilon_f) \equiv (Mkp_f(\mathbf{s}, \varepsilon_f), A_f(\mathbf{s}, \varepsilon_f))$ represents a Markovian strategy for firm f, and $E_{\mathbf{s},\sigma_f,\sigma_{-f}}[.]$ denotes firm f's expectations conditional on all firms choosing Markovian strategies, and on observable states. It can be shown that each player's value function is unique by verifying Blackwell's sufficient conditions for a contraction mapping.

Potential entrants. Potential entrants are defined as firms are defined as players with no products for sale. These players are assumed to decide whether to enter the market. In case entry takes place, the potential entrant will have a set of products for sale I assume that a potential entrant decides on entry after a shock which determines future average product quality ν_{ft}. Thus, a potential entrant's state space is assumed to be $S \times \Gamma_{3f}$, where Γ_{3f} corresponds to the shock determining future average quality.

These shocks are assumed private information of the entrants. It is assumed that potential entrants get a payoff of zero by the time they decide on entry.

1.2.1. Product Market Profits

In this subsection I model product market profits. The modeling of product market profits proceeds in two steps. First, I model product demand. Second, I assume price competition in a differentiated products market to model supply. Finally I derive each firm's profit as a function of observed states and absolute markups.

Demand. The purchasing behavior in several markets is influenced by consumer heterogeneity. Several studies in the literature provide empirical evidence on behalf of this claim (e.g., Berry, Levinsohn and Pakes (1995), Nevo (2001)). Consequently, the modeling of demand should control for consumer heterogeneity to the extent possible without compromising tractability in dynamic oligopoly models. I deal with this issue as follows.

Let D_{rft} be a dummy variable which equals one if consumer r is part of firm f's network and is zero otherwise. I follow the characteristics-based discrete choice approach described in McFadden (1981). Let J_t be the number of products available at time t. The utility function for a buyer r interested in acquiring product j at time t is assumed to be[2]

$$U_{rjt} = \gamma X_{jt} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} + \Lambda_{f(j)} D_{rtf(j)} + \epsilon_{rjt}$$

[2] I am implicitly assuming that buyers will purchase at most one unit of product. An alternative modeling strategy would be to account for multiple purchases as proposed by Hendel (1999).

where X_{jt} is a K-dimensional column vector of product characteristics, p_{jt} is the price of product j at time t (in constant dollars), ξ_{jt} regards unobserved (by the researcher) product attributes, $N_{f(j)t}$ corresponds to the number of consumers at time t who already purchased from $f(j)$ (i.e. the firm which commercializes product j), $D_{rf(j)t}$ is a dummy variable which equals one if consumer r is part of $f(j)$'s network, and ϵ_{rjt} is a stochastic term. I assume that $D_{rf(j)t}$ is not observed by the econometrician, but its distribution is. ψ measures the magnitude of network effects, while $\Lambda_{f(j)}$ measures switching costs associated with the firm which produces product j.

Consumers can choose an "outside good", whose utility is $U_{r0t} = \xi_{0t} + \epsilon_{r0t}$. I allow for the mean utility of the outside alternative, ξ_{0t}, to evolve over time. This would be the case, for example, in the supercomputer market. The outside alternative would represent not only other types of computers that buyers may consider powerful enough to meet their computing needs (e.g., mainframes, workstations), but also the flow utility from using a high-end computer one already owns.

The following assumption is imposed for not only for computational convenience, but also to ensure that assumption A2 holds:

Assumption A3: $\forall j = 0, 1, .., J_t$, ϵ_{rjt} is identically and independently distributed extreme-value type I.

Nevo and Rossi (2007) prove that, under assumption A3, firms will restrict attention to constant absolute-markup pricing strategies. Under assumption A3, the market share for product j at time t conditional on consumer state D_{rtf} is given by

$$q_{jt}(p|D_{rt}) = \frac{\exp(\gamma X_{jt} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_{l} \exp(\gamma X_{lt} + \psi N_{d(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{d(l)} D_{rd(l)t})}$$

and the conditional outside good share is

$$q_{0t}(p|D_{rt}) = \frac{1}{1 + \sum_{l} \exp(\gamma X_{lt} + \psi N_{d(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{d(l)} D_{rd(l)t})}$$

Note that this specification also nests the case where a consumer may be part of multiple firm networks. In what follows, I assume that consumers may belong to at most one network. That is, if $D_{rtf} = 1$ for a given f , then $D_{rti} = 0 \; \forall i \neq f$.

Under this assumption, using the Law of Total Expectations, the aggregate demand for a given product j at time t is given by

$$M_t q_{jt}(p) = \sum_{k=1}^{F} N_{kt} q_{jt}(p|D_{rkt} = 1)$$

where $M_t \equiv \sum_{k=1}^{F} N_{kt}$ is the market size as described in standard models of differentiated product's industries (e.g., Nevo (2001), Berry, Levinsohn and Pakes (1995)), and $q_{jt}(p)$ is the unconditional probability purchase of product j.

In order to ensure that this framework allows for both consumer heterogeneity and tractability of the dynamic model, it is necessary to verify that product demand can be written only as a function of observed states and actions. The following claim is the first step to address this issue.

Claim 1 The probability of purchase of each product conditional on consumer r's type vector $D_{rt} = (D_{rt1}, ..., D_{rtF})$ depends only on the vector of equilibrium markups, observed states, and product attributes. That is, $q_{jt}(p|D_{rt}) = q_{jt}(\mathbf{Mkp}_t|D_{rt})$, $\forall j = 1, ..., J_t$.

Proof. Following Nevo and Rossi (2007), the conditional probability of purchase can be written as

$$q_{jt}(p|D_{rt}) = \frac{\exp(-\alpha Mkp_{f(j)t})\exp(\gamma X_{jt} + \psi N_{f(j)t} - \alpha mc_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_l \exp(-\alpha Mkp_{i(l)t})\exp(\gamma X_{lt} + \psi N_{d(l)t} - \alpha mc_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{d(l)} D_{rd(l)t})}$$

Since markups are firm-specifc, the denominator can be simplified further, implying a product market share of

$$q_{jt}(\mathbf{Mkp}_t, \mathbf{s}_t|D_{rt}) = \frac{\exp(Mkp_{f(j)t})\exp(\gamma X_j + \psi N_{f(j)t} - \alpha mc_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt})\exp(\nu_{dt})\exp(\psi N_{dt})\exp(\Lambda_d D_{rtd})}$$

where $\nu_{ft} = \ln\left(\sum_{l \in \mathcal{F}_{it}} \exp(\gamma X_{lt} - \alpha mc_{lt} + \xi_{lt} - \xi_{0t})\right)$ is the markup-adjusted inclusive value metric of Nevo and Rossi (2007).

Inclusive values have been widely used as state variables in dynamic models since its development in McFadden (1981). In the context of the multinomial logit demand system, a firm-specific inclusive value consists on the buyer's expected utility from purchasing a firm's product. That is, for every firm f, the inclusive value is defined as

$$\omega_{ft} = \ln\left(\sum_{l\in\mathcal{F}_{it}} \exp(\gamma X_{lt} - \alpha p_{lt} + \xi_{lt} - \xi_{0t})\right)$$

Of course, one could consider inclusive values specific to sets other than firm product portfolios. Melnikov (2001) considers the inclusive value from buying any existing product for modeling consumer demand for printers. More recently, Hendel and Nevo (2006) used inclusive values from buying specific quantities to estimate demand for laundry detergent. However, this firm-specific quality measure is dependent on prices, which are assumed to be chosen by firms. Therefore, in order to use firm-specific inclusive values in my dynamic model, I must adjust for its endogeneity in prices. Following Nevo and Rossi (2007), one can use the fact that $Mkp_{ft} \equiv p_{jt} - mc_{jt}$, $\forall j \in \mathcal{F}_{ft}$, to write the firm inclusive value as

$$\omega_{ft} = -\alpha Mkp_{ft} + \ln\left(\sum_{l\in\mathcal{F}_{it}} \exp(\gamma X_{lt} - \alpha mc_{lt} + \xi_{lt} - \xi_{0t})\right)$$

where the last parcel corresponds to the *markup-adjusted inclusive value.*

Supply. Each firm f in the industry produces some subset $\mathcal{F}_{ft}$ of the $j = 1, ..., J_t$ products commercialized at time t. Firm f 's flow profits at time t are defined by

$$\Pi_f(\mathbf{p}_t, \mathbf{s}_t) = M_t \sum_{j\in\mathcal{F}_{ft}} (p_{jt} - mc_{jt}) q_{jt}(p, \mathbf{s}_t) - C_f \mathbf{1}\{\exp(\nu_{ft}) > 0\}$$

where $C_f 1\{\exp(\nu_{ft}) > 0\}$ corresponds to fixed production costs conditional on the firm being active in the market (i.e., it has at least one product being commercialized). The marginal cost of producing product j at time t, mc_{jt}, is assumed to be time-varying and dependent on product characteristics. M_t corresponds to the market size state at time t, which is defined as being the number of consumers potentially interested on purchasing a product. The last component of flow market profits is $q_{jt}(p, \mathbf{s}_t)$, which denotes the probability of purchasing product j at time t, given prices of all products and observable states $\mathbf{s}_t$.

Firms are assumed to compete in prices. However, the pricing decisions will affect not only the profits at time t, but also future payoffs. Intuitively, firms pricing decisions will also take into account the fact that prices affect purchase probabilities, which in turn affect the evolution of some observable states. For example, a firm may charge prices below marginal costs to attract consumers to its network. Hence, static price competition models will fail to correctly describe firm pricing behavior. I use the results from the previous section to both simplify the profit function and account for these issues.

First, I replace $(p_{jt} - mc_{jt})$ by the firm-specific absolute markup Mkp_{ft} in the flow profit. Second, I use the fact that

$$M_t q_{jt}(p) = \sum_{k=1}^{F} N_{kt} q_{jt}(p | D_{rkt} = 1)$$

to write flow profits as

$$\Pi_f(\mathbf{Mkp}_t, \mathbf{s}_t) = Mkp_{ft} \sum_{k=1}^{F} N_{kt} \left(\sum_{j \in \mathcal{F}_{ft}} q_{jt}(\mathbf{Mkp}_t, \mathbf{s}_t | D_{rkt} = 1) \right) - C_f \mathbf{1}\{\exp(\nu_{ft}) > 0\}$$

Again following Nevo and Rossi (2007), it can be shown that the probability of purchase from a given firm f's conditional on consumer r's type vector $D_{rt} = (D_{rt1}, ..., D_{rtF})$ depends only on the vector of equilibrium markups and observed states. The sum of purchase probabilities across firm f's products can be simplified to

$$\sum_{j \in \mathcal{F}_{ft}} q_{jt}(\mathbf{Mkp}_t, \mathbf{s}_t | D_{rt}) = \frac{\exp(-\alpha Mkp_{ft}) \exp(\nu_{ft}) \exp(\psi N_{ft}) \exp(\Lambda_f D_{rtf})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt}) \exp(\nu_{dt}) \exp(\psi N_{dt}) \exp(\Lambda_d D_{rtd})}$$

which is a function of markups and observed states only. In addition, the conditional probability of choice of the outside alternative is also a function of markups and observed states only. This probability is given by

$$q_{ot}(\mathbf{Mkp}_t, \mathbf{s}_t | D_{rt}) = \frac{1}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt}) \exp(\nu_{dt}) \exp(\psi N_{dt}) \exp(\Lambda_d D_{rtd})}$$

Under these results, assuming the existence of a dynamic equilibrium in markups, the flow profit depends from product characteristics only via markup-adjusted inclusive. In particular, the equilibrium profit function is given by

$$\begin{aligned} \Pi_f(\mathbf{Mkp}_t, \mathbf{s}_t) &= Mkp_{ft} \sum_{k=1}^{F} N_{kt} \frac{\exp(-\alpha Mkp_{ft} + \nu_{ft} + \psi N_{ft} + \Lambda_f 1\{k = f\})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt} + \nu_{dt} + \psi N_{dt} + \Lambda_d 1\{k = d\})} \\ &\quad - C_f 1\{\exp(\nu_{ft}) > 0\} \end{aligned}$$

Note that this profit function is well defined even if only $F_t^A < F$ are active in the market. By definition $\exp(\nu_{ft}) = 0$ for firms with no product portfolio, which implies profits equal to zero.

1.2.2. Quality Investment and Costs

One of the major challenges on analyzing multiproduct industries is how to deal with product introduction and scrappage. This paper allows for changes in the firm product portfolio by considering percentage variations on firms' current markup-adjusted inclusive values. I exploit the fact that, under the assumptions imposed so far, markup-adjusted inclusive values are sufficient statistics for firm quality. For ease of exposition, define $\varphi_{ft+1} \equiv \exp(\nu_{ft+1})$ as the exponential of a firm's markup-adjusted inclusive value at time $t+1$. In addition, denote $\mathcal{F}_{ft}^{IN}$ and $\mathcal{F}_{ft}^{OUT}$ as the set of products that firm f decides to launch and scrap, respectively. Then the set of products that firm f commercializes at time $t+1$ is $\mathcal{F}_{ft+1} = \{\mathcal{F}_{ft}^{IN} \cup \mathcal{F}_{ft}\} \setminus \mathcal{F}_{ft}^{OUT}$. Then, φ_{ft+1} can be decomposed as

$$
\begin{aligned}
\varphi_{ft+1} \quad = \quad & \left(\sum_{j \in \mathcal{F}_{ft}} \exp(\gamma X_{j,t+1} - \alpha mc_{j,t+1} + \xi_{j,t+1} - \xi_{0,t+1}) \right) \\
& + \left(\sum_{l \in \mathcal{F}_{ft}^{IN}} \exp(\gamma X_{lt+1} - \alpha mc_{lt+1} + \xi_{lt+1} - \xi_{0t+1}) \right) \\
& - \left(\sum_{k \in \mathcal{F}_{ft}^{OUT}} \exp(\gamma X_{k,t+1} - \alpha mc_{k,t+1} + \xi_{k,t+1} - \xi_{0,t+1}) \right)
\end{aligned}
$$

The rate of quality investment is defined as

$$A_{ft} = \frac{\left(\sum_{l\in\mathcal{F}_{ft}^{IN}} \exp(\gamma X_{lt} - \alpha mc_{lt} + \xi_{lt} - \xi_{0t})\right) - \left(\sum_{k\in\mathcal{F}_{ft}^{OUT}} \exp(\gamma X_{k,t} - \alpha mc_{k,t} + \xi_{k,t} - \xi_{0,t})\right)}{\varphi_{ft}}$$

Intuitively, A_{ft} represents the ratio of net amount of product quality the firm added (or subtracted) to the existing product quality stock. The fact that A_{ft} may be negative means that the firm may choose to degrade its products' quality. This can apply to cases where the firm introduces new products, although of not enough quality to compensate the one of scrapped goods. Therefore, the model is rich enough to allow for several complex product introduction and scrappage strategies. Note that A_{ft} is bounded below by -1, which corresponds to scrappage of all products in $\mathcal{F}_{ft}$ with no new products being launched. Hence, $A_{ft} = -1$ can be interpreted as a market exit decision. The details on how A_{ft} impacts φ_{ft+1} can be found in the next section.

I assume that the firm incurs in costs $C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{ft}))$ of adjusting its quality level (either by increasing or decreasing it). I further assume that $C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{ft}))$ satisfies the property

$$\frac{\partial C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{ft})}{\partial A_{ft}} = C_1(A_{ft}, \mathbf{s}_t) + C_2(A_{ft}, \mathbf{s}_t)\varepsilon_{ft}$$

This is a special case of the form $C_A(A_{ft}, \mathbf{s}_t, A_{ft}\varepsilon_{ft}))$ where the function has Lipschitz derivatives. Jenkins, Liu, Matzkin and McFadden (2004) assume this form in their dynamic game and prove the existence of pure-strategy MPNE.

1.2.3. State transitions

The specification of the continuation value in the firms' Bellman equations requires assumptions on state transition functions. I impose the following auxiliary assumptions:

Assumption A4: Private shocks on quality investment costs are independently and identically distributed over time and players.

Assumption A5: For every firm f, the private information state ε_{ft} is assumed independent of observed states $\mathbf{s}_t$.

These assumptions are motivated by computational and tractability concerns. Allowing for serial correlation implies a significant (and unaffordable) increase in computational burden. The details on each observable state's transitions are presented in the next subsections.

Markup-adjusted inclusive values. The transition of industry markup-adjusted inclusive values requires examining incumbents and potential entrants separately. I start by rewriting the exponential of the markup-adjusted inclusive value as

$$\begin{aligned}
\varphi_{ft+1} &= \exp(\xi_{0,t+1}-\xi_{0,t})\left(\sum_{j\in\mathcal{F}_{ft}}\exp(\gamma X_{j,t}-\alpha mc_{j,t}+\xi_{j,t}-\xi_{0,t})\exp(u_{j,t+1})\right)\\
&\quad+\exp(\xi_{0,t+1}-\xi_{0,t})\left(\sum_{l\in\mathcal{F}_{ft}^{IN}}\exp(\gamma X_{lt}-\alpha mc_{lt}+\xi_{lt}-\xi_{0t})\exp(u_{l,t+1})\right)\\
&\quad-\exp(\xi_{0,t+1}-\xi_{0,t})\left(\sum_{k\in\mathcal{F}_{ft}^{OUT}}\exp(\gamma X_{k,t}-\alpha mc_{k,t}+\xi_{k,t}-\xi_{0,t})\exp(u_{k,t+1})\right)
\end{aligned}$$

where

$$u_{k,t+1} \equiv \gamma(X_{k,t+1} - X_{k,t}) - \alpha(mc_{k,t+1} - mc_{k,t}) + (\xi_{k,t+1} - \xi_{k,t})$$

Intuitively, this variable corresponds to the markup-adjusted quality variation of consuming product k from period t to $t+1$. For ease of exposition and without loss of generality, I assume that (i) a product's observed characteristics do not change after its introduction, i.e., $X_{k,t+1} = X_{k,t}$, $\forall t, k$, (ii) $\xi_{k,t} = \xi_k + \Delta\xi_{k,t}$, and (iii) marginal costs are a time-varying stochastic function given by

$$mc_{k,t} = mc_0 + mc_t(X_{k,t}) + \eta_{kt}^{MC}$$

Under these assumptions, the variable $u_{k,t+1}$ simplifies to

$$u_{k,t+1} \equiv y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) + \zeta_{k,t+1} + (\Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta_{k,t+1}^{MC} - \eta_{k,t}^{MC}))$$

where $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$ is the projection of $-\alpha(mc_{t+1}(X_{k,t}) - mc_t(X_{k,t}))$ onto the space of observed states and actions. Intuitively, $\zeta_{k,t+1}$ corresponds to the parcel of marginal cost variation that is not explained by firms' strategic interactions and observed states.

In order to have a well-defined transition for markup-adjusted inclusive values, auxiliary assumptions are necessary. I pose the following conditions:

Assumption A6: The process $\zeta_{k,t+1} + \Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta_{k,t+1}^{MC} - \eta_{k,t}^{MC})$ is common to all products within the firm's portfolio. That is, for every firm f, we have

$$\varrho_{ft+1} \equiv \zeta_{k,t+1} + \Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta^{MC}_{k,t+1} - \eta^{MC}_{k,t})$$

Assumption A6 implies that the temporal evolution of a firm's product attributes can be encoded into a firm-specific shock. The latter may include shocks common to all firms which might not be observed by the econometrician (e.g., imput price fluctuations).

Assumption A7: There exists a constant d_0 for which the process $\xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1} - d_0$ is a martingale. That is, $E[\xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1} - d_0|\mathbf{s}_t, \mathbf{A}_t] = 0$ for every firm f.

Assumption A7 is necessary for stationary markup-adjusted inclusive values. It allows, for example, for a random walk with positive drift in $\xi_{0,t+1}$ (i.e., a exponential growth of the outside alternative mean utility over time) with ϱ_{ft+1} stationary. Assumption A8 also nests the case where the series $\xi_{0,t+1}$ and ϱ_{ft+1} are not stationary, but $\xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1}$ is.

Under these assumptions, the transition of an active firm f's markup-adjusted inclusive value can be written as

$$\ln(\varphi_{ft+1}) = d_0 + y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) + \ln(\varphi_{ft}) + \ln(1 + A_{ft}) + \eta_{ft}$$

where $\eta_{ft} \equiv \xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1} - d_0$ is a zero-mean shock. Note that $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$ can be chosen to guarantee stationary markup-adjusted inclusive values. For example, if $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) = -d_1 \ln(\varphi_{ft}) + d_2 \ln(1 + A_{ft})$. The resulting transition equation for each markup-adjusted inclusive value is given by

$$\upsilon_{ft+1} = d_0 + (1 - d_1)\upsilon_{ft} + (1 + d_2)\ln(1 + A_{ft}) + \eta_{ft}$$

We are left to define the transition for potential players. By definition, these players occupy slots in the state space where $\exp(\nu_{ft}) = 0$. I assume that these inactive firms enter the market (i.e., will have $\exp(\nu_{ft+1}) > 0$) exogenously. Again using the notation $\varphi_{ft+1} = \exp(\nu_{ft+1})$, the transition of inclusive values for inactive players is given by

$$\varphi_{ft+1} = \max\{0, w(\mathbf{s}_t) + \eta_{ft}^{IN}\}$$

where $w(\mathbf{s}_t)$ is a function of all observed states and η_{ft}^{IN} is a zero-mean shock. The inclusive-value transition function for all players in the game can be concisely written as

$$\begin{aligned}\exp(\nu_{ft+1}) &= \mathbf{1}\{\exp(\nu_{ft}) > 0\}\exp(d_0 + y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) + \nu_{ft} + \ln(1 + A_{ft}) + \eta_{ft}) \\ &\quad + (1 - \mathbf{1}\{\exp(\nu_{ft}) > 0\})\max\{0, w(\mathbf{s}_t) + \eta_{ft}^{IN}\}\end{aligned}$$

Networks. In what follows, I build on the work of Dube, Hitsch and Rossi (2006). The latter considers consumer switching costs in the context of packaged goods without any network effects. I allow not only for both switching costs and network effects, but also for (i) arrival of new consumers to the market and (ii) dynamic network formation. The transition function for $\mathbf{N}_t = [N_{0t}, N_{1t}, ..., N_{Ft}]^T$ is given by

$$\mathbf{N}_{t+1} = Q(\mathbf{Mkp}_t, \mathbf{s}_t)\mathbf{N}_t + \tilde{\mathbf{N}}_{0t+1}$$

where $\tilde{\mathbf{N}}_{0t+1} = [\Delta N_{0t+1}, 0, ..., 0]^T$ corresponds to the column vector with the number of new consumers arriving to the market in the first cell and zero otherwise. This construction implicitly assumes that thereare no stochastic terms in firm networks. This implies that there are no consumers leaving a network for exogenous reasons. $Q(\mathbf{Mkp}_t, \mathbf{s}_t)$ is a matrix whose entries $Q_{i,j}$ are defined by the probability of a consumer joining firm i's network when he currently belongs to firm j's network. These probabilities can be derived from the purchase probabilities from a given firm. However, one must account for the fact that a consumer who belongs to a firm network still belongs to that network if he chooses the outside alternative. Under these assumptions, every entry in Q is defined as follows

$$Q_{i,0} = P_{i,0}, \forall i = 0, 1, ..., F$$

$$Q_{i,i} = P_{i,i} + P_{0,i}, \forall i = 1, ..., F$$

$$Q_{0,i} = 0, \forall i = 1, ..., F$$

$$Q_{i,j} = P_{i,j}, \forall i \neq j, i, j = 1, ..., F$$

where $P_{i,j}$ corresponds to the probability of buying a product from firm i given that the consumer is part of firm j's network. These correspond to

$$P_{i,j} = \frac{\exp(-\alpha Mkp_{it}) \exp(\nu_{it}) \exp(\psi N_{it}) \exp(\Lambda_i 1\{i = j\})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt}) \exp(\nu_{dt}) \exp(\psi N_{dt}) \exp(\Lambda_d 1\{d = j\})}$$

For the case where all network series are observed, ΔN_{0t+1} can be estimated by using the property that

$$\sum_{d=0}^{F} N_{dt+1} = \sum_{d=0}^{F} N_{dt} + \Delta N_{0t+1}$$

However, they may be cases where the series N_{0t} is not observed. In that case, auxiliary assumptions are necessary to recover both N_{0t} and ΔN_{0t+1}. This is the case analyzed in Chapter 2, in the context of the supercomputer industry. The discussion of this case is therefore deferred to Chapter 2.

1.2.4. Equilibrium concept

The structure of the game allows me to assume that firms choose markups rather than specific retail prices. Therefore, using assumption A1, I restrict attention to a pure-strategy equilibrium where players choose Markovian strategies of the form $\sigma_f(\mathbf{s}, \varepsilon_f) \equiv \{Mkp_f(\mathbf{s}, \varepsilon_f), A_f(\mathbf{s}, \varepsilon_f)\}$. Let $V_f(\mathbf{s}, ,\varepsilon_f|\sigma_f, \sigma_{-f})$ be firm f's expected discounted payoffs when he chooses strategy σ_f and his rivals choose σ_{-f}. A Markov-Perfect Nash Equilibrium (MPNE) of this dynamic oligopoly game can be defined as follows

Definition 1 (MPNE): A Markovian strategy profile $\sigma^* \equiv (\sigma_1^*, ..., \sigma_F^*)$ is an MPNE if, for every firm f, σ_f^* solves f's problem given σ_{-f}^*. That is, $V_f(\mathbf{s},\varepsilon_f|\sigma_f^*, \sigma_{-f}^*) \geq V_f(\mathbf{s},\varepsilon_f|\hat{\sigma}_f, \sigma_{-f}^*)$ for all $\mathbf{s}$ and alternative strategy $\hat{\sigma}_f$.

Existence of equilibrium can be shown in two ways. The first is to invoke Theorem 1 of Jenkins, Liu, Matzkin and McFadden (2004), since the conditions for this theorem to hold are met by the assumptions imposed so far. The second is to applying Brower's fixed point theorem after writing the system for first-order conditions on markups and quality investment in fixed point form for all active firms. The fact each firm's value function is unique ensures the existence of at least one equilibrium[3]. Uniqueness can in principle be verified either by directly proving that the fixed point form of the system is a contraction, or by checking if the system satisfies Blackwell's sufficient conditions. Unfortunately, MPNE uniqueness verification using these and other proof strategies are particularly difficult and still an open area of research (see Doraszelski and Satterthwaite (2007) for a discussion). In addition, uniqueness conditions would require assumptions on the game parameters, which can only be verified upon the estimation of the dynamic game. The next section addresses the methods to recover these parameters from the data.

1.3. Estimation methods

The structural estimation of the model combines the two-step method proposed by Bajari, Benkard and Levin (2006) with Ai and Chen's (2003) nonparametric estimator. Several other dynamic models' estimation methods have been introduced in the literature (e.g. Aguirregabiria and Mira (2006), Berry, Ostrovsky and Pakes (2005), Hotz, Miller, Sanders and Smith (1994), Jenkins, Liu, Matzkin and McFadden (2004)) after Rust's (1987) pioneer work on dynamic discrete choice[4]. Even though the method of Bajari et al. (2006) has been considered in other dynamic oligopoly studies (e.g., Beresteanu and Ellickson (2006), Ryan (2006), Sweeting (2007)), this Chapter will follow a modified framework. I start by providing a general overview of my estimation strategy, explaining

[3]If each firm's value function were not unique, existence of equilibrium could instead be done by invoking Kakutani's fixed point theorem.
[4]A comprehensive survey of these methods can be found in Ackerberg et al. (2006)

its differences with respect to the method of Bajari et al. and the reasons for modifying their method. Then I discuss the technical details of each step of the estimation procedure.

1.3.1. An overview of the estimation

The first-step of my estimation strategy closely follows the one of Bajari, Benkard and Levin (2006). It consists on estimating the flow profit parameters, as well as the observable state transitions and policies. Unlike Bajari et al., the players' optimal policies will not be used for forward simulation of continuation values. The reason is that the later will not be necessary for my second step of estimation. Instead, the policy estimates and its residuals are used to integrate out private information of rivals when forming moments for the second round of estimation.

After choosing appropriate instruments, profit function parameters are recovered by using Nevo's (2001) demand estimation method. Markup and marginal cost series can be obtained by projecting prices onto a firm-specific markup function. These estimates are then used to compute Nevo and Rossi's (2007) markup-adjusted inclusive values for all firms, quality investments and firm network transitions. The laws of motion for markup-adjusted inclusive values and N_{0t} can be consistently estimated with standard methods upon imposing structure on the function $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$ and on the shock ΔN_{0t+1}, respectively. Finally, markup and quality investment policies can be consistently estimated using semi- or nonparametric methods (see Jenkins et al. (2004) for a discussion).

In the second stage of estimation, I use the framework of Ai and Chen (2003) to estimate the parameters not recovered in the first stage. This method estimates models which can be fully described by moment conditions of the form

$$E\left[\Psi(Z,\theta_0,h_o(.))|X\right]=0$$

where $\Psi(.)$ is a vector of known functions, θ_0 is a vector of model parameters (e.g., the remaining parameters of the dynamic model), $h_o(.)$ is a vector of unknown but identified functions, Z is a vector of both endogenous and exogenous variables, and X is a set of exogenous variables containing all the exogenous controls in Z. The fact that firm optimality conditions on markups and quality investment decisions depend on integral transforms of both value functions and rival strategies (which usually do not have a closed form) suggests the usage of a method where these functions can be approximated arbitrarily well. Ai and Chen (2003) propose a method where θ_0 and $h_o(.)$ are consistently estimated by replacing the latter by a vector of sieve approximating functions in the moment conditions. Denoting $\alpha=(\theta_0,h_o)$, and letting $\hat{m}(X,\alpha)$ be the a consistent nonparametric estimator of $m(X,\alpha)\equiv E\left[\Psi(Z,\theta_0,h_o(.))|X\right]$ (e.g., its sample analog), Ai and Chen's (2003) sieve minimum distance (SMD) estimator is defined as

$$\hat{\alpha}\equiv\arg\min_{\alpha}\frac{1}{n}\sum_{i=1}^{n}\hat{m}(X_i,\alpha)'\left[\hat{\Sigma}(X_i)\right]^{-1}\hat{m}(X_i,\alpha)$$

My estimation strategy to recover the remaining parameters of the dynamic game, θ_0, can be outlined in two steps. First, I form moments using (i) first-order conditions on both markups and quality investment, (ii) envelope conditions on both observed and privately-known states, and (iii) Bellman equations evaluated at the firm's observed policy choices. To integrate these conditions with respect to rival private information, I replace their policies recovered in the first round of estimation in the firm moment conditions. Integration can be performed using standard approximation methods (e.g., Gauss-Legendre quadrature - see Judd (1998) and Miranda and Fackler (2002) for surveys) with respect to policy function residual distributions. Second, I replace firm value functions with sieves approximations and apply Ai and Chen's estimator. This approach has several advantages. First, forward simulation of firm continuation values given first step estimates is not necessary. That procedure is at the core of Bajari, Benkard and Levin's (2007) method, being also the main source of its computational burden. Second, unlike Bajari, Benkard and Levin, a flow payoff function linear in parameters is not necessary to ensure an affordable computational burden. θ_0 enters on both the flow payoff function and the approximating functions $\hat{h}$. As the latter are sieves approximating functions, the estimation problem effectively becomes a parametric one. This implies an estimation burden similar to minimum distance estimators (e.g. GMM). Finally, the estimates of $\hat{h}$ are useful for simulation purposes. Since $\hat{h}$ are value function estimates, these can be used for equilibrium computations in counterfactual policy experiments if no model parameters are changed. That is, the intermediate step of MPNE computation where the value function is computed can be avoided by treating the estimated value function as the true function. Moreover, if a researcher needs to compute the value function during a simulation routine (e.g., if the structural parameters of the game are changed), the estimated value function can still be useful for computation time reduction by using it as a good initial guess.

1.3.2. First-step estimates

Profit function parameters. There are two alternative estimation strategies to recover the profit function parameters. The first consists on the method introduced by Berry, Levinsohn and Pakes (1995). In this method, a framework which enables one to estimate demand and supply parameters for differentiated products markets is proposed. Like other related methods (e.g., Berry (1994)), it relies on the ability of observed product characteristics to explain consumer utility. The second alternative method is Nevo's (2001) estimator, which deals with cases where that ability is questionable by adding brand fixed effects. I follow Nevo's method to estimate demand parameters.

The dynamic model proposed in this Chapter assumes that the constant absolute markup charged in equilibrium by every firm is observable. However, only product prices are available in most datasets, not the charged markups. I deal with this problem by using the fact that prices must equal markup plus marginal cost. One may question how can one empirically separate markups from marginal costs. The identification relies on the assumption that, within each given period, all products that a firm commercializes are charged the same constant absolute markup. Under this assumption, at any given time t, all variation in prices across the firm's product portfolio must come from differences in products' marginal costs. Therefore, marginal costs and firm markups are identified.

I start by describing demand estimation. Recall that a product's market share conditional on the consumer's network type is given by

$$q_{jt}(p|D_{rt}) = \frac{\exp(\gamma X_{jt} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_{l} \exp(\gamma X_{lt} + \psi N_{d(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{d(l)} D_{rd(l)t})}$$

I further assume that $\xi_{jt} = \xi_j + \Delta\xi_{jt}$[5]. In addition, I decompose consumer r's utility on consuming product j at time t as

$$U_{rjt} = \delta_{jt} + \Lambda_{f(j)} D_{rtf(j)} + \epsilon_{rjt}$$

where $\delta_{jt} = \gamma X_j + \psi N_{f(j)t} - \alpha p_{jt} + \xi_j + \Delta\xi_{jt}$. is the mean valuation of product j at time t. Apparently, one should could consistently estimate the demand parameters by minimizing the distance between observed and predicted market shares. However, if there exists correlation of the error term $\Delta\xi_{jt}$ with other variables in δ_{jt}, this approach will yield inconsistent estimates. Even if suitable instruments are available, the fact that $\Delta\xi_{jt}$ enters the product market share function nonlinearly precludes this method. I instead follow the methodologies of Berry (1994), Berry, Levinsohn and Pakes (1995) and Nevo (2001), where this error term can be obtained by using the definition of δ_{jt} upon inverting the product market share functions. These are given by

$$\begin{aligned} q_{jt}(p) &= \int \frac{\exp(\gamma X_j + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_{l} \exp(\gamma X_l + \psi N_{d(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{d(l)} D_{rd(l)t})} \\ &\quad \times dP(D_{rt1}, ..., D_{rtF} | \mathbf{s}_t) \end{aligned}$$

where $P(D_{rt1}, ..., D_{rtF} | \mathbf{s}_t)$ corresponds to the distribution of consumer network types induced by the observed network states. In our model, this consists on a multinomial distribution, where the probability of drawing $D_{rtf} = 1$ is given by $\dfrac{N_{ft}}{M_t}$, $\forall f = 1, ..., F_t$.

[5]One could also assume a decomposition where time effects are present in ξ_{jt}. However, these are not separable from ξ_{0t}, and therefore not identified.

Upon getting the solutions $\delta_{jt}(S_{.,t}, \Lambda)$ to the implicit system of equations

$$q_{.t}(\delta_{.,t}; \Lambda) = s_{.,t}$$

where $s_{j,t}$ is the observed market share for product j at time t, the error term is given by

$$\Delta\xi_{jt} = \delta_{jt}(S_{.,t}, \Lambda) - (\gamma X_j + \psi N_{f(j)t} - \alpha p_{jt} + \xi_j)$$

The term $\Delta\xi_{jt}$ can be treated as econometric error term. However, this term should be correlated with prices, since manufacturers take into account all product characteristics in their pricing decisions. Therefore, valid instruments for both prices and some of the product's attributes should are necessary. Determining how relevant these endogeneity issues are for estimation is an empirical issue. The computation of $q_{.t}(\delta_{.,t}; \Lambda)$ requires integration with respect to each D_{rtf}. This can be done by replacing the integral by a sum across random draws from the joint distribution $P(D_{rt1}, ..., D_{rtF}|\mathbf{s}_t)$.

An important issue is how to empirically separate the time-invariant components of utility (i.e, γX_j and ξ_j). Nevo (2001) shows that both γ and ξ_j are identified. In addition, they can be recovered from the data by first replacing $\gamma X_j + \xi_j$ by brand-specific dummies d_j in the mean utilities δ_{jt}, and then regress the estimates $\hat{d}_j$ on the observed attributes X_j using Chamberlain's (1982) minimum-distance method. However, this method assumes that $E[\xi_j|X_j] = 0$ for every product j. This may fail to hold in some applications (see Chapter 2 for a concrete example). Again, these endogeneity concerns are an empirical matter. For this reason, if instruments are required for consistent estimation, the estimator for γ will be the two-stage least-squares estimator of

$$\hat{d}_j = \gamma X_j + \xi_j$$

In case no endogeneity concerns arise, estimates of γ can be recovered using a GLS regression. In any case, fixed-effect estimates are given by $\hat{\xi}_j = \hat{d}_j - \hat{\gamma} X_j$.

Letting λ denote the set of all demand parameters to be estimated, and defining $\Delta\xi(\lambda)$ be the vector of errors evaluated at a given value of λ, one can estimate the true value of λ with the GMM estimate

$$\hat{\lambda} = \arg\min_{\lambda} \Delta\xi(\lambda)' Z \Sigma^{-1} Z' \Delta\xi(\lambda)$$

where is a set of instruments orthogonal to $\Delta\xi(\lambda)$ and Σ is a consistent estimate of $E[Z'(\Delta\xi)(\Delta\xi)'Z]$.

Markups, marginal costs and quality investments. Markups and marginal costs can be empirically recovered as follows. By definition, we have

$$p_{jt} = Mkp_{f(j)t} + mc_{jt}$$

where $Mkp_{f(j)t}$ stands for the absolute markup charged by the firm which produces product j at time t. The fact that $Mkp_{f(j)t} = Mkp_{ft}$, $\forall j \in \mathcal{F}_{ft}$, $\forall f = 1, ..., F$, allows me to separate markups from marginal costs, since the latter depend of the specific attributes of product j. Assuming that marginal costs are given by $mc_{j,t} = mc_0 + mc_t(X_{j,t}) + \eta_{jt}^{MC}$, markups can be recovered from the residuals of the regression equation

$$p_{jt} = mc_0 + mc_t(X_{j,t}) + m \cdot B_{jt} + \eta_{jt}^{Mkp}$$

where $B_{jft} = \mathbf{1}\{j \in \mathcal{F}_{ft}\}$, $B_{jt} = [B_{j1t}, ..., B_{jFt}]^T$ and m is a $1 \times F$ vector of coefficients. The latter measure the markup charged by firm f at time t.

After recovering marginal cost function, one can compute quality investment $A_{f,t}$ by taking the following steps:

- For each firm f, check the set of firm products at $t+1$, $\mathcal{F}_{f,t+1}$, and verify which products do not belong to $\mathcal{F}_{f,t}$. Denote this set of products as $\mathcal{F}_{f,t}^{IN}$.
- If $\mathcal{F}_{f,t}^{IN} \neq \varnothing$, use demand and marginal cost function estimates to compute

$$\left(\sum_{l \in \mathcal{F}_{ft}^{IN}} \exp(\gamma X_{lt} - \alpha mc_{lt} + \xi_{lt} - \xi_{0t}) \right)$$

- Verify which products belong to $\mathcal{F}_{f,t}$ but not to $\mathcal{F}_{f,t+1}$. Denote this set of products as $\mathcal{F}_{f,t}^{OUT}$.
- If $\mathcal{F}_{f,t}^{OUT} \neq \varnothing$, use the demand and marginal cost function estimates to compute

$$\left(\sum_{k \in \mathcal{F}_{ft}^{OUT}} \exp(\gamma X_{k,t} - \alpha mc_{k,t} + \xi_{k,t} - \xi_{0,t}) \right)$$

- Compute $A_{f,t}$ by using the formula provided in the previous section.

For the second stage of estimation, policy function estimates are necessary. These can be obtained by projecting each $A_{f,t}$ and $Mkp_{f,t}$ on the space of observed states, $\mathbf{s}_t$. The estimation methods to obtain these best-reply functions will depend on the specific application at hand. One example can be found in Chapter 2 in the context of the supercomputer industry. I proceed by assuming that the researcher has estimated the functions $A_{f,t}(\mathbf{s}_t, \hat{\varepsilon}_{f,t}^{A})$ and $Mkp_{f,t}(\mathbf{s}_t, \hat{\varepsilon}_{f,t}^{MKP})$ from the data, where $\hat{\varepsilon}_{f,t}^{A}$ and $\hat{\varepsilon}_{f,t}^{MKP}$ are the policy estimation residuals.

1.3.3. Second-step estimates

The first step of estimation recovers all but the quality investment cost parameters. I use the sieve minimum distance (SMD) estimator of Ai and Chen (2003) to obtain estimates on these parameters. For ease of notation, let $\Delta(S)$ denote the dimension of the observed state space, and g_k as the k^{th} entry in the observed state vector at period $t+1$. For the necessary moment conditions, it is useful to start with the first-order condition on quality investment. However, I must account for the fact that the support of this variable is bounded below by -1. I use the Karush-Kuhn-Tucker condition (see Miranda and Fackler (2002) pp 191-193) to derive the first-order condition. This corresponds to

$$\int_{\varepsilon_{-f}} \left(-\frac{\partial C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft}; \theta))}{\partial A_{ft}} + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial A_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu_f^A \mathbf{1}\{A_{ft} = -1\} = 0$$

where μ_f^A is a nonpositive Lagrange Multiplier. Letting $C_1(A_{ft}, \mathbf{s}_t; \theta)$ and $C_2(A_{ft}, \mathbf{s}_t; \theta)$ be such that

$$\frac{\partial C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{ft}; \theta)}{\partial A_{ft}} = C_1(A_{ft}, \mathbf{s}_t; \theta) + C_2(A_{ft}, \mathbf{s}_t; \theta)\varepsilon_{ft}$$

then the Karush-Kuhn-Tucker condition on quality investment simplifies to

$$\frac{\int\limits_{\varepsilon_{-f}} \left(-C_1(A_{ft}, \mathbf{s}_t; \theta) + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial A_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu_f^A \mathbf{1}\{A_{ft} = -1\}}{C_2(A_{ft}, \mathbf{s}_t; \theta)} = \varepsilon_{ft}$$

Since the right-hand side of the equation has expected value zero, one can consider the expectation of the left-hand side as a moment condition. For this purpose, I replace (i) $F(\varepsilon_{-f})$ by the distributions of residuals of first-stage policy estimates and (ii) firm f's integrated value function by the approximant described below.

A moment condition involving the first-order condition in absolute markup can be derived in a similar way. This condition can be written as

$$\int\limits_{\varepsilon_{-f}} \left(\frac{\partial \Pi_{ft}(\mathbf{Mkp}_{ft}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t))}{\partial Mkp_{ft}} + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial Mkp_{ft}} \right] \right) dF(\varepsilon_{-f}) = 0$$

Again replacing the integrated value function V_f by its approximant, and integrating with respect to ε_{-f} using the best-reply functions estimated in the first-stage, a second moment condition can be used for estimation.

Bellman equations can also be used as moments, upon (i) replacing the firm's integrated value function by a sieves approximation h_0 and (ii) removing the max operator by replacing the firm's optimal policies by its observed values in the data. I define the integrated value function approximant by the tensor product

$$h_f(\mathbf{s}_t, \theta) \equiv \sum_{k_1=0}^{K} ... \sum_{k_{\Delta(S)}=0}^{K} c_{f,k_1,...,k_{\Delta(S)},k_{\varepsilon f}}(\theta) T_{k_1}(s_1) \times ... \times T_{k_{\Delta(S)}}(s_{\Delta(S)})$$

where $T_k(.)$ is an univariate first-kind Chebyshev polynomial of order k. Following Miranda and Fackler (2002), this product can be written as

$$h_f(\mathbf{s}_t, \theta) = [T_{\Delta(S)}(s_{\Delta(S)}) \otimes T_{\Delta(S)-1}(s_{\Delta(S)-1}) \otimes ... \otimes T_1(s_1)] c_f(\theta)$$

where c_f is a $K^{\Delta(S)} \times 1$ column vector of coefficients and $T_i(.)$ is a $1 \times K$ row vector. Using the more compact notation $h_f(\mathbf{s}_t, \theta) = T_f(\mathbf{s}_t) c_f(\theta)$, the following holds after integrating the firm's problem with respect to its private information:

$$\begin{aligned} T_f(\mathbf{s}_t) c_f(\theta) &= \int_{\varepsilon} (\pi(\sigma_f(\mathbf{s}_t, \varepsilon_f), \sigma_{-f}(\mathbf{s}_t, \varepsilon_f), \mathbf{s}_t, \varepsilon_f; \theta) \\ &\quad + \beta E\left[T(\mathbf{s}_{t+1}) c_f(\theta) | \sigma_f(\mathbf{s}_t, \varepsilon_f), \sigma_{-f}(\mathbf{s}_t, \varepsilon_f), \mathbf{s}_t, \varepsilon_f; \theta\right]) dF(\varepsilon) \end{aligned}$$

As before, one may form a moment condition by using estimated policies and its residuals to integrate out private information. Moreover, the private information term ε_f can be replaced by the formula derived for the first moment condition.

Finally, one can use envelope conditions on the states from firm's problem as extra moments. The latter can be derived using the methods of Miranda and Fackler (2002). For each component of the state vector $\mathbf{s}_{i,t} \in \mathbf{s}_t$ (e.g., N_{0t}) , we have

$$\int_{\varepsilon} \left(\frac{\partial \pi(\sigma_f(\mathbf{s}_t, \varepsilon_f), \sigma_{-f}(\mathbf{s}_t, \varepsilon_f), \mathbf{s}_t, \varepsilon_f; \theta)}{\partial \mathbf{s}_{i,t}} + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial \mathbf{s}_{i,t}} \right] \right) dF(\varepsilon) = \frac{\partial V_f(\mathbf{s}_t)}{\partial \mathbf{s}_{i,t}}$$

Like in other moment conditions, I replace integrated value functions and unobserved private values to form an additional moment condition. If the total number of moment conditions is not enough to identify all the parameters in θ, one can use the fact that the error terms in the equations outlined above are assumed independent of observed states to form extra moments For example, one can consider Chebyshev polynomials of observed states as instruments.

At this point, the researcher is left to apply Ai and Chen's (2003) SMD method. However, one must account for the fact that the second step of estimation takes first-step estimates as given when computing standard errors. This can be accomplished by either (i) using the correction methods of Murphy and Topel (1985) or (ii) bootstrapping samples of the data with replacement and estimating the dynamic parameters with each subsample.

1.4. Concluding remarks

This Chapter proposes a dynamic structural model of competition where firms invest in product quality and set product markups. Unlike other studies in the switching costs literature, the proposed framework accounts for the influence of firm network size on both consumer utility and the distribution of switching costs across consumers. Strong assumptions imposed by static competition models with switching costs and network effects are avoided by building on recent refinements to the dynamic oligopoly games literature. Firms will account for network sizes when computing optimal strategies, and therefore competition models with switching costs ignoring network effects may lead to misleading conclusions. The difficulty of modeling pricing decisions in the presence of new product launch and exit is circumvented by restricting the firms' pricing strategies to constant absolute markups for each product in the firm's portfolio. This restriction can be weakened at the cost of increasing each firm's state space. The model can be estimated by combining the methods of Bajari, Benkard and Levin (2006) and Ai and Chen (2003). This hybrid methodology offers several computational advantages in terms of parameter estimation and simulation of counterfactuals.

Several important questions raised in economic policy debates could be addressed with the model. For example, one could simulate the consequences of a given subsidy scheme for product quality investment. One could also assess what market structures would yield maximal welfare. Another application of interest would be the measurement of both network externalities and switching costs effects on competition and welfare. Another issue of particular interest is the impact of mergers on welfare and technological progress dynamics. I address some of these questions on the remaining chapters in the context of the supercomputer industry.

CHAPTER 2

Extending the Frontier: A Structural Model of Investment and Technological Competition in the Supercomputer Industry

2.1. Introduction

Technological innovation plays a central role in economic growth. However, there is no consensus on whether competition helps technological innovation. In addition, it is hard to quantify the gains (or losses) from increased competition in high-technology industries, particularly when network effects and switching costs are present. In this Chapter, I aim to fill this gap. I focus on one high-technology industry - supercomputers - and evaluate the evolution of the best computational speed available in the supercomputer market under three different market structures: monopoly, duopoly and a three-firm market. In addition, I quantify the temporal evolution of welfare under these three different industry sizes.

I propose a structural model where firms set prices, invest in product quality, and innovate on their most advanced technology. Using data from the supercomputer industry, I estimate the model parameters with a variant of the two-step method of Bajari, Benkard and Levin (2007). I then use the estimated parameters to simulate the outcomes of three different market structures. This is done in two steps: (i) computation of the Markov-Perfect Nash Equilibrium (MPNE) of the game under monopoly, duopoly and three-firm market, and (ii) comparison of total welfare and the maximal available computational speed in the supercomputer industry implied by these industry structures for a period of five consecutive years. I also examine the impact of the threat of entry on a monopolist's innovating behavior. This is also done in two steps: (i) simulation of maximum computing speed when only one firm can operate in the market, and (ii) comparison of the results with the ones from a monopolist facing free entry.

I find that increased levels of competition have positive effects on the evolution of supercomputer technology. In line with Aghion, Harris, Howitt and Vickers (2001), increases in market competition induce more innovation due to the "selection effect". That is, the incremental payoff from innovating is higher when a firm is in "neck-and-neck" competition with technologically similar rivals. Firms therefore extend their technological frontier primarily to escape competition with "neck-and-neck" rivals. I also find that increased competition increases welfare in the supercomputer market, even though the marginal increase in welfare is decreasing in the number of competitors. Finally, I find that fixed innovation costs are the most important component of technology investment expenditures. Both the firm and its rivals' technological states impact fixed costs considerably. In particular, this component of the innovation costs is decreasing in the aggregate frontier of a firm's most advanced rivals.

My model builds on the theoretical model of Ericson and Pakes (1995) and extends the framework developed in Chapter 1. The inclusion of frontier innovation in the firms strategy sets raises the problem of how to deal with nonstationary states and controls. I follow the methods of variable scaling in the growth literature surveyed in Stockey and Lucas (1989) to guarantee stationary measures of firm technological state and innovation investment. A second difficulty is how to deal with multiproduct firms. I build on the model proposed by Nevo and Rossi (2007), in which a markup-adjusted inclusive value is included as state variable. This allows me to control for product market profits while keeping a computationally manageable state space dimension.

To estimate the model I use a variant of the two step method proposed by Bajari, Benkard and Levin (2007). In the first step, I recover the profit function parameters from demand and supply estimation. In addition, I estimate policies and state transitions nonparametrically as functions of observable states. In the second step, I solve for the continuation values of firms using function approximation methods (see Miranda and Fackler (2002)) and the policy and transition estimates from the first stage. Instead of using the continuation values in the inequality sampling method proposed by Bajari et al. (2007), I include them in a GMM estimation procedure where I use firm optimality conditions directly. I am able to separate the benefits of frontier extension from its costs by using the variations on the number of a firm's technologically similar rivals. According to Aghion, Harris, Howitt and Vickers (2001) and Aghion, Bloom, Blundell, Griffiths and Howitt (2003), increases in market competition induce more innovation due to a "selection effect". That is, the incremental payoff from innovating is higher when a firm is in "neck-and-neck" competition with technologically similar rivals. Firms therefore innovate on their most advanced technology primarily to escape competition with "neck-and-neck" rivals. A higher number of similar rivals does not impact the costs of innovation but increases the incremental payoff from innovating due to the "selection effect". Therefore, the positive correlation between investment rates and number of technologically similar rivals successfully identifies the benefits from expanding the technological frontier.

The understanding of the forces driving innovation and the measurement of competition effects require focus on a specific industry. Ideally, one would want to analyze an industry whose technological advances influence the evolution of other industries' technologies. Bresnahan and Trajtenberg (1995) document 'too late, too little' innovation in industries producing this type of technology - named General Purpose Technology (GPT). Moreover, the positive feedbacks of advances in GPTs are usually dispersed through the economy, and can therefore work as a trigger for economic growth. One of these industries, supercomputers, is particularly well suited to study the technological evolution patterns under different market structures for several reasons. First, there is a standard measure for supercomputer performance - the LINPACK Benchmark. A supercomputer, commonly mentioned as High Performance Computer (HPC), is a general purpose computer that is faster than commercial competitors and has sufficient central memory to compute problem sets of general scientific interest. The technological state of a firm in this industry can be measured by the maximal supercomputing speed that the firm has ever produced. Second, the technological evolution patterns of the firms in this industry exhibit evidence of "racing behavior", where players aim to leapfrog opponents as documented by Khanna (1995). Therefore, the supercomputer industry provides an almost ideal ground to study innovation behavior under oligopoly. Third, supercomputer technology brings several important benefits that are only partially considered by manufacturers, thus justifying governmental intervention. Supercomputers are crucial to advances in many sciences and play an important role on the production of other goods (e.g. cars, airplanes, medicines, intelligence creation and national defense).

A recent body of the industrial organization literature studies the incentives for technical progress in oligopolistic industries. (e.g., Bresnahan and Greenstein (1999), Athey and Schmutzler (2001), Aghion et al. (2001, 2003), Scotchmer (1996), Evans and Schmalensee (2002)). More recently, Segal and Whinston (2004) analyze the effects of specific antitrust policies in industries where innovation is of central importance to competition. Most of these studies restrict attention to duopoly games, or abstract from important dynamic strategic considerations of players (e.g. dependence of innovation decisions on the firms' technological states). I introduce a model where such restrictions are relaxed.

The rest of the Chapter is organized as follows. In section 2.2 I describe the supercomputer industry, mentioning its key differences compared to other high-end computers. In Section 2.3 I present the dynamic oligopoly model, describing firm behavior on pricing and innovation investment. The details about the data and the empirical strategy for model estimation are presented in section 2.4. Section 2.5 addresses the estimation results, while section 2.6 describes both the policy experiments and the simulation results. Section 2.7 concludes with results and contributions summary, as well as possible extensions for the paper. All derivations and computational details can be found in Appendix B.

2.2. The supercomputer industry

A supercomputer refers to "those computing systems (hardware, systems software, and applications software) that provide close to the best currently achievable sustained performance on demanding computational problems"[1]. Supercomputers are very expensive durables whose expected useful lifetime is five years. Also known as high-performance computer (HPC), its key components are processors, memory and interconnect bandwidth.

[1] Source: "Getting Up to Speed: the Future of Supercomputing" (2005).

The processors are assigned the tasks of performing the instructions programmed in the supercomputer (e.g., simulation of a nuclear explosion). In order to temporarily store data or results for processing in intermediate program instructions, all supercomputers must contain a memory sector which must be connected with the processors. An interconnect bandwidth is required to control the traffic of information between the memory sector and the processors. A supercomputer's computing speed is measured in total floating-point operations per second (FLOPS). Even though there are several benchmarks for speed measurement of a supercomputer, the consensus measure is the LINPACK Benchmark. This is the maximum possible computation speed that can be achieved by the supercomputer when it is instructed to solve a system of linear equations with 1000 equations and 1000 unknowns[2].

Despite their similarities, there are two main differences between a supercomputer and a mainframe. The first is that, upon delivery and installation, the supercomputer is configured for selected purposes, channeling all its power to execute these tasks as fast as possible. High performance computers are assigned very specific tasks (e.g., simulation of car accidents for new vehicle models, simulation of nuclear explosions) to be performed one at a time, whereas a mainframe uses its power to performs several programs simultaneously. The second main difference between a supercomputer and a mainframe is the system architecture. High performance computers are designed to execute specific programs as efficiently as possible, implying a deviation from the regular mainframe architecture designed for simultaneous task processing.

[2]A comprehensive discussion on LINPACK Benchmark's technical details can be found in Dongarra (2006), and Dongarra, Luzczek and Petitet (2001).

There are two main types of supercomputer architectures: vector systems (VS) and massive parallel processing systems (MPP). The vector systems are characterized by special central processors (which is typically produced by the manufacturing firm), along with built-in memory chips and interconnect bandwidth. Processing takes place by seeding in several tasks at the same time. The processor is specially designed to interpret the vector of tasks and to perform instructions simultaneously. This technique, initially proposed to the late supercomputer engineer Seymour Cray, is known as vector processing. The VS are known for its computing efficiency (i.e., high computation speed due to optimized combination of processors, memory and bandwidth), as well as for acceptable installation area requirements, reasonable power consumption and maintenance costs. However, it is usually expensive compared to MPPs. This is primarily due to the high costs required to build special purpose processors.

This description sharply contrasts with the one for MPP. Instead of special purpose processor, manufacturers of MPPs use off-the-shelf processors to build a processing sector for the machine. A similar procedure is followed for building the memory and bandwidth sectors of the supercomputer. As such, the production of an MPP is not as demanding in R&D efforts as for VS. For this reason, this type of system is usually not as expensive as VS. However, MPPs have its own caveats. First, it is not as efficient as VS. It relies on aggregating processors along with memory and bandwidth components, while in VS these components are conceived in order to maximize computing efficiency. Second, despite its relatively cheap price, MPPs require far more installation space, maintenance costs and power consumption.

Despite these disadvantages, MPPs have become popular in supercomputing since the early 1990's for two reasons. First, the prices of off-the-shelf microprocessors have fallen dramatically over time, while its processing power has steadily increased over time. Second, despite the fact that progress on bandwidth and memory technology was not as remarkable as for processors, there was enough scope for increases in FLOPS by adding extra processors, for a given combination of memory and bandwidth. This advantage over VS is no longer available, as the current pace of memory and bandwidth advances can no longer sustain gains in speed from increasing processor power. The most up-to-date supercomputer models aim to combine the advantages of both types of architectures.

The supercomputer industry is characterized by few units being sold and considerable revenues. The joint revenue of the top nine competitors in this industry was $4655 million in 1997, reaching its peak in 2000 with $6083 million[3]. The vast majority of buyers are governmental institutions, which account for at least 70% of total supercomputer sales since 1990. Since the rise of this industry in 1953, several firms have entered and exited the market. Most of the entry and exit is due to short-lived competitors, while most of the important players in this market remained active. The most relevant competitors who kept their importance over time have been Cray Inc., Hewlett-Packard, IBM, NEC, Fujitsu, Hitachi, SGI, and Sun Microsystems. The benefits of being among the best HPC suppliers are not restricted to product market profits. Innovations in supercomputer technology are often used for improvements in other high-tech products the firm may be producing. For example, advances in computational speed for HPC have been incorporated in personal

[3]Source: IDC 2002

computers during the last ten years. Except for Cray, all other supercomputer sellers have other product lines.

Table 2.1 provides yearly statistics on the maximum computing speed ever produced (henceforth *firm technological frontier*) for the most important firms in the market. Several important aspects can be inferred. First, significant increases on the maximal computing speed in the industry tend to be followed by periods of no expansion. In general, the more significant this expansion is, the more years are necessary to beat that record. Second, the data indicates that once a firm reaches quality leadership, it will not expand its frontier significantly (if at all) unless a rival leapfrogs upon it. Therefore, the incentives on firms to invest on frontier expansion are contingent on the firms' frontier position. Finally, firms tend to expand their frontiers more significantly over time. This suggests that firms can benefit from the technological advances of their rivals. Hence, since the technological advances of a firm leverage rival innovations, firms are likely to underinvest.

The scope of application of supercomputers is immense. According to the classification available at the supercomputing rating organization TOP500, there are 26 application areas for supercomputers. Some representative examples application areas are Aerospace, Automotive, Finance, Defense, Geophysics, Semiconductors, Weather and Climate Research, and Telecommunications.

Due to the fact that supercomputers are extremely expensive durables (costing typically millions of dollars), buyers usually acquire at most one unit per year. Potential buyers may either choose among existing models (henceforth *"off-the-shelf" HPCs*) or order a machine whose computing power is suited for her specific needs (henceforth *custom supercomputers*). There are some differences between buying an existing model and

Table 2.1. Technological Frontiers of the most important firms (in LINPACK Benchmark Giga-FLOPS)

Year	Cray	HP	IBM	NEC	Fujitsu	Hitachi	SGI	Intel	Sun	Ind. Max.
1990	2.171	-	0.54	23.2	4	1.817	-	2.6	-	23.2
1991	2.171	-	0.54	23.2	4	1.817	-	13.9	-	23.2
1992	13.7	-	1.457	23.2	4	1.817	-	15.2	-	23.2
1993	13.7	1.6	5.8	23.2	124	7.4	1.284	143.4	-	143.4
1994	100.5	3.306	66.3	23.2	124	27.5	4.142	143.4	-	143.4
1995	100.5	7.408	88.4	60.72	124	28.4	26.653	143.4	17.91	143.4
1996	341	15.01	88.4	66.53	229	368.2	341	143.4	17.91	368.2
1997	815	51.3	151.8	244	229	368.2	815	1338	26.45	1338
1998	891	51.3	547	244	229	368.2	891	1338	272.1	1338
1999	1166	189.3	2144	244	492	873	1166	2379	420.44	2379
2000	1166	189.3	4938	303	886	1035	1166	2379	420.44	4938
2001	1166	431.7	4938	1192	886	1709.1	1166	2379	420.44	4938
2002	1166	2916	7304	35860	5406	1709.1	1166	2379	1226.4	35860
2003	2932.9	8633	7304	35860	5406	1709.1	11652	2379	1226.4	35860
2004	5895	8633	70720	35860	8728	3319	51870	2379	1439	70720
2005	36190	8633	280600	35860	8728	3319	51870	2379	3146	280600

ordering a custom supercomputer. If the consumer is interested on ordering a custom model, potential manufacturers are called to submit proposals, indicating the prototype details and the price to be charged. Upon testing the prototypes, the buyer chooses the supplier and makes a public announcement of the winning proposal details. For the case of off-the-shelf models, the potential buyer contacts the supercomputer suppliers, in an attempt to obtain discounts over the list prices. After checking if the available models suit his computing needs, the consumer decides whether to place an order. Discounts over list price are frequent, but depend on the client characteristics. According to the available information on discounting practices in this industry, discounts for private and

Table 2.2. Network statistics for the most important firms

year	Cray	HP	IBM	NEC	Fujitsu	Hitachi	SGI	Intel	Sun	Total Networks
1990	22	0	2	9	9	9	0	1	0	56
1991	52	0	3	10	19	9	0	11	0	121
1992	76	0	4	17	48	11	0	15	0	200
1993	125	0	5	25	56	12	0	27	0	329
1994	165	3	20	30	69	19	9	38	0	504
1995	161	14	82	30	63	20	115	60	0	689
1996	156	24	139	31	61	21	217	43	1	815
1997	178	41	177	50	73	26	227	37	0	927
1998	229	93	177	56	73	26	267	35	142	1212
1999	162	50	263	50	71	27	249	32	187	1200
2000	145	119	311	60	77	29	248	32	247	1376
2001	131	89	402	57	74	32	213	27	321	1459
2002	124	255	392	51	77	35	213	25	205	1494
2003	112	293	452	56	72	33	200	23	278	1655
2004	123	333	473	59	70	32	199	23	217	1684
2005	106	467	571	56	68	32	188	23	204	1876

non-government clients are in the range of 10%-20% over the list price. Government-related institutions are typically given a discount rate between 20%-30%[4]. Like other technological goods, supercomputers have experienced sharp decreases in both nominal and real price/quality ratio.

Table 2.2 describes the evolution of firm networks for the most important players in the industry in the period 1990-2005. A firm network is here defined as the number of consumers whose last purchase was a product from that firm[5]. Several relevant facts can be inferred from Table II. First, firms entering the market tend to considerably expand their networks shortly after entry occurs. In this market, this seems explained by "market

[4]As an example, the latest available contract between Cray and the General Services Administration (GSA) specified a discount rate of about 22% for several supercomputer models and related components.
[5]Network values for the early years in the sample were computed using the few existing information on purchases prior to 1990. This information is available from TOP500 early surveys.

tipping" behavior from entrants (e.g., SGI, Sun Microsystems) in an attempt to create a significant installed base via low prices.

Second, firms tend to have less volatility in their network size once it has reaches a "reasonable" dimension. Given that the total number of consumers being part of a network has been growing since 1990 (see last column), this suggests that firms seek to create a consumer loyal base which can be harvested in future years by charging higher prices. Finally, firms fiercely dispute their network size with their opponents. Since there is imperfect lock-in in this market (i.e., consumers may change supplier by incurring in switching costs), firms may not only attract buyers who never purchased supercomputers but also dispute the customers of their rivals. This hints for the importance of network size in consumer purchasing decisions. That is, firms will dispute their rivals' networks in an attempt to decrease their products' purchase probabilities. Network size should therefore be a key driver of firm market shares.

2.3. The model

In this section I extend the model presented in Chapter 1 to describe firm behavior in the supercomputer industry. Time is assumed discrete with an infinite horizon, and indexed by $t \in 1, 2, ..., \infty$. Letting F be the maximum number of firms that can operate in this market, I denote $F_t^A \leq F$ the number of incumbent firms at time t. An incumbent firm is defined as every player with at least one product for sale to consumers. At each period, incumbents simultaneously decide on retail prices, quality investment and frontier expansion. Firm quality is defined as the markup-adjusted expected utility from buying from that firm. As in Chapter 1, quality investment is defined as the percentage increase (or decrease) in that metric to the extent controlled by the firm. Frontier investment is defined as being the number of computing speed units (measured in Gigaflops according to the LINPACK Benchmark) added to the maximal computing speed that the firm has ever produced (henceforth *firm frontier*). Firms can exit the market by scrapping all their existing products and introducing none. Empty firm slots are taken by potential entrants, whose entry decision is assumed exogenous and dependent only on observed states.

Observable states. Following the framework developed in Chapter 1, I assume that all payoff-relevant features of firms can be encoded into a state vector. All firms observe the number of consumers who never purchased a supercomputer, denoted N_{0t}, and firm-specific states: κ_{ft}, N_{ft}, ν_{ft} and Ξ_{ft}, for all $f = 1, ..., F$ players and period $t \in 1, 2, ..., \infty$. N_{ft} (henceforth *firm network*) represents the number of consumers whose last purchase was from firm f. ν_{ft} denotes Nevo and Rossi's (2007) markup-adjusted expected utility from buying from firm f. The latter is a function of all the firm's products characteristics, which are jointly denoted by Ξ_{ft}. κ_{ft} (henceforth *firm frontier rank*) represents the ratio of firm f's frontier over the best frontier available in the industry at time t. Ideally, I would like to consider the firm frontier itself as state variable. Unfortunately, the fact that this variable grows without bound makes its choice as state variable impractical. Instead, I consider a scaled version of this variable, κ_{ft}, by following the variable rescaling methods surveyed by Stockey and Lucas (1989)[6]. Letting h_{ft} be the maximal computing speed that the firm has ever produced up to time t, κ_{ft} is defined as

[6]As explained later in the paper, this stationary scaled variable embodies the dependence of firm frontier on innovation.

$$\kappa_{ft} \equiv \frac{h_{ft}}{\max\limits_{i=1,...,F} \{h_{it}\}}$$

I denote $\tilde{\mathbf{s}}_t \in \tilde{S}$ as the vector of all observable states at time t.

Incumbent firms. Let $\mathbf{p}_{ft}$ denote the vector of prices charged by firm f at time t for its products. In addition, let I_{ft} denote the frontier investment rate of firm f, defined as the ratio between investment and the firm frontier at time t[7]. I denote A_{ft} as the quality investment of firm f. At the time of the investment and pricing decisions, each firm observes a pair of private information shocks: a shock in frontier extension costs, ε_{1ft},, and a shock in quality investment costs, ε_{2ft}. Before defining each firm's intertemporal optimization problem, it is convenient to use the framework developed in Chapter 1 to further simplify the model. Let Mkp_{jft} be the equilibrium absolute markup that firm f charges for its product j at time t. In addition, define $\mathcal{F}_{ft}$ as the set of product firm f commercializes at time t, and denote mc_{jt} the marginal cost of producing product j at time t. Under the conditions on product demand proposed by Nevo and Rossi (2007) - discussed in the demand section of this Chapter - the following assumption holds:

Assumption A1 (Constant absolute markup per firm): In equilibrium, each firm f restricts attention to constant markup strategies Mkp_{ft} for each product $j \in \mathcal{F}_{ft}$. That is, $p_{jt} - mc_{jt} = p_{lt} - mc_{lt}$, $\forall j, l \in \mathcal{F}_{ft}$, $\forall f = 1, ..., F_t^A$.

[7]The motivation for choosing investment rates rather than levels as a firm control is that the data on firm maximal computing speed increases is nonstationary. Firms tend to expand their frontiers more significantly over time. Considering investment rates rather than investment levels successfully circumvents this problem.

This assumption aims to restrict the firm strategy set to allow for a varying number of supercomputer models in the firm's product portfolio. However, this assumption may not be necessary in industries where neither the number of firm products changes nor pricing strategies affect state transitions[8].

Before defining each firm's intertemporal optimization problem, I impose the following assumption:

Assumption A2 (Markovian pure strategies): In equilibrium, all players' choices are deterministic functions of payoff-relevant information.

Formally, this corresponds to a mapping between the state space and firm actions. Assumption A1 suggests that product-specific attributes should not play a role on the players' strategy functions. Therefore, I assume that $\{\Xi_{ft}\}_{f=1}^{F_t^A}$ are not payoff-relevant information and therefore removed from the set of observed states. Hence, the relevant set of observed states is given by $S \equiv \left(N_{0t}, \{\kappa_{ft}, N_{ft},\ \nu_{ft}\}_{f=1}^{F}\right)$.

Under these assumptions, an incumbent's state space is $S \times \Gamma_{1f} \times \Gamma_{2f}$, where Γ_{if} corresponds to the space of realizations of ε_{ift}. In what follows, I assume $\Gamma_{if} = \mathbb{R}, \forall f = 1, ..., F_t^A$ and $\forall i = 1, 2$. Under assumptions A1 and A2, a player's strategy is defined by the map $\sigma_f : S \times \Gamma_{1f} \times \Gamma_{2f} \longrightarrow (\mathbf{Mkp}_f, A_f, I_f)$. Assumption A1 implies that one can replace prices by an absolute constant markup in the firm's problem. Therefore, an incumbent's flow payoff can be written as

[8]However, even in this case assumption A1 may hold as a result of spot price competition. See Nevo and Rossi (2007) for details.

$$\pi_f(\mathbf{Mkp}_t, A_{ft}, I_{ft}, \mathbf{s}_t, \varepsilon_{ft}) \equiv \Pi_f(\mathbf{Mkp}_t, \mathbf{s}_t) + \Upsilon_f(\mathbf{s}_t) - C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}) - C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft})$$

where $\mathbf{Mkp}_t \equiv \{\mathbf{Mkp}_{ft}\}_{f=1}^{F_t^A}$ is a vector of markups charged by the F_t^A firms, $\Pi_f(\mathbf{Mkp}_t, \mathbf{s}_t)$ corresponds to firm f's flow market profits, and $\Upsilon_f(\mathbf{s}_t)$ is a function representing the benefits (other than flow profits) for firm f from the current state of the industry $\mathbf{s}_t$. As described in more detail below, these benefits outside profits aim to control, for example, for spillovers of frontier expansions to other lines of product within the firm. $C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft})$ represents frontier extension costs, while $C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft})$ represents costs of quality investment.

Incumbents are assumed to decide quality investment, frontier expansion and absolute markups in order to maximize the expected sum of discounted payoffs. Neither the researcher nor individual firms observe the private shocks of other firms. A firm's intertemporal optimization problem can be written in recursive form, for any Markovian profile of strategies $\sigma = (\sigma_1, ..., \sigma_F)$ of Markovian strategies.. Therefore, the Bellman equation of the incumbent firm is

$$\begin{aligned} V_f(\mathbf{s}, \varepsilon_f) &= \int_{\varepsilon_{-f}} \underset{\sigma_f(\mathbf{s},\varepsilon_f)}{Max} \{\pi_f(\sigma_f(\mathbf{s}, \varepsilon_f), \sigma_{-f}(\mathbf{s}, \varepsilon_{-f}), \mathbf{s}, \varepsilon_f) \\ &\quad + \beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}} \left[V_f(\mathbf{s}', \varepsilon_f') | \mathbf{s}, \sigma_f(\mathbf{s}, \varepsilon_f), \sigma_{-f}(\mathbf{s}, \varepsilon_{-f}) \right] \} \, dF(\varepsilon_{-f}) \end{aligned}$$

where $\sigma_f(\mathbf{s}, \varepsilon_f) \equiv (Mkp_f(\mathbf{s}, \varepsilon_f), I_f(\mathbf{s}, \varepsilon_f), A_f(\mathbf{s}, \varepsilon_f))$ represents a Markovian strategy for firm f, and $E_{\mathbf{s},\sigma_f,\sigma_{-f}}[.]$ denotes firm f's expectations conditional on all firms choosing Markovian strategies, and on observable states.

Potential entrants. Potential entrants are defined as firms are defined as players with no products for sale. These players are assumed to decide whether to enter the market. In case entry takes place, the potential entrant will have both a set of products for sale and a technological frontier in the following period. I assume that a potential entrant decides on entry after observing a pair of shocks which determine future average quality and technological frontier. Thus, a potential entrant's state space is assumed to be $S \times \Gamma_{3f} \times \Gamma_{4f}$, where Γ_{3f} corresponds to the shock determining future average quality and Γ_{4f} is the shock that determines the entrant's future technological state. These shocks are assumed private information of the entrants. It is assumed that potential entrants get a payoff of zero by the time they decide on entry.

2.3.1. Product Market Profits

In this subsection I model flow profits in the supercomputer market. The modeling of product market profits proceeds in two steps. First, I model supercomputer demand. Second, I assume price competition in differentiated products to model the supply side. Finally I derive each firm's market profit as a function of observed states and absolute markups.

Demand. I follow the framework developed in Chapter 1 to control for consumer heterogeneity. Let J_t be the number of supercomputer models available at time t, and define D_{rft} as a dummy variable which equals one if consumer r is part of firm f's network and is zero otherwise. The utility function for a buyer r interested in acquiring supercomputer j at time t is assumed to be[9]

$$U_{rjt} = \gamma x_j + \lambda MPP_j + \tau_1 \kappa_{f(j)t} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} + \Lambda_{f(j)} D_{rtf(j)} + \epsilon_{rjt}$$

where x_j is the observed quality of product j, defined as the computing speed measured in Teraflops (1000 Gigaflops) according to the LINPACK Benchmark (*henceforth* Rmax), MPP_j is a dummy variable which equals unity if supercomputer j belongs to the family of massive parallel processing systems, p_{jt} is the price of supercomputer j at time t, ξ_{jt} regards its unobserved quality, $N_{f(j)t}$ corresponds to the number of consumers at time t whose last purchase was from $f(j)$ (i.e. the firm which commercializes product j), $D_{rf(j)t}$ is a dummy variable which equals one if consumer r is part of $f(j)$'s network, $\kappa_{f(j)t}$ corresponds to the computing speed ranking of the firm which commercializes product j at time t[10], and ϵ_{rjt} is a zero-mean stochastic term. I assume that $D_{rf(j)t}$ is not observed by the econometrician, but its distribution is. τ_1, ψ and $\Lambda_{f(j)}$ are parameters which measure valuation of technology ranking, network externalities and switching costs associated with the firm which produces product j, respectively.

[9] I am implicitly assuming that buyers will purchase at most one supercomputer. An alternative modeling strategy would be to account for multiple purchases as proposed by Hendel (1999). Except for some selected customers, the TOP500 data indicates that the vast majority of consumers will purchase at most one supercomputer per period.

[10] Note that this state variable is bounded between 0 and 1. $\kappa_{f(j)t} = 0$ corresponds to the extreme case of inactivity in the industry, while $\kappa_{f(j)t} = 1$ indicates that firm f has produced the most powerful supercomputer up to time t.

The richness of the TOP500 data allows for additional characteristics to be included in the mean utility (e.g., number of processors, operative system and interconnect type). There are, however, two reasons why a more parsimonious model is preferred. First, most of these variables are strongly correlated with the observed quality measure Rmax, which raise colinearity problems in estimation. Second, it is reasonable to assume that supercomputer buyers are primarily concerned with performance metrics and not so much with the machine specifics (e.g., number of processors).

Consumers can choose an "outside good", whose utility is $U_{r0t} = \xi_{0t} + \epsilon_{r0t}$. This outside alternative may represent not only other types of computers that buyers may consider powerful enough to meet their computing needs (e.g., mainframes, workstations), but also the flow utility from using a high-end computer one already owns. Therefore, it is reasonable to assume that the mean utility of the outside alternative, ξ_{0t}, evolves over time.

The following assumption is imposed for not only for computational convenience, but also to ensure a constant absolute markup per firm in equilibrium (see Nevo and Rossi (2007) for a proof):

Assumption A3: $\forall j = 0, 1, .., J_t$, ϵ_{rjt} is identically and independently distributed extreme-value type I.

Under assumption A3, the market share for product j at time t conditional on consumer type vector is given by

$$q_{jt}(p|D_{rt}) = \frac{\exp(\gamma x_j + \lambda MPP_j + \tau_1 \kappa_{f(j)t} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_l \exp(\gamma x_l + \lambda MPP_l + \tau_1 \kappa_{f(l)t} + \psi N_{f(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{f(l)} D_{rtf(j)})}$$

and the conditional outside good share is

$$q_{0t}(p|D_{rt}) = \frac{1}{1 + \sum_l \exp(\gamma x_l + \lambda MPP_l + \tau_1 \kappa_{f(l)t} + \psi N_{f(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{f(l)} D_{rtf(j)})}$$

I assume that consumers may belong only to a firm network. That is, if for a given f we have $D_{rtf} = 1$, then $D_{rti} = 0 \; \forall i \neq f$. By applying the results on demand derivation from Chapter 1, the probability of purchase of product conditional on consumer type is given by

$$\begin{aligned} q_{jt}(\mathbf{Mkp}_t, \mathbf{s}_t | D_{rt}) \quad = \quad & \frac{\exp(-\alpha Mkp_{f(j)t} + \tau_1 \kappa_{f(j)t} + \psi N_{f(j)t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt}) \exp(\tau_1 \kappa_{dt}) \exp(\nu_{dt}) \exp(\psi N_{dt}) \exp(\Lambda_d D_{rtd})} \\ & \times \exp(\gamma x_j + \lambda MPP_j - \alpha mc_{jt} + \xi_{jt} - \xi_{0t}) \end{aligned}$$

where $\nu_{ft} = \ln \left(\sum_{l \in \mathcal{F}_{it}} \exp(\gamma x_l + \lambda MPP_l - \alpha mc_{lt} + \xi_{lt} - \xi_{0t}) \right)$ is the markup-adjusted inclusive value metric introduced by Nevo and Rossi (2007).

Supply. Each firm f in the industry produces some subset $\mathcal{F}_{ft}$ of the $j = 1, ..., J_t$ products commercialized at time t. Firm f 's flow profits at time t are defined by

$$\Pi_f(\mathbf{p}_t, \mathbf{s}_t) = M_t \sum_{j \in \mathcal{F}_{ft}} (p_{jt} - mc_{jt}) q_{jt}(p, \mathbf{s}_t) - C_f \mathbf{1}\{\exp(\nu_{ft}) > 0\}$$

where $C_f 1\{\exp(\nu_{ft}) > 0\}$ corresponds to fixed production costs conditional on the firm being active in the market (i.e., it has at least one product being commercialized)[11]. The marginal cost of producing product j at time t, mc_{jt}, is assumed to be time-varying and dependent on product characteristics. M_t corresponds to the market size state at time t, which is defined as being the number of consumers potentially interested on purchasing a supercomputer. The last component of flow market profits is $q_{jt}(p, \mathbf{s}_t)$, which denotes the probability of acquiring supercomputer j at time t, given prices and observable states.

Firms are assumed to compete in prices. However, the pricing decisions will affect not only the profits at time t, but also future payoffs. Intuitively, firms pricing decisions will also take into account the fact that prices affect purchase probabilities, which in turn affect state transitions. For example, a firm may engage into "market tipping" strategies to attract consumers to its own network via low prices. Hence, static price competition models will fail to correctly describe firm pricing behavior. I use the results from the both previous section and Chapter 1 to both simplify the profit function and account for these issues.

[11]Note that when firm f is a potential entrant, we have $\mathcal{F}_{ft} = \varnothing$. Hence $\nu_{ft} = -\infty$ and so $\exp(\nu_{ft}) = 0$.

Under the assumptions imposed so far, and assuming the existence of a dynamic equilibrium in markups, the equilibrium profit function is given by

$$\Pi_f(\mathbf{Mkp}_t, \mathbf{s}_t) = Mkp_{ft} \sum_{k=0}^{F} N_{kt} \frac{\exp(-\alpha Mkp_{ft} + \tau_1 \kappa_{ft} + \nu_{ft} + \psi N_{ft} + \Lambda_f \mathbf{1}\{k = f\})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt} + \tau_1 \kappa_{dt} + \nu_{dt} + \psi N_{dt} + \Lambda_d \mathbf{1}\{k = d\})} - C_f \mathbf{1}\{\exp(\nu_{ft}) > 0\}$$

Note that this profit function is well defined even if only $F_t^A < F$ are active in the market. By definition $\exp(\nu_{ft}) = 0$ for firms with no product portfolio, which implies profits equal to zero.

2.3.2. Frontier benefits outside profits

Supercomputer firms are commonly reported to incorporate their advances in computing power into other lines of product (e.g. mainframes, PCs).Consequently, product market profits alone may fail to account for all the benefits the firm derives from its technological frontier. This section describes the modeling of these additional benefits and its empirical identification. One would expect these benefits outside profits to be increasing on the firm's technological frontier. However, if there are several rivals with a similar technological state, the firm only gets a share of that benefit. For example, if two firms are competing on both the supercomputer and the PC market, the benefit of creating improved PCs by incorporating advances in supercomputing is lower if the rival is able to do the same. Therefore I assume that the benefit is decreasing in the number of "technologically similar" rivals. This motivates the following functional form for frontier benefits

$$\Upsilon_f(\mathbf{s}_t) = \frac{\Theta_f(\mathbf{s}_t;\chi)}{2\pi\sqrt{1-\rho^2}\sum_{i=1}^{F}\phi\left(\frac{v_{it}-v_{ft}}{h_A},\frac{\kappa_{it}-\kappa_{ft}}{h_I};\rho\right)} \qquad \forall f = 1,..,F$$

where $\Theta_f(\mathbf{s}_t;\chi)$ is a function with parameters χ, and $\phi(.,.;\rho)$ is the standard bivariate normal distribution with correlation parameter ρ. Although seemingly arbitrary, this specification controls for firm heterogeneity and externalities on outside-profit frontier benefits. Its intuition can be described as follows. Firms are assumed to derive an out-of-profit benefit of $\Theta_f(\mathbf{s}_t;\chi)$ given their current technological state. Since both v_{ft} and κ_{ft} are continuous variables, I cannot restrict attention to the firms whose technological frontier is exactly the same as of firm f. Hence, the bandwidth parameters h_I and h_A are required to define how "similar" the rivals are under some distance metric. I assume that the latter consists on a weighted average of the distance of firm f's from its rivals it terms of both markup-adjusted inclusive values and frontier state. Hence, the bandwidth parameters h_I and h_A define "how close" is a rival by scaling its quality and frontier distance with respect to firm f. I choose a Gaussian kernel as the weighting function[12], although other kernels could be considered (see Härdle (1990) and Pagan and Ullah (1999) for detailed discussions on these methods). The correlation coefficient ρ allows me to control for positive relationship between firm quality and frontier rank state.

[12]The normal density is scaled by $2\pi\sqrt{1-\rho^2}$ to ensure that each firm with exactly the same state values as firm f would count as one similar rival. This guarantees that, in the case of monopoly, the payoff $\Theta_f(\mathbf{s}_t;\chi)$ is divided by one.

One important question from an empirical viewpoint is how to identify the parameters χ. The answer to this question lies on the variation in the number of a firm's similar rivals. This variation which allows me to separate the benefit from innovating from cost functions. According to the "selection effect" introduced in the literature by Aghion et al. (2001, 2003), firm's incentives to innovate are increasing in the number of technologically similar rivals. As this number does not affect the costs of innovation, the benefit function can be identified from the data. Hence, one can estimate $\Theta_f(\mathbf{s}_t;\chi)$ by either (i) using a nonparametric estimator, such as a sieves approximation with Chebyshev polynomials, or (ii) assuming a parametric form for $\Theta_f(\mathbf{s}_t;\chi)$. I follow approach (ii) by assuming that $\Theta_f(\mathbf{s}_t;\chi) = \chi_1\kappa_{ft} + \chi_2\kappa_{ft}^2$.

2.3.3. Frontier extension costs

I assume that, for all firms $f = 1, ..., F_t^A$, the functional form for frontier extension costs is

$$
\begin{aligned}
C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}) \quad = \quad & 1\{I_{ft} > 0\} \Big(c_0 + c_1\kappa_{ft} + (c_2 + \varepsilon_{1ft})I_{ft} + c_3 I_{ft}^2 \\
& + c_4 \sum_{i=1}^{F_t} \mathbf{1}\{\kappa_{it} \leq \kappa_{ft}\}\, 2\pi\sqrt{1-\rho^2}\phi\left(\frac{v_{it}-v_{ft}}{h_A}, \frac{\kappa_{it}-\kappa_{ft}}{h_I}; \rho\right) \\
& + c_5 \sum_{i=1}^{F_t} \mathbf{1}\{\kappa_{it} > \kappa_{ft}\}\, 2\pi\sqrt{1-\rho^2}\phi\left(\frac{v_{it}-v_{ft}}{h_A}, \frac{\kappa_{it}-\kappa_{ft}}{h_I}; \rho\right)\Big)
\end{aligned}
$$

The intuition behind this specification can be described as follows. Investment costs only take place if the firm extends its technological frontier. If the firm does expand its frontier, it pays fixed costs which depend on (i) a fixed amount c_0, (ii) the firm's technological state κ_{ft}, and (iii) the quality and technology states of rivals. The proposed specification aims to control firm heterogeneity and externalities on frontier expansion. The parameters c_4 and c_5 control for the amount of learning that a firm derives from technological laggards and leaders, respectively. The cost function is assumed to be quadratic in I_{ft}, in order to capture convexities in frontier investment expenditures. The parameters c_2 and c_3 therefore provide information on marginal costs of innovation. Marginal costs are assumed to depend also on a private shock ε_{1ft}, but not dependent on firm heterogeneity. The average investment rate conditional on positive investment is approximately 300%, which suggests low marginal innovation costs.

2.3.4. Quality investment costs

Following the framework presented in Chapter 1, the rate of quality investment is defined as

$$
\begin{aligned}
A_{ft} \quad = \quad & \frac{\left(\sum_{l \in \mathcal{F}_{ft}^{IN}} \exp(\gamma x_l + \lambda MPP_l - \alpha mc_{lt} + \xi_{lt} - \xi_{0t})\right)}{\varphi_{ft}} \\
& \frac{- \left(\sum_{k \in \mathcal{F}_{ft}^{OUT}} \exp(\gamma x_k + \lambda MPP_k - \alpha mc_{kt} + \xi_{kt} - \xi_{0t})\right)}{\varphi_{ft}}
\end{aligned}
$$

where $\varphi_{ft} \equiv \exp(\nu_{ft})$. $\mathcal{F}_{ft}^{IN}$ and $\mathcal{F}_{ft}^{OUT}$ denote the sets of products that firm f decides to launch and scrap, respectively. I assume that, for all firms $f = 1, ..., F_t^A$, the functional form for quality investment costs is given by:

$$C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft}; \theta)) = \mathbf{1}\{A_{ft} > 0\}(c_0^A A_{ft} + c_1^A A_{ft}^2) + \mathbf{1}\{A_{ft} < 0\}(c_2^A A_{ft} + c_3^A A_{ft}^2) + \varepsilon_{2ft} A_{ft}$$

This specification can be described as follows. If the firm decreases its markup-adjusted inclusive value, it pays (or receives) a linear payoff $c_2^A A_{ft} + c_3^A A_{ft}^2$. In the limit case of scrapping all products ($A_{ft} = -1$), it receives a payoff of $-c_2^A + c_3^A$. In case $A_{ft} = 0$, the firm does not incur in any expenses. If it augments its existing quality, it pays an amount which is assumed quadratic in A_{ft}. Marginal costs are assumed to depend also on a private shock ε_{2ft}, which is assumed to be the only source firm heterogeneity in quality investment costs.

2.3.5. State transitions

The specification of the continuation value in the firms' Bellman equations requires assumptions on state transition functions. Like in the framework developed in Chapter 1, I impose the following auxiliary assumptions:

Assumption A4: Private shocks on (i) quality investment costs and (ii) frontier extension costs are independently and identically distributed over time and players.

Assumption A5: For every firm f, the private information state ε_{ft} is assumed independent of observed states $\mathbf{s}_t$.

These assumptions are motivated by computational and tractability concerns. Allowing for serial correlation implies a significant (and unaffordable) increase in computational burden.

Markup-adjusted inclusive values. Like in Chapter 1's framework, the dynamics of each active firm's average quality measure depend on three components: (i) the firm's quality investment, (ii) the evolution of the average outside alternative's value ξ_{0t}, and (iii) temporal changes on the characteristics inherent to the products the firm still commercializes (e.g. changes in the marginal costs). For the case of supercomputers, observed characteristics (i.e., reported Rmax and machine architecture) of the product remain unchanged upon introduction in the market. Hence, I assume that (i) a product's observed characteristics do not change after its, introduction, (ii) $\xi_{k,t} = \xi_k + \Delta\xi_{k,t}$, and (iii) marginal costs are a time-varying stochastic function given by

$$mc_{k,t} = mc_0 + mc_t(x_k, MPP_k) + \eta_{kt}^{MC}$$

Under these assumptions we have

$$u_{k,t+1} \equiv y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) + \zeta_{k,t+1} + (\Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta_{k,t+1}^{MC} - \eta_{k,t}^{MC}))$$

where $u_{k,t+1}$ is the total variation in product k's markup-adjusted mean utility (as defined in Chapter 1), $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$ is the projection of

$$-\alpha(mc_{t+1}(x_k, MPP_k) - mc_t(x_k, MPP_k))$$

onto the space of observed states and actions. Intuitively, $\zeta_{k,t+1}$ corresponds to the parcel of marginal cost variation that is not explained by firms' strategic interactions and observed states. The function $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$ is assumed to measurable with respect to the Lebesgue measure.

In order to have a well-defined transition for markup-adjusted inclusive values, auxiliary assumptions are necessary. Following the framework developed in Chapter 1, I pose the following assumptions:

Assumption A6: The process $\zeta_{k,t+1} + \Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta^{MC}_{k,t+1} - \eta^{MC}_{k,t})$ is common to all products within the firm's portfolio. That is, for every firm f, we have

$$\varrho_{ft+1} \equiv \zeta_{k,t+1} + \Delta\xi_{k,t+1} - \Delta\xi_{k,t} - \alpha(\eta^{MC}_{k,t+1} - \eta^{MC}_{k,t}).$$

Assumption A6 implies that the temporal evolution of a firm's product attributes can be encoded into a firm-specific shock. The latter may include shocks common to all firms which might not be observed by the econometrician (e.g., input price fluctuations). This assumption means that the temporal evolution of each firm's product attributes is a result of firms' quality investments decision, markups, industry state and a random shock.

Assumption A7: There exists a constant d_0 for which the process $\xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1} - d_0$ is a martingale. That is, $E[\eta_{ft}|\mathbf{s}_t, \mathbf{A}_t] = 0$ for every firm f, where

$$\eta_{ft} \equiv \xi_{0,t+1} - \xi_{0,t} - \varrho_{ft+1} - d_0$$

Assumption A7 is sufficient for stationary markup-adjusted inclusive values, conditional on the functional form for $y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t)$. For this purpose, I assume that[13]

$$y_f(\mathbf{s}_t, \mathbf{Mkp}_t, \mathbf{A}_t) = -d_1 \ln(\varphi_{ft}) + d_2 \ln(1 + A_{ft})$$

Under these assumptions, the transition of an active firm f's markup-adjusted inclusive value can be written as

$$\upsilon_{ft+1} = d_0 + (1 - d_1)\upsilon_{ft} + (1 + d_2)\ln(1 + A_{ft}) + \eta_{ft}$$

We are left to define the transition for potential players. By definition, these players occupy slots in the state space where $\exp(\nu_{ft}) = 0$. I assume that these inactive firms enter the market (i.e., will have $\exp(\nu_{ft+1}) > 0$) exogenously. Again using the notation $\varphi_{ft+1} = \exp(\nu_{ft+1})$, the transition of inclusive values for inactive players is given by

$$\varphi_{ft+1} = \max\{0, w_1(\mathbf{s}_t) + \varepsilon_{3ft}\}$$

[13] Of course, other alternative functional forms could be considered to ensure stationary markup-adjusted inclusive values. The proposed conditions are sufficient for this purpose, but not necessary.

where $w(\mathbf{s}_t)$ is a function of all observed states and ε_{3ft} is a zero-mean shock. The inclusive-value transition function for all players in the game can be concisely written as

$$\begin{aligned}\exp(\nu_{ft+1}) &= \mathbf{1}\{\exp(\nu_{ft}) > 0\}\exp(d_0 + (1-d_1)\upsilon_{ft} + (1+d_2)\ln(1+A_{ft}) + \eta_{ft}) \\ &\quad + (1-\mathbf{1}\{\exp(\nu_{ft}) > 0\})\max\{0, w_1(\mathbf{s}_t) + \varepsilon_{3ft}\}\end{aligned}$$

Networks. Chapter 1 presented a framework valid when all networks - including the number of consumers who never purchased the product, N_{0t} - is observed. However, the series N_{0t} is not observed in the supercomputer market. In this case, auxiliary assumptions are necessary to recover this series. There are two suitable methodologies to recover N_{0t}. The first is to follow the approach of Benkard (2004), in which consumers are assumed to optimally reallocate their supercomputer stocks every period. Under this assumption, the number of consumers potentially interested on purchasing a supercomputer is well approximated by the total of units sold within the last T^* periods, where T^* corresponds to the expected lifetime of the product. The second approach is to use the specific properties of network transitions by applying Kalman's filter. This latter approach instead requires assumptions on the initial values of both the series and the new consumer arrival process. I follow Benkard's approach.

The transition function for the network vector $\mathbf{N}_t = [N_{0t}, N_{1t}, ..., N_F]'$ is given by

$$\mathbf{N}_{t+1} = Q(\mathbf{Mkp}_t, \mathbf{s}_t)\mathbf{N}_t + \tilde{\mathbf{N}}_{0t+1}$$

where $\tilde{\mathbf{N}}_{0t+1} = [\Delta N_{0t+1}, 0, ..., 0]^T$ corresponds to the column vector with the number of new consumers arriving to the market in the first cell and zero otherwise. This construction implicitly assumes that there are no stochastic terms in firm networks. This implies that there are no consumers leaving a network for exogenous reasons. $Q(\mathbf{Mkp}_t, \mathbf{s}_t)$ is a matrix whose entries $Q_{i,j}$ are defined by the probability of a consumer joining firm i's network when he currently belongs to firm j's network. These probabilities can be derived from the purchase probabilities from a given firm, which were derived in the section 3.1.2. However, one must account for the fact that a consumer who belongs to a firm network still belongs to that network if he chooses the outside alternative. Under these assumptions, every entry in Q is defined as follows

$$Q_{i,0} = P_{i,0}, \forall i = 0, 1, ..., F$$

$$Q_{i,i} = P_{i,i} + P_{0,i}, \forall i = 1, ..., F$$

$$Q_{0,i} = 0, \forall i = 1, ..., F$$

$$Q_{i,j} = P_{i,j}, \forall i \neq j, i, j = 1, ..., F$$

where $P_{i,j}$ corresponds to the probability of buying a product from firm i given that the consumer is part of firm j's network. These correspond to

$$P_{i,j} = \frac{\exp(-\alpha Mkp_{it}) \exp(\tau_1 \kappa_{it}) \exp(\nu_{it}) \exp(\psi N_{it}) \exp(\Lambda_i \mathbf{1}\{i = j\})}{1 + \sum_{d=1}^{F} \exp(-\alpha Mkp_{dt}) \exp(\tau_1 \kappa_{dt}) \exp(\nu_{dt}) \exp(\psi N_{dt}) \exp(\Lambda_d \mathbf{1}\{d = j\})}, \forall i, j$$

By definition, the sum of all networks is equal to the market size. Under Benkard's (2004) assumption, this series is observed. Hence, ΔN_{0t+1} is identified by using the property that

$$\sum_{d=0}^{F} N_{dt+1} = \sum_{d=0}^{F} N_{dt} + \Delta N_{0t+1}$$

I assume that the arrival of new consumers to the supercomputer market at $t+1$ is a linear function of the market size at period t plus a random term. That is,

$$\Delta N_{0t+1} = \phi_0 + (\phi_1 - 1)M_t + \eta_{t+1}^{N_0}$$

where $M_t \equiv \sum_{d=0}^{F} N_{dt}$. Hence, the total market size transition is given by

$$M_{t+1} = \phi_0 + \phi_1 M_t + \eta_{t+1}^{N_0}$$

I assume that $0 \leq \phi_1 < 1$ in order to ensure a stationary process for M_t. Intuitively, that means that the number of consumers arriving to the market should not be growing without bound. Even though the market size series computed under Benkard's assumption suggest this pattern, assessing whether $0 \leq \phi_1 < 1$ holds is an empirical issue. Therefore, the robustness of this assumption is discussed in the empirical section of this Chapter.

Under the proposed law of motion for the net increase in the number of consumers in this market, the transition function of networks is given by

$$\mathbf{N}_{t+1} = Q(\mathbf{Mkp}_t, \mathbf{s}_t)\mathbf{N}_t + \left[\left(\phi_0 + (\phi_1 - 1)\mathbf{1}^T\mathbf{N}_t + \eta_{t+1}^{N_0}\right), 0, ..., 0\right]^T$$

where $\mathbf{1}$ is a $1 \times F$ row vector of ones.

Firm frontier ranks. At this point, I am only let to specify the transition of the technological frontier state. For ease of exposition, I start by focusing on the case where there are no potential entrants (i.e., all firms have $\kappa_{ft} > 0$). The following proposition establishes the state transition when all firms are incumbents.

Proposition 1. *If for every firm f, $\kappa_{ft} > 0$, then the transition of κ_{ft+1} conditional on player's actions and observed states is deterministic and given by*

$$\kappa_{ft+1} \equiv \frac{(1 + I_{ft})\kappa_{ft}\mathbf{1}\{A_{ft} > -1\}}{\max\limits_{i=1,...,F} \{(1 + I_{it})\kappa_{it}\mathbf{1}\{A_{it} > -1\}\}}$$

Proof. Let h_{ft} be firm f's frontier at time t. Next period's frontier, h_{ft+1}, is given by $h_{ft}(1 + I_{ft})$ if the firm doesn't scrap all products (i.e., if $A_{ft} > -1$) and zero otherwise. The industry's frontier in the next period if simply the maximal h_{it+1} after investment decisions. By definition, we have

$$\kappa_{ft+1} \equiv \frac{h_{ft+1}\mathbf{1}\{A_{ft} > -1\}}{\max\limits_{i=1,...,F} \{(h_{it+1}\mathbf{1}\{A_{it} > -1\}\}} = \frac{(1 + I_{ft})h_{ft}\mathbf{1}\{A_{ft} > -1\}}{\max\limits_{i=1,...,F} \{(1 + I_{it})h_{it}\mathbf{1}\{A_{it} > -1\}\}}$$

The transition of κ_{ft+1} is the obtained by dividing both the numerator and the denominator by $\max\limits_{i=1,\ldots,F}\{h_{it}\}$. $\square$

However, if some of the F players are not incumbents, one must extend the above transition to allow for cases where $\kappa_{ft}=0$. Let h^{E}_{ft+1} be the maximal computing speed that a potential entrant f produces, should he enter at period $t+1$.

Defining $W_{f,t+1}=\ln(h^{E}_{ft+1})-\ln\left(\max\limits_{i=1,\ldots,F}\{h_{it}\}\right)$, I assume the following:

Assumption A8: $W_{f,t+1}$ is a stationary random variable with mean $w_2(\mathbf{s}_t)$, where $w_2(\mathbf{s}_t)$ is a measurable function with respect to the Lebesgue measure.

In what follows, it is convenient to write $W_{f,t+1}$ as the sum of its mean with a zero-mean shock, i.e., $W_{f,t+1}=w_2(\mathbf{s}_t)+\varepsilon_{4ft}$. Since I am assuming an exogenous entry process, this assumption poses structure on the technological frontier rank of new active firms. In particular, the technological frontier of an entrant at $t+1$ is given by

$$\begin{aligned}\kappa_{ft+1} &= \frac{h^{E}_{ft+1}\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}}{\max\limits_{i=1,\ldots,F}\left\{(h_{it+1}\mathbf{1}\{A_{it}>-1\},h^{E}_{it+1}\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}\right\}}\\ &= \frac{\exp(W_{f,t+1})\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}}{\max\limits_{i=1,\ldots,F}\{(1+I_{it})\kappa_{it}\mathbf{1}\{A_{it}>-1\},\exp(W_{i,t+1})\mathbf{1}\{w(\mathbf{s}_t)+\varepsilon_{3ft}>0\}\}}\end{aligned}$$

Hence, the transition of technological states for all players in the game can be concisely written as

$$\kappa_{ft+1} = \frac{(1+I_{ft})\kappa_{ft}\mathbf{1}\{A_{ft}>-1\}\mathbf{1}\{\kappa_{it}>0\}}{\max\limits_{i=1,\dots,F}\{(1+I_{it})\kappa_{it}\mathbf{1}\{A_{it}>-1\},\exp(w_2(\mathbf{s}_t)+\varepsilon_{4ft})\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}\}} + \frac{\exp(w_2(\mathbf{s}_t)+\varepsilon_{4ft})\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}\mathbf{1}\{\kappa_{it}=0\}}{\max\limits_{i=1,\dots,F}\{(1+I_{it})\kappa_{it}\mathbf{1}\{A_{it}>-1\},\exp(w_2(\mathbf{s}_t)+\varepsilon_{4ft})\mathbf{1}\{w_1(\mathbf{s}_t)+\varepsilon_{3ft}>0\}\}}$$

2.3.6. Equilibrium concept

This section concerns the definition of equilibrium for the dynamic model just outlined. Using the assumptions imposed in the section 3, I restrict attention to a pure-strategy equilibrium where firms choose Markovian strategies of the form $\sigma_f(\mathbf{s},\varepsilon_f) \equiv \{Mkp_f(\mathbf{s},\varepsilon_f), I_f(\mathbf{s},\varepsilon_f), A_f(\mathbf{s},\varepsilon_f)\}$. and Let $V_f(\mathbf{s},,\varepsilon_f|\sigma_f,\sigma_{-f})$ be firm f's expected discounted payoffs when he chooses strategy σ_f and his rivals choose σ_{-f}. A Markov-Perfect Nash Equilibrium (MPNE) of this dynamic oligopoly game can be defined as follows.

Definition 2. ***(MPNE):*** *A Markovian strategy profile $\sigma^* \equiv (\sigma_1^*, ..., \sigma_F^*)$ is an MPNE if, for every firm f, σ_f^* solves f's problem given σ_{-f}^*. That is,*

$$V_f(\mathbf{s},\varepsilon_f|\sigma_f^*,\sigma_{-f}^*) \geq V_f(\mathbf{s},\varepsilon_f|\hat{\sigma}_f,\sigma_{-f}^*)$$

for all $\mathbf{s}$ and alternative strategy $\hat{\sigma}_f$.

Existence of a pure-strategy MPNE can be done by either (i) invoking theorems 1 and 2 of Jenkins, Liu, Matzkin and McFadden (2004) or (ii) applying Brower's fixed point theorem after writing the system for first-order conditions of all firms on markups, quality investment and frontier extension rates in fixed point form. The fact each firm's value function is unique ensures the existence of at least one equilibrium[14]. Uniqueness can in principle be verified either by directly proving that the fixed point form of the system is a contraction, or by checking if the system satisfies Blackwell's sufficient conditions. Unfortunately, MPNE uniqueness verification using these and other proof strategies are particularly difficult and still an open area of research (see Doraszelski and Satterthwaite (2007) for a discussion). However, uniqueness is not necessary for estimation purposes, whose details are addressed in the next section.

2.4. Estimation methods

The structural estimation of the model consists on a variant of the two-step method proposed by Bajari, Benkard and Levin (2006). Several other dynamic model estimation methods have been introduced in the literature (e.g., Aguirregabiria and Mira (2006), Berry, Ostrovsky and Pakes (2005), Hotz, Miller, Sanders and Smith (1994)) after Rust's (1987) pioneer work on dynamic discrete choice models[15]. Even though the method of Bajari et al. (2006) has been considered in other dynamic oligopoly studies (e.g., Beresteanu and Ellickson (2006), Ryan (2006)), this Chapter will follow a modified framework. I start by providing a general overview of my estimation strategy, explaining how the methods outlined in Chapter 1 can be applied to my model for the supercomputer industry. Then I discuss the technical details of each estimation step.

[14]If each firm's value function were not unique, existence of equilibrium could instead be done by invoking Kakutani's fixed point theorem.

[15]A comprehensive survey of these methods can be found in Ackerberg et al. (2006).

2.4.1. An overview of the estimation

Following Bajari, Benkard and Levin (2006), the first step of my estimation strategy consists on estimating the flow profit parameters, as well as the observable state transitions and policies. After choosing appropriate instruments, demand parameters are recovered by using Nevo's (2001) demand estimator. Markup and marginal cost series are obtained using the assumption of constant absolute markups per firm. I regress prices on a proposed functional form for marginal costs and firm-period specific dummies. Estimated marginal costs and demand parameters are used to compute markup-adjusted inclusive values, quality investments and firm network transitions. Optimal policy functions are nonparametrically estimated using Chebyshev polynomials of observed states, following the lines of Jenkins, Liu, Matzkin and McFadden (2004).

In the second stage of estimation, I deviate from Bajari, Benkard and Levin (2006). I solve for the continuation values of firms using function approximation methods (see Miranda and Fackler (2002)) and the policy and transition estimates from the first stage. Instead of using the continuation values in the inequality sampling method proposed by Bajari et al. (2007), I include them in a GMM estimation procedure where I use firm optimality conditions directly. I form moments using (i) first-order conditions on markups, frontier extension rates and quality investment, and (ii) the condition that the firm will only extend its frontier if it yields a higher value than not extending. To integrate these conditions with respect to rival private information, I replace their policies recovered in the first round of estimation in the firm moment conditions. Integration is performed using Gauss-Legendre quadrature with respect to policy function residual distributions. This approach has several advantages. First, forward simulation of firm continuation values given first step estimates is not necessary. That procedure is at the core of Bajari, Benkard and Levin's (2007) method, being also the main source of its computational burden. Second, unlike Bajari, Benkard and Levin, a flow payoff function linear in parameters is not necessary to ensure an affordable computational burden. Since the structural parameters enter both in the flow payoff function and in the continuation value approximation, the estimation problem effectively becomes a parametric one. This implies an estimation burden similar to minimum distance estimators. Finally, the estimates of continuation values may be useful for simulation purposes. These can be used for equilibrium computations in counterfactual policy experiments where structural parameters are not modified. That is, the intermediate step of MPNE computation where the value function is computed can be avoided by treating the estimated value function as the true function. Computational time is reduced at the cost of introducing estimation error in the simulation routine.

2.4.2. First-step estimates

Profit function parameters. There are two alternative estimation strategies to recover the profit function parameters. The first consists on the method introduced by Berry, Levinsohn and Pakes (1995). In this method, a framework which enables one to estimate demand and supply parameters for oligopolistic differentiated products markets is proposed. Like other related methods (e.g., Berry (1994)), it relies on the ability of observed product characteristics to explain consumer utility. The second alternative method is Nevo's (2001) estimator, which deals with cases where that ability is questionable by adding brand fixed effects. Since the observable supercomputer characteristics available in my dataset may fail to capture important unobserved attributes (e.g., whether the machine performs particularly well on solving certain types of problems), I follow Nevo's method to estimate demand parameters.

The model described in section 3 assumes that the (constant) absolute markup charged by every firm is observable. However, only product prices are available, not the charged markups. So the question is how to empirically separate markups from marginal costs given price information only. The answer to this identification issue relies on the assumption that, within each given period, all products that a firm commercializes are charged the same constant absolute markup. Under this assumption, at any given time t, all variation in prices across the firm's product portfolio must come from differences in products' marginal costs. Therefore, marginal costs and firm markups are identified.

I start by describing demand estimation. A product's market share conditional on the the consumer's network type is given by

$$q_{jt}(p|D_{rt}) = \frac{\exp(\gamma x_j + \lambda MPP_j + \tau_1 \kappa_{f(j)t} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_{jt} - \xi_{0t} + \Lambda_{f(j)} D_{rtf(j)})}{1 + \sum_l \exp(\gamma x_l + \lambda MPP_l + \tau_1 \kappa_{f(l)t} + \psi N_{f(l)t} - \alpha p_{lt} + \xi_{lt} - \xi_{0t} + \Lambda_{f(l)} D_{rtf(j)})}$$

In line with the structure imposed for state transitions, I assume that $\xi_{jt} = \xi_j + \Delta\xi_{jt}$[16]. In addition, I decompose consumer r's utility on consuming product j at time t as

$$U_{rjt} = \delta_{jt} + \Lambda_{f(j)} D_{rtf(j)} + \epsilon_{rjt}$$

[16]One could also assume a decomposition where time effects are present in ξ_{jt}. However, these are not separable from ξ_{0t}, and therefore not identified.

where

$$\delta_{jt} = \gamma x_j + \lambda MPP_j + \tau_1 \kappa_{f(j)t} + \tau_2 \mathbf{1}\left\{\kappa_{f(j)t} = 1\right\} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_j + \Delta\xi_{jt}$$

is the mean valuation of product j at time t. Apparently, one should could consistently estimate the demand parameters by either minimizing the distance between observed and predicted market shares. However, if there exists correlation of the error term $\Delta\xi_{jt}$ with other variables in δ_{jt}, this approach will yield inconsistent estimates. Even if suitable instruments are available, the fact that $\Delta\xi_{jt}$ enters the product market share function nonlinearly precludes this method. I instead follow the methodologies of Berry (1994), Berry, Levinsohn and Pakes (1995) and Nevo (2001), where this error term can be obtained by using the definition of δ_{jt} upon inverting the product market share functions. The model's prediction for the later is given by

$$q_{jt}(p) = \int q_{jt}(p|D_{rt})dP(D_{rt1}, ..., D_{rtF}|\mathbf{s}_t)$$

where $P(D_{rt1}, ..., D_{rtF}|\mathbf{s}_t)$ corresponds to the distribution of consumer network types induced by the observed network states. In our model, this consists on a multinomial distribution, where the probability of drawing $D_{rtf} = 1$ is given by $\dfrac{N_{ft}}{M_t}$, $\forall f = 1, ..., F$[17]. Denoting $s_{j,t}$ as the observed market share for product j at time t, I define $\delta_{jt}(S_{.,t}, \Lambda)$ as the

[17]Since I observe all the market purchases and network sizes, I could instead write the predicted market share by replacing the integral by a sum across all consumers. For computational reasons, however, it is preferable to numerically integrate using a few draws from the true distribution of consumer states.

(unique) solution to the implicit system of equations $q_{.t}(\delta_{.,t};\Lambda) = s_{.,t}$[18]. The computation of $q_{.t}(\delta_{.,t};\Lambda)$ requires integration with respect to each D_{rtf}. This can be done by replacing the integral by a sum across random draws from the joint distribution $P(D_{rt1},...,D_{rtF_t}|\mathbf{s}_t)$. Given the solutions $\delta_{jt}(S_{.,t},\Lambda)$, the error term is given by

$$\Delta\xi_{jt} = \delta_{jt}(S_{.,t},\Lambda) - (\gamma x_j + \lambda MPP_j + \tau_1\kappa_{f(j)t} + \psi N_{f(j)t} - \alpha p_{jt} + \xi_j)$$

The term $\Delta\xi_{jt}$ can be treated as econometric error term. Since manufacturers take into account all product characteristics in their pricing decisions, This term should be correlated with prices. For the case of supercomputers there is also reason to believe that some of the observed product attributes are correlated with the error term. For example, supercomputer's unobserved characteristics may be contributing to the system's speed on solving a system of 1000 equations with 1000 unknowns, which is an observable machine attribute. Therefore, valid instruments for both prices and some of the product's attributes should are necessary. Determining how relevant these endogeneity issues are for estimation is an empirical issue. Hence, the details about the need and choice of instruments will be left for section 2.5.

An important issue is how to empirically separate the time-invariant components of the mean utility function. Nevo (2001) shows that they can be recovered from the data by first replacing $\gamma x_j + \lambda MPP_j + \xi_j$ by brand-specific dummies d_j in the mean utilities δ_{jt}, and then regress the estimates $\hat{d}_j$ on the observed attributes x_j and MPP_j using Chamberlain's (1982) minimum-distance method. However, this method assumes that

[18]The proof of uniqueness of this solution can be found in Berry, Levinsohn and Pakes (1995)

$E[\xi_j|x_j] = 0$ for every product j, which may fail to hold by the same reasons outlined in the previous paragraph. Again, these endogeneity concerns are an empirical matter. For this reason, if instruments are required for consistent estimation, the estimator for γ will be the two-stage least-squares estimator of

$$\hat{d}_j = \gamma X_j + \xi_j$$

In case no endogeneity concerns arise, estimates of γ can be recovered using a GLS regression. In any case, fixed-effect estimates of unobserved product attributes are given by $\hat{\xi}_j = \hat{d}_j - \hat{\gamma} X_j$.

Letting λ denote the set of all demand parameters to be estimated, and defining $\Delta\xi(\lambda)$ be the vector of errors evaluated at a given value of λ, one can estimate the true value of λ with the GMM estimate

$$\hat{\lambda} = \arg\min_{\lambda} \Delta\xi(\lambda)' Z \Sigma^{-1} Z' \Delta\xi(\lambda)$$

where Z is a set of instruments orthogonal to $\Delta\xi(\lambda)$ and Σ is a consistent estimate of $E[Z'(\Delta\xi)(\Delta\xi)'Z]$.

Markups, marginal costs and quality investments. Markup and marginal cost series are obtained using the assumption of constant absolute markups per firm. I assume that marginal costs of producing product j at time t are given by

$$mc_{j,t} = m_0 + m_1 MPP_j + m_t x_j + \eta_{jt}^{MC}$$

where m_t corresponds to time-varying coefficients on observed LINPACK Benchmark of product j. This specification aims to control for a potential temporal decrease on the marginal costs of producing a given computing speed. The coefficients can be estimated by running the linear regression

$$p_{jt} = m_0 + m_1 MPP_j + m_t x_j + m^{Mkp} \cdot B_{jt} + \eta_{jt}^{Mkp}$$

where $B_{jft} = \mathbf{1}\{j \in \mathcal{F}_{ft}\}$, $B_{jt} = [B_{j1t}, ..., B_{jFt}]^T$ and m^{Mkp} is a $1 \times F$ vector of coefficients. The latter measure the markup charged by firm f at time t.

Estimated marginal costs and demand parameters are used to compute Nevo and Rossi's (2007) markup-adjusted inclusive values for all firms, quality investments and firm network transitions. The estimates are also used to compute the quality investment series $A_{f,t}$.

Policy functions. For the second stage of estimation, policy function estimates are necessary. One way to estimate quality investment and markup policies is to project each $A_{f,t}$, $I_{f,t}$ and $Mkp_{f,t}$ on the space of observed states, $\mathbf{s}_t$. The fact that I only have 16 time periods precludes estimating firm-specific policy functions. Fortunately, the structure of the flow payoff suggests that one can restrict attention to symmetric and anonymous policy functions, as defined by Doraszelski and Satterthwaite (2007). These concepts are defined as follows:

Definition 3. ***(Symmetry):*** *A function f_n is said to be symmetric if, for all $n = 1, ..., N$,*

$$f_n(z_1, ..., z_n, ..., z_N) = f_1(z_n, ..., z_1, ..., z_N)$$

Definition 4. ***(Anonymity):*** *A function f_n is said to be anonymous if*

$$f_1(z_1, z_2..., z_k, ..., z_l, ..., z_N) = f_1(z_1, z_2..., z_l, ..., z_k, ..., z_N)$$

for all $k \geq 2$ and $l \geq 2$.

The fact that the flow payoff function satisfies these two properties motivates an econometric strategy where optimal policy functions are assumed to have the same properties. Hence, one way to obtain consistent estimates of the policy functions is to use sieve functions on firm-specific states and the auxiliary variables $\kappa_{-ft} \equiv \sum\limits_{d \neq f} \kappa_{dt}$, $\exp(\nu_{-ft}) \equiv \sum\limits_{d \neq f} \exp(\nu_{dt})$ and $N_{-ft} \equiv \sum\limits_{d \neq f} N_{dt}$. I estimate the following regression for the markup policy by by OLS:

$$\begin{aligned} Mkp_{f,t} &= \sum_{k_1=0}^{K_1} ... \sum_{k_7=0}^{K_7} \chi^{Mkp}_{k_1,...,k_7} T_{k_1}(N_{0t}) \times T_{k_2}(\kappa_{ft}) \times T_{k_3,}(N_{ft}) \times T_{k_4}(\exp(\nu_{ft})) \times T_{k_5}(\kappa_{-ft}) \\ &\quad \times T_{k_6,}(N_{-ft}) \times T_{k_7}(\exp(\nu_{-ft})) + \varepsilon^{Mkp}_{f,t} \end{aligned}$$

where $T_k(.)$ corresponds to a first-kind Chebyshev polynomial or order k, and $\chi^{Mkp}_{k_1,...,k_7}$ denote the coefficients on the tensor product of these polynomials. The orders of approximation in each dimension, K_i, are an empirical issue and therefore discussed in the results section.

Both quality investment and frontier expansion policies are recovered using nonparametric methods. However, I must control for the fact that the support for these policies is bounded below. In particular, we have $A_{f,t} \geq -1$ and $I_{f,t} \geq 0$. The latter restriction is particularly important due to lumpy frontier extension behavior seen in the data. Even though firms decide optimal investment rates by comparing marginal benefits to marginal costs of innovation at their frontier, they may not invest due to considerable fixed costs of frontier extension. This motivates the estimation of each policy by using the Tobit I model. Like for the markup policy, I consider Chebyshev polynomials on observed states as regressors. The proposed specification for the investment policy is therefore $I_{ft} = \max\{0, w_{f,I}(\mathbf{s}_t, \chi^I) + \varepsilon^I_{f,t}\}$, where $\varepsilon^I_{f,t}$ is a zero-mean normally distributed error and

$$\begin{aligned} w_{f,I}(\mathbf{s}_t, \chi^I) &= \sum_{k_1=0}^{K_1} ... \sum_{k_7=0}^{K_7} \chi^I_{k_1,...,k_7} T_{k_1}(N_{0t}) \times T_{k_2}(\kappa_{ft}) \times T_{k_3,}(N_{ft}) \times_{k_4} (\exp(\nu_{ft})) \\ &\times T_{k_5}(\kappa_{-ft}) \times T_{k_6,}(N_{-ft}) \times T_{k_7}(\exp(\nu_{-ft})) \end{aligned}$$

I propose an analogous specification for $A_{f,t}$. That is, $A_{ft} = \max\{-1, w_{f,A}(\mathbf{s}_t, \chi^A) + \varepsilon^A_{f,t}\}$ where $\varepsilon^A_{f,t}$ is a zero-mean normally distributed error and

$$\begin{aligned} w_{f,A}(\mathbf{s}_t, \chi^A) &= \sum_{k_1=0}^{K_1} ... \sum_{k_7=0}^{K_7} \chi^A_{k_1,...,k_7} T_{k_1}(N_{0t}) \times T_{k_2}(\kappa_{ft}) \times T_{k_3,}(N_{ft}) \times T_{k_4}(\exp(\nu_{ft})) \\ &\times T_{k_5}(\kappa_{-ft}) \times T_{k_6,}(N_{-ft}) \times T_{k_7}(\exp(\nu_{-ft})) \end{aligned}$$

State transitions. The estimation of profit function parameters allows me to compute all network transitions but the outside alternative network, N_{0t}. The parameters on the law of motion of this observed state are recovered by running the regression $M_{t+1} = \phi_0 + \phi_1 M_t + \eta_{t+1}^{N_0}$, where the definition market size M_t follows the approach of Benkard (2004). The distribution of the error $\eta_{t+1}^{N_0}$ is a necessary element in the second stage of estimation. However, it can only be assessed upon estimation of the proposed equation. In the next section I will proceed by assuming that this distribution has been inferred.

The transitions of both markup-adjusted inclusive values and technological states depend on potential entrant-specific shocks, ε_{3ft} and ε_{4ft}. Recall that, for any potential entrant f, we have

$$\exp(\upsilon_{ft+1}) = \max\{0, w_1(\mathbf{s}_t) + \varepsilon_{3ft}\}$$

$$W_{f,t+1} = w_2(\mathbf{s}_t) + \varepsilon_{4ft} if \exp(\upsilon_{ft+1}) > 0$$

where $W_{f,t+1} = \ln(h_{ft+1}^E) - \ln\left(\max_{i=1,\ldots,F}\{h_{it}\}\right)$. Assuming that ε_{3ft} is a zero-mean normal shock, the first equation can be estimated using a Tobit Type I specification. The number of sample observations for this regression is given by the number of empty slots available at every period t. The equation on $W_{f,t+1}$ can be estimated by OLS conditional on the set of inactive players who decided to enter at period t.

At this point, I am left to recover the transitions of markup-adjusted inclusive values for incumbents. The functional form assumed for this transition is

$$\upsilon_{ft+1} = d_0 + (1 - d_1)\upsilon_{ft} + (1 + d_2)\ln(1 + A_{ft}) + \eta_{ft}$$

which can be recovered with panel data linear regression techniques. In this Chapter I will use GLS random-effects regression.

2.4.3. Second-step estimates

The first step of estimation recovers all parameters except the ones of frontier extension costs, benefits outside profits, and quality investment costs. I denote θ as the vector of parameters which were not estimated in the first step. The first-order conditions on markups, quality investment and frontier extension provide information on all parameters except fixed innovation costs. In order to obtain information to estimate fixed frontier extension costs, I must impose the condition that if a positive frontier innovation rate is chosen, then it must be better than no investment. However, these conditions depend on firm continuation values, which are not available to the researcher. To deal with this difficulty I exploit the structure of the Bellman equations to estimate a sieves approximation to value functions. For ease of exposition, the details on value function approximation can be found in Appendix B. I proceed by deriving the moment conditions used in the GMM objective function. Auxiliary derivations for this conditions can also be found in Appendix B.

For ease of notation, let g_k be the k^{th} entry in the observed state vector at period $t+1$, whose dimension is denoted $\Delta(S)$. For the necessary moment conditions, I start with the first-order condition on absolute markup. This condition can be written as

$$\int_{\varepsilon_{-f}} \Bigg(\frac{\partial \Pi_{ft}(\mathbf{Mkp}_{ft}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t))}{\partial Mkp_{ft}}$$

$$+\beta E \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial Mkp_{ft}} | \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t \right] \Bigg) dF(\varepsilon_{-f}) = 0$$

Replacing integration with respect to ε_{-f} by integration of policy estimate's residuals and (ii) the integrated value function by its approximant, we obtain a moment condition by considering the expectation of this equation's left-hand side.

One can also derive a first-order condition for both quality investment and frontier extension. However, I must account for the fact that the support of these variables is bounded below. I use the Karush-Kuhn-Tucker condition (see Miranda and Fackler (2002) pp 191-193) to derive moment conditions. I start with frontier extension investment, which must be nonnegative. Optimality in this type of investment must satisfy the condition

$$\int_{\varepsilon_{-f}} \left(-\frac{\partial C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta)}{\partial I_{ft}} + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial I_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu_{ft}^{I} \mathbf{1}\{I_{ft} = 0\} = 0$$

where μ_{ft}^I is a nonpositive Lagrange Multiplier. Unfortunately, ε_{1ft} can only be isolated when frontier extension investment is strictly positive. I deal with this difficulty by considering the random variable

$$\tilde{\varepsilon}_{1ft} \equiv \mathbf{1}\{I_{ft} > 0\}\varepsilon_{1ft} + \mu_{ft}\mathbf{1}\{I_{ft} = 0\}$$

as the error term in the Karush-Kuhn-Tucker condition above. Letting μ^I be the expected value of $\tilde{\varepsilon}_{1ft}$, this condition can be written as

$$\int\limits_{\varepsilon_{-f}} \left(-\mathbf{1}\{I_{ft} > 0\}(2c_6 I_{ft} + c_5) + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial I_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu^I = \tilde{\varepsilon}_{1ft} - \mu^I$$

Since the right-hand side of the equation has expected value zero, one can consider the expectation of the left-hand side as an extra moment condition. Like in the first-order condition in markups, I replace the integrated value function V_f by its approximant and integrate private information using the best-reply functions estimated in the first-stage.

A moment condition for quality investment is analogously derived. This corresponds to

$$\int\limits_{\varepsilon_{-f}} \left(-\frac{\partial C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft}; \theta))}{\partial A_{ft}} + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial A_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu^A_{ft} \mathbf{1}\{A_{ft} = -1\} = 0$$

I again replace $F(\varepsilon_{-f})$ by the distributions of residuals of first-stage policy estimates. Firm f's integrated value function is also replaced by its approximant to form a moment condition. Like for frontier investment, I demean the moment condition by a constant μ^A, since the error term for this condition has a similar structure.

In principle, one could estimate the fixed cost parameters by matching the probability of no frontier investment with a consistent estimator (e.g. the probability of no frontier extension implied by the first-stage Tobit estimate for this policy). However, this requires the knowledge of the frontier investment rate that the firm is considering when comparing the value of extending its frontier with the value of not investing. That rate is only observed in the data when the firm does extend the frontier. I deal with this difficulty by replacing that rate with $E\left[I_{ft}|I_{ft}>0,\mathbf{s}_t\right]$ whenever the firm is observed not to extend its technological frontier. An estimate of $E\left[I_{ft}|I_{ft}>0,\mathbf{s}_t\right]$ can be computed from the Tobit Type I approximation recovered in the first stage of estimation. Thus, denoting the right-hand side of the positive investment condition (see appendix B for details) when no frontier innovation is observed by $R(E\left[I_{ft}|I_{ft}>0,\mathbf{s}_t\right],\mathbf{s}_t;\theta)$, the moment condition that identifies fixed innovation costs is

$$
\begin{aligned}
&E\left[\mathbf{1}\{I_{ft}=0\}\Pr(\varepsilon_{1ft}>R(E\left[I_{ft}|I_{ft}>0,\mathbf{s}_t\right],\mathbf{s}_t;\theta))\right.\\
&\left.-\mathbf{1}\{I_{ft}=0\}\Phi\left(\frac{-w_{f,I}(\mathbf{s}_t,\hat{\chi}^I)}{\hat{\sigma}_\varepsilon^I}\right)\right]=0
\end{aligned}
$$

where $\Phi(.)$ denotes the standard Normal distribution. The term in $\Phi(.)$ evaluates the probability of no investment implied by the Tobit estimates of the investment policy. In order to empirically implement this condition, a probability distribution for ε_{1ft} is required. However, this distribution is unknown. To deal with this difficulty, I replace the probability parcel in the moment condition by the integral of Gallant and Tauchen's (1989, 1992) density estimator. The latter consists on an approximation via Hermite expansion to the true density, having the normal density as the leading term in that expansion. The conditional density is scaled by an appropriate function so that it integrates to unity. Gallant and Nychka (1987) proved that this technique is rich enough to accommodate densities from a large class that includes densities with fat, t-like tails, densities with tails that are thinner than normal, and skewed densities. Formally, the density estimator is

$$f_K(\varepsilon_{1ft}) \equiv \left(\frac{[P_D(z_{1ft})]^2}{\int [P_D(u)]^2 \phi(u) du} \right) \frac{\phi(z_{1ft})}{\sigma_{\varepsilon_1}}$$

where $z_{1ft} \equiv \frac{\varepsilon_{1ft}}{\sigma_{\varepsilon_1}}$ and $P_D(z_t) \equiv \sum_{i=0}^{D} \omega_i z_t^i, \;\; \omega_0 = 1.$

At this point, I am left to form a GMM estimator for θ. I use the fact that the error terms in the equations outlined above are assumed independent of observed states to form enough moments to identify all the parameters in θ. In particular, I consider Chebyshev polynomials of all observed states as instruments for the GMM weighting matrix. I denote Z as the vector of these instruments. In addition, I denote all auxiliary parameters (e.g. polynomial coefficients in Gallant and Tauchen's density, variance of ε_{1ft}) by ω. Letting $\Psi(\mathbf{s}_t, \theta, \omega, h_f(\mathbf{s}_t, \theta))$ be the vector of moment conditions derived above and $\alpha \equiv (\theta, \omega)$, the GMM estimator is defined as

$$\hat{\alpha} \equiv \arg\min_{\alpha} \Psi(\mathbf{s}_t, \theta, \omega, h_f(\mathbf{s}_t, \theta))' Z \Sigma^{-1} Z' \Psi(\mathbf{s}_t, \theta, \omega, h_f(\mathbf{s}_t, \theta))$$

where Σ is a consistent estimate of $E[Z'\Psi\Psi'Z]$. The computation of standard errors must account for the fact that the second step of estimation takes first-step estimates as given. Corrected standard errors can be computed by either (i) using the correction methods of Murphy and Topel (1985) or (ii) bootstrapping samples of the data with replacement and estimating the dynamic parameters with each subsample. I followed the second procedure for computing standard errors.

2.5. Data and estimation results

2.5.1. The data

The data required for estimation using the methods described in the previous section consists of the following variables: market shares and prices in each year, supercomputer characteristics, network sizes (i.e., N_{ft}), technological rankings of firms (i.e., κ_{ft}) and frontier expansion rates. The data ranges from 1990 to 2005. The details on the data construction and the definition of its variables are presented in Appendix A.

Tables 2.3 and 2.4 provide yearly descriptive statistics of the most relevant variables in the data set. Except for the first three columns (measured in units), all entries on Table 2.3 correspond to sales-weighted averages, where real prices are in $1M units, Rmax is measured in Giga-FLOPS, and processors speed is measured in MHz. Firmrank κ_{nt} corresponds to the ratio between a firm's maximal Rmax production record and the industry's maximum.

Some relevant trends can be detected. First, there is a considerable growth in total sales up to 1997, followed by a stable evolution of the quantity sold. This pattern follows closely the evolution of the average real price statistic, which experiences a sharp decline up to the mid 1990s and then varies between $6M and $10M. This suggests that firms engaged into "market tipping" to attract new consumers in the early 1990s and then "harvested" their consumer base via higher prices. Second, average network size per

Table 2.3. Yearly Sales-Weighted Means

Year	Quantity	Network	Price	Rmax	No. Proc.	Proc. speed	Firm rank (κ_{ft})
1990	73	7	11.549	1.578	215.246	140.058	0.129
1991	101	13.444	7.793	1.733	151.435	184.961	0.230
1992	169	16.667	8.455	2.508	51.443	137.896	0.428
1993	275	25.308	7.363	5.212	189.646	112.010	0.281
1994	400	33.6	4.708	6.249	87.955	108.876	0.390
1995	405	45.933	3.423	6.590	43.049	107.524	0.398
1996	407	54.333	6.245	19.391	61.985	180.838	0.322
1997	584	92.7	9.581	42.021	111.603	246.689	0.177
1998	407	134.667	8.457	51.801	130.081	275.771	0.378
1999	510	133.333	9.471	94.920	192.558	339.015	0.401
2000	477	125.091	8.789	145.313	184.972	389.066	0.478
2001	408	145.9	7.333	242.820	227.902	536.621	0.387
2002	449	124.5	8.968	638.469	347.791	1006.972	0.126
2003	388	110.333	6.865	836.986	341.118	1650.017	0.229
2004	489	120.286	9.552	1889.182	614.918	1951.751	0.473
2005	463	125.067	9.961	4329.1	1382.551	2529.151	0.417

firm grows considerably until 1998, and fluctuates around the value for that year in the following periods. This suggests that until the mid 1990s the growth of network size per firm was primarily due to increases in new consumers, which is in line with the described pattern on quantity sold. The wave-like evolution in average network following that period is likely to be explained by variations both in average prices and number of firms in the market. Finally, there is a exponential growth on the average LINPACK record (i.e. Rmax) of supercomputers, and a similar evolution applies to the average number of processors and processor speed used in the machines. Finally, the wave-like evolution of the average firmrank indicates that firms tend to quickly approach the technological leader until some industry player extends its frontier considerably. The descriptive statistics presented on Table 2.4 not only reinforce this evidence, but also indicate that firms tend to innovate considerably, even though no innovation is particularly frequent.

Table 2.4. Summary Statistics

Variable	Mean	Median	Std	Min	Max
Price	7.952	5.660	9.791	0.238	317.126
Rmax	633.465	51.2	4683.404	0.422	280600
No. processors	298.336	96	2092.405	1	131072
Processor speed	713.410	333.3	893.827	7	3600
Product market share(%)	0.2	0.04	0.4	0.031	5.3
Firm market share(%)	4.1	4.5	3.1	0.031	26.1
Frontier expansion rate (%)	296.4	47.8	857.2	0	3100.7
Firm rank (κ_{ft})	0.424	0.240	0.368	.006	1
Firm network (N_{ft})	81.760	101.517	51.386	0	571

2.5.2. Estimation Results

I estimate demand and supply equations separately. In order to assess the need for instruments in demand estimation, I run a preliminary OLS regression of the logit demand equation. The first column of Table 2.5 presents the results of this regression. Prices are measured in $1M units (constant 1998 dollars), network size is measured in hundreds of consumers, and Rmax is in Teraflop (1000 Gigaflops) units. The coefficients on product characteristics are unintuitive. For example, one would expect marginal utility to be increasing in observed quality. This suggests that this variable could be correlated with unobserved attributes. Moreover, the coefficient on price has the expected sign, but implies that all supercomputer models have inelastic demands. This contradicts profit maximizing behavior. Therefore, the OLS results indicate that instrumental variables are required for consistent estimation of demand.

In addition to the price endogeneity problem, I must investigate whether observed quality (Rmax) is endogenous, or if its coefficient in the OLS regression is a result of inconsistent estimation. For this purpose, I run two 2SLS random-effects regressions: one where I only instrument for prices, and another where both Rmax and price are assumed endogenous[19]. Instruments must be carefully chosen in order to avoid correlation with unobserved attributes. One alternative is to assume that product characteristics are predetermined and therefore suitable instruments. Even though the richness of the Top500 data would allow me to consider the attributes ignored in the utility specification as instruments, only three of those were considered: processor speed, a dummy on whether the processor was produced by Cray, and a dummy variable for cluster systems (i.e. supercomputers which result of clustering less powerful systems). The speed of processors included in supercomputers is primarily explained by the exogenous evolution in processor technology during the 1990s. As processors are a key input of supercomputers, processor speed should be correlated with both price and observed quality. Moreover, the fact that some manufacturers use processors produced within the firm for some of their supercomputers (e.g. Cray) should be influencing the retail price. However, a dummy variable for Cray processors is unlikely to be correlated with the supercomputer unobserved characteristics, since about 7% of the machines in the sample use Cray processors[20]. The clustering property is shared by about 15% of the supercomputers in the sample, so it is unlikely to be correlated with the machine-specific unobserved attributes. Most of other variables in the data (e.g., number of processors) are not likely to be valid instruments. For example, supercomputers with more processors may be preferred for applications where

[19]Fixed-effects regressions could be run instead. However, the each supercomputer's computing speed, Rmax, is time-invariant. Hence it vanishes in regular fixed-effects regressions. Since the purpose here is solely to investigate whether one should instrument for Rmax, I use random-effects specification.

[20]As a robustness check, I run the IV logit regression without the dummy for Cray processors. The effect in the final estimates is negligible. However, this dummy variable is significant in the first-stage regressions. Therefore I preferred the instrument set where this dummy is included.

it is convenient to assign separate parts of the problem being solved to each processor. Treating (some) characteristics of the product as exogenous is as reasonable here as in previous work (e.g., Berry, Levinsohn and Pakes(1995), Nevo (2000,2001), Song (2006)). In order to ensure that no endogeneity concerns arise, it would be convenient to consider alternative sources of instruments. Cost shifters are an obvious candidate. McCallum (2002) provides series for memory prices for the period 1957-2006. Based on this series, I compute the average price of a Megabyte of memory (in 1998 dollars) for 1990-2005, and use it in the instrument set along with cluster dummy, Cray processor dummy and processor speed.

The results for instrumental variable logit regressions are shown on the last two columns of table 2.5. The third column corresponds to the case where I only instrument for prices, while the last one shows estimates when I instrument for both Rmax and price. Even thought the coefficient for Rmax has now the expected sign in the third column, the price coefficient still implies about 95% inelastic demands under the logit model. Since there are also other coefficients with unintuitive sign (e.g., firm frontier ranking), this evidence suggests that instruments for other potentially endogenous variables are necessary. The results in the last column indicate that Rmax is indeed a variable whose endogeneity must be controlled for. When instrumenting for both prices and computing speed, all the parameters have the expected sign and magnitude. The negative coefficients of the time effects are in line with the rapid obsolescence of supercomputers due to rapid innovation in the industry and incorporation of technical advances in other goods (e.g. mainframes, workstations).

Table 2.5. Results from Logit Model

	OLS Logit		IV Logit (i)		IV Logit (ii)	
Variable	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Price (in $1M)	-0.004	0.002	-0.035	0.025	-0.438	0.254
Rmax (in Teraflops)	-0.008	0.028	0.053	0.057	2.600	1.964
Network size (in 100)	0.065	0.018	0.087	0.028	0.452	0.251
MPP dummy	-0.225	0.037	-0.159	0.063	0.680	0.567
Firm rank	-0.329	0.058	-0.120	0.165	1.833	1.164
1991	-0.797	0.183	-0.889	0.196	-1.996	0.762
1992	-1.345	0.169	-1.509	0.210	-3.520	1.254
1993	-1.892	0.159	-2.072	0.223	-4.767	1.687
1994	-2.317	0.154	-2.596	0.269	-6.493	2.326
1995	-2.770	0.157	-3.054	0.280	-7.337	2.575
1996	-3.279	0.153	-3.462	0.219	-7.179	2.186
1997	-3.528	0.151	-3.428	0.170	-6.269	1.589
1998	-3.824	0.154	-3.912	0.181	-7.671	2.250
1999	-3.862	0.153	-3.996	0.199	-8.860	2.883
2000	-3.976	0.156	-4.160	0.229	-9.423	3.167
2001	-3.997	0.158	-4.233	0.250	-10.529	3.684
2002	-4.107	0.163	-4.225	0.205	-9.928	3.562
2003	-4.353	0.163	-4.578	0.269	-11.289	4.276
2004	-4.101	0.165	-4.409	0.324	-13.236	5.674
2005	-4.195	0.180	-4.539	0.362	-15.636	7.466
Constant	-3.292	0.141	-3.049	0.264	-2.624	3.470
Adjusted R^2	0.598		n.a.		n.a.	
%inelastic demands	100%		95.78%		9.97%	
Number of obs.	2204		2204		2204	

I estimate the full demand system using Nevo's (2001) estimator considering the same set of instruments. I draw 100 "consumers" for every year in the sample to numerically integrate the purchase probability with respect to consumer state vector D_{rt}. In order to get a parsimonious specification, I assume that switching costs are only significant if a consumer purchases from the nine most important firms: Cray, HP, IBM, NEC, Fujitsu, Hitachi, SGI, Intel and Sun Microsystems[21]. Estimates are shown on table 2.6. As discussed in the previous section, estimates for the constant, Rmax and MPP coefficients are recovered by running 2SLS regression of brand dummies on those variables using the same instrument set. All coefficients are significant at a 5% significance level except for MPP dummy and switching costs on HP, Fujitsu, Intel and Sun. Nevertheless, all coefficients have the expected sign. The price coefficient became more negative than under any of the above logit specifications. As before, the time effects suggest rapid obsolescence of outside alternative, suggesting that supercomputer substitutes become more attractive to consumers over time. Consumers seem to face higher switching costs upon purchasing from Cray and Japanese producers NEC, Fujitsu and Hitachi.

I estimate supply parameters (i.e., marginal costs and markups) by regressing prices of each product j on both the marginal cost function and dummy variables $B_{jft} = \mathbf{1}\{j \in \mathcal{F}_{ft}\}$. The coefficients on the latter are estimates of the markup charged by firm f at time t. Due to the fact that only 2204 observations are available, I must restrict the set of markup dummies to have a parsimonious specification for the regression equation. Hence, I use dummies for the firms which operate (or operated) in the market and had the highest market share[22]. Markups for other firms are derived by averaging regression

[21] Due to the short size of other firms' networks, consumers belonging to their network are rarely found in each of the 100 random "consumers" drawn each year. Therefore, it is hard to obtain estimates of switching costs incurred by consumers in their networks with acceptable precision.

[22] These firms are: Cray, Compaq, IBM, Sun, HP, SGI, Thinking Machines Corporation, Intel, Fujitsu, Hitachi, Kendall Square Research, MasPar, Meiko, nCube, NEC, Dell and Linux Networx

Table 2.6. Estimates from Full Model of Demand

Variable	Coef.	Std. Err.
Price (in $1M)	-0.588	0.194
Rmax (in Teraflops)	2.430	1.114
Network size (in 100)	0.529	0.179
MPP dummy	0.530	0.367
Firm rank	1.977	0.818
Sw. cost Cray	2.013	0.839
Sw. cost HP	0.512	0.478
Sw. cost IBM	1.129	0.411
Sw. cost NEC	2.145	0.549
Sw. cost Fujitsu	2.098	1.213
Sw. cost Hitachi	2.132	0.733
Sw. cost SGI	1.342	0.481
Sw. cost Intel	0.237	0.179
Sw. cost Sun	0.781	0.399
1991	-2.096	0.823
1992	-3.620	1.754
1993	-4.517	1.447
1994	-6.273	2.877
1995	-7.477	2.679
1996	-7.262	2.357
1997	-6.430	1.911
1998	-7.874	2.561
1999	-8.868	3.012
2000	-9.311	3.541
2001	-10.299	3.755
2002	-9.927	3.811
2003	-10.078	4.321
2004	-13.111	5.911
2005	-15.577	7.811
Constant	-2.874	1.470

residuals per firm and period[23]. Table 2.7 displays the estimates for marginal costs (coefficients on B_{jft} are omitted). The coefficients correspond to $1M units in 1998 constant dollars.

[23]This specification was robust to alternatives where other firm/period dummies were included in the regression.

Table 2.7. Marginal Costs Estimates (markup dummies' estimates omitted)

	without MPP dummy		with MPP dummy	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
Rmax*1990	2041.267	290.4437	2041.267	286.2619
Rmax*1991	1138	277.1594	1139.393	273.169
Rmax*1992	1697.555	227.9918	1649.904	224.792
Rmax*1993	420.4942	83.61033	363.5546	82.7285
Rmax*1994	389.6954	53.30058	345.6589	52.83499
Rmax*1995	323.2644	47.73932	294.9246	47.19174
Rmax*1996	276.1672	8.417946	265.7746	8.40281
Rmax*1997	168.7572	4.433223	163.8737	4.413922
Rmax*1998	127.7273	5.457506	122.1691	5.42582
Rmax*1999	67.61001	3.615437	64.78073	3.581755
Rmax*2000	27.94773	2.400988	27.56561	2.366925
Rmax*2001	16.83415	1.728735	16.59066	1.70413
Rmax*2002	4.771295	0.877638	4.125026	0.868952
Rmax*2003	6.483846	0.87091	6.300686	0.858691
Rmax*2004	1.814095	0.60428	2.305664	0.598897
Rmax*2005	1.138372	0.204216	0.921823	0.203178
Constant	3.207091	0.929455	2.292312	0.923533
MPP	-	-	3.004809	0.384797
Adjusted R^2	0.7623		0.767	

From the results, one can infer that marginal costs for producing a Teraflop are decreasing over time under both specifications. This evidence agrees with two facts about supercomputer inputs: (i) processors are increasingly cheaper and of improved speed, and (ii) memory prices per Megabyte are falling over time[24], even though the improvements in memory technology are far less considerable than the ones for processors. The MPP dummy essentially affects the constant part of the marginal costs, but adds very little explanatory power. However, since its coefficient is statistically significant, the specification with MPP dummy was preferred.

[24]Source: http://www.jcmit.com/mem2006.htm

Another important implication to be derived from the marginal cost estimates concerns its relation with Moore's Law. Even though Moore's original statement concerned transistor technology, it is common to cite Moore's Law to refer to the rapidly continuing advance in computing power per unit cost. In the context of computing technology, Moore's Law prescribes that the computing power per unit cost doubles every 18 months[25]. With the estimates of marginal cost at hand, it is possible to check this law by estimating the temporal evolution of marginal costs for a fixed level of computing power. The results for marginal costs suggest a decreasing exponential evolution over time, I fit the following equation by OLS

$$\ln(mc_t) = \mu_0 + \mu_1 t + v_t \quad t = 1990, ..., 2005$$

where mc_t corresponds to marginal costs of producing 1 Teraflop at time t. This series corresponds to the estimates on each coefficient columns of table 2.7. I estimate the above equation under the two marginal cost specifications considered on table 2.7 (i.e., with and without MPP dummy) as a robustness check. Table 2.8 presents the results for the temporal evolution marginal costs of producing 1 Teraflop. The results for the two sets of estimates are very similar, so I will focus the analysis on the results of the second column. The instantaneous rate of decrease in marginal costs is 40.4%, which implies an annual decrease in marginal costs of approximately 33.22%. So the ratio of

[25]This is the most popular version of Moore's Law, even though Moore's original claim was that transistor density of integrated circuits, with respect to minimum component cost, doubles every 24 months.

Table 2.8. Marginal Costs and Moore's Law

	without MPP dummy		with MPP dummy	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
Year	-0.440	0.020	-0.404	0.018
Constant	883.077	40.670	811.096	35.893
Adjusted R^2	0.9688		0.9711	

computing power over unit cost is increasing by 67.79% per year. Therefore, the marginal cost estimates closely resemble Moore's Law.

I estimate the markup policy by fitting the sieves function proposed in the previous section to the markup series recovered after supply estimation. The results are displayed on Table 2.9. No Chebyshev polynomials of order bigger than two were considered due both to overfitting concerns and to lack of significance those terms in all preliminary regressions. Even though the coefficients of this regression are hard to interpret, the results deserve some comment. First, the coefficients on N_{0t} and N_{ft} are apparently at odds with economic intuition: one should expect firms to charge higher markups the more consumers belong to either the outside alternative network or its own network. However, the higher order polynomials on these variables have positive and significant coefficients. This suggest that firms would charge low markups to tip the market for low own and outside network values, but charge significant markups when these are of "reasonable" size.

The choice of regressors for the Tobit type I estimates of quality investment and frontier extension rates followed similar guidelines as for markup policy estimation. Upon computing quality investment using marginal cost and demand estimates, I estimate its

Table 2.9. Markup Policy Estimates

	Estimates	Std. Error
Constant	0.113	0.982
N_{0t}	-1.229	0.349
κ_{ft}	0.382	0.128
N_{ft}	-1.401	0.853
$\exp(\nu_{ft})$	1.037	0.231
κ_{-ft}	-0.034	0.053
N_{-ft}	-0.129	0.091
$\exp(\nu_{-ft})$	-0.101	0.073
$T_2(N_{0t})$	2.486	0.649
$T_2(\kappa_{ft})$	2.085	0.053
$T_2(N_{ft})$	6.303	2.001
$T_2(\exp(\nu_{ft}))$	-0.122	0.070
$T_2(\kappa_{-ft})$	-0.0034	0.001
$T_2(N_{-ft})$	-0.023	0.059
$T_2(\exp(\nu_{-ft}))$	-0.122	0.002
$N_{ft} \times N_{-ft}$	4.872	0.452
$\exp(\nu_{ft}) \times \exp(\nu_{-ft})$	-2.230	3.327
$\kappa_{ft} \times \kappa_{-ft}$	1.560	0.549
$\exp(\nu_{ft}) \times N_{ft}$	1.263	0.490
$\exp(\nu_{-ft}) \times N_{ft}$	-0.251	0.138
$\exp(\nu_{ft}) \times N_{-ft}$	-0.908	0.215
No. obs	218	
Adj. R^2	0.791	

policy using the Tobit specification proposed in the previous section. The frontier extension policy is also estimated using a Tobit Type I. Results are presented on Tables 2.10 and 2.11, respectively.

The estimation of demand parameters allows me to compute the transition of all firm networks except for $N_{0,t}$. I estimate the remaining parameters of outside alternative's network transition (i.e, ϕ_0 and ϕ_1) by using the market size transition equation derived

Table 2.10. Quality Investment Policy Estimates

Variable	Estimates	Std. Error
Constant	0.294	0.095
N_{0t}	0.311	0.216
κ_{ft}	-0.176	0.099
N_{ft}	0.019	0 .055
$\exp(\nu_{ft})$	-0.993	0.025
κ_{-ft}	-1.773	1.590
N_{-ft}	-1.670	0.121
$\exp(\nu_{-ft})$	-1.369	0.750
$T_2(N_{0t})$	14.895	4.936
$T_2(\kappa_{ft})$	-5.703	1.729
$T_2(N_{ft})$	1.389	0.668
$T_2(\exp(\nu_{ft}))$	3.569	1.436
$T_2(\kappa_{-ft})$	3.303	0.946
$T_2(N_{-ft})$	2.140	0.812
$T_2(\exp(\nu_{-ft}))$	2.265	1.360
$\kappa_{ft} \times \kappa_{-ft}$	5.145	0.386
No. Obs:	177	
Log-Likelihood	-319.423	

on the previous section. This equation consists on $M_{t+1} = \phi_0 + \phi_1 M_t + \eta_{t+1}^{N_0}$, where M_t denotes market size. Following Benkard (2004), I defined M_t as the number of of new and used supercomputers in use in year t, which is approximated by the sum of all supercomputers sold between year $t-5$ and t (see Appendix A for details). In principle, the specification proposed for this state's transition can be estimated using OLS. However, one may suspect that the shock $\eta_{t+1}^{N_0}$ is correlated with M_t. I use the Markovian structure of the game to propose M_{t-1} as an instrument for M_t. The results of this regression using both OLS and 2SLS are presented on Table 2.12. The AR(1) specification in market size seems to capture market size dynamics rather well. The inclusion of other regressors (e.g., sum of firm ranks and markup-adjusted inclusive values) would not only bring little

Table 2.11. Frontier Extension Policy Estimates

Variable	Estimates	Std. Error
Constant	-6.294	11.905
N_{0t}	0.401	0.166
κ_{ft}	-8.176	0.099
N_{ft}	-0.289	0 .099
$\exp(\nu_{ft})$	-7.493	4.025
κ_{-ft}	1.443	0.590
N_{-ft}	7.257	5.111
$\exp(\nu_{-ft})$	4.369	1.890
$T_2(N_{0t})$	-1.895	1.136
$T_2(\kappa_{ft})$	5.113	1.888
$T_2(N_{ft})$	2.088	1.206
$T_2(\exp(\nu_{ft}))$	3.569	1.222
$N_{ft} \times N_{-ft}$	-4.333	2.312
$\exp(\nu_{ft}) \times \exp(\nu_{-ft})$	-5.301	3.212
$\kappa_{ft} \times \kappa_{-ft}$	3.211	1.782
No. Obs:	177	
Log-Likelihood	-284.333	

Table 2.12. Market Size Regression Estimates

	OLS REGRESSION		2SLS REGRESSION	
Variable	Coef.	Std. Err.	Coef.	Std. Err.
Constant	348.685	117.898	455.250	127.931
M_t	0.927	0.047	0.884	0.047
Adjusted R^2	0.9688		0.978	
Number of obs.	15		14	

explanatory power to the market size transition, but also raise overfitting concerns. The estimates in the two columns differ more significantly in the constant term. I considered the 2SLS estimates for the remainder of the work[26].

[26]This decision was also based on an Hausman test, which rejected the hypothesis of equal coefficients in both regressions.

Table 2.13. Potential Entrant's Average Quality Transition Estimates

Variable	Estimates	Std. Error
Constant	0.051	0.009
N_{0t}	0.401	0.166
$sum(N_{ft})$	-0.419	0 .094
$sum(\exp(\nu_{ft}))$	-7.493	36.025
$sum(\kappa_{ft})$	-34.443	22.590
No. Obs:	66	
Log-Likelihood	-192.015	

The transitions for markup-adjusted inclusive values and technological states depend not only on observed states and actions, but also on the shocks which impact market entry decisions of inactive players. As described in the previous section, the evolution of state transitions for potential entrants are described by the equations

$$\exp(v_{ft+1}) = \max\{0, w_1(\mathbf{s}_t) + \varepsilon_{3ft}\}$$

$$W_{f,t+1} = w_2(\mathbf{s}_t) + \varepsilon_{4ft} if \exp(v_{ft+1}) > 0$$

where

$$W_{f,t+1} \equiv \ln(h^E_{ft+1}) - \ln\left(\max_{i=1,\dots,F}\{h_{it}\}\right)$$

Assuming that ε_{3ft} is a zero-mean normal shock, the first equation is estimated using a Tobit Type I specification. The equation on $W_{f,t+1}$ is estimated by OLS conditional on the set of inactive players who decided to enter at period t. Table 2.13 and 2.14 display the results for these regressions, respectively.

Table 2.14. Entrant's Frontier Transition Estimates

Variable	Estimates	Std. Error
Constant	-2.294	1.395
N_{0t}	0.291	0.109
$sum(N_{ft})$	-0.045	0.012
$sum(\exp(\nu_{ft}))$	-0.009	0 .004
$sum(\kappa_{ft})$	-8.783	0.425
No. obs	41	
Adj. R^2	0.635	

Table 2.15. Incumbent's Average Quality Transition Estimates

	Estimates	Std. Error
Constant	-.7961032	.2786951
ν_{ft}	.8352611	.0552814
$\ln(1 + A_{ft})$	1.3761631	.620071
No. obs	177	
Adj. R^2	0.675	

I am left to estimate the markup-adjusted inclusive value transitions of incumbent firms. The proposed specification for this transition is

$$\upsilon_{ft+1} = d_0 + (1 - d_1)\upsilon_{ft} + (1 + d_2)\ln(1 + A_{ft}) + \eta_{ft}$$

which I estimate by random-effects GLS. The results are displayed on Table 2.15. The proposed specification seems to be capturing average quality dynamics reasonably well. As expected, average quality dynamics seem to depend primarily from past average quality values and on the logarithm of gross quality investment rate $1+A_{ft}$. As for the the market size, the Bera-Jarque normality test didn't reject the normality hypothesis for the error term. So I assume that the error term of both regressions is normality distributed with zero mean in the second stage of estimation.

Table 2.16. Dynamic Parameter Estimates

	Variable	Parameter	Estimate	Std. Error
Benefit outside profit terms				
	κ_{ft}	χ_1	612.348	289.312
	κ_{ft}^2	χ_2	32.918	19.423
Bandwidth parameters				
	$\kappa_{it}-\kappa_{ft}$	h_I	0.231	0.056
	$\upsilon_{it}-\upsilon_{ft}$	h_A	2.098	1.571
Frontier extension costs				
	Constant	c_0	78.092	30.897
	Firm rank (κ_{ft})	c_1	-42.234	19.434
	Innovation rate (I_{ft})	c_2	4.761	2.082
	Square Innovation rate (I_{ft}^2)	c_3	0.971	0.512
	Tech. gap on laggards	c_4	-2.871	2.109
	Tech. gap on advanced rivals	c_5	-8.598	4.052
Quality investment costs				
	A_{ft} if $A_{ft} > 0$	c_0^A	0.231	0.115
	A_{ft}^2 if $A_{ft} > 0$	c_1^A	0.093	0.067
	A_{ft} if $A_{ft} < 0$	c_2^A	2.096	0.781
	A_{ft}^2 if $A_{ft} < 0$	c_3^A	-15.012	4.671
Fixed costs	$1\{\exp(\upsilon_{ft}) > 0\}$	C_f	83.902	50.412
Auxiliary parameters	Lagrange multipliers	μ^I	-2.889	1.923
		μ^A	-0.551	0.449
	Innovation shock variance	σ_{ε_1}	39.041	11.231

The GMM estimates of these dynamic parameters an presented on Table 2.16. The approximation parameter for the Chebyshev polynomials in value fuction approximant, K, was set to 3. This implies 2187 coefficients for the integrated value function approximation. I randomly sampled 2187 state vectors for computing estimates of these coefficients for each given trial value of the dynamic parameters. Based on the empirical correlation between firm frontier rank and markup-adjusted inclusive value, I calibrate $\rho = 0.3156$. I also set the discount factor to $\beta = 0.95$. Bootstrap standard errors were generated from sampling 100 bootstrap samples of the data with replacement and estimating the dynamic parameters with each subsample. I chose a polynomial of order 4 for the density estimator of Gallant and Tauchen. These were not statistically significant in preliminary estimations and therefore removed from the final specification. Estimates suggest that the benefit of being at a certain technological state commands a significant portion of the firm's payoff. For example, the parameters of the benefit function imply that that a technological leader would collect an annual benefit of about \$645M (in 1998 dollars) if he were the only firm in the market. Estimates also indicate that fixed innovation costs are the most significant portion of the overall innovation cost. However, these costs depend considerably on the firm's frontier positioning. Firms can benefit both from the knowledge of its most advanced rivals and from laggards, but both effects are very small compared to the other fixed cost components. More advanced firms tend to have significant reductions in innovation costs. Marginal costs of frontier innovation are also small compared with the fixed cost components. The costs of quality investment are also relatively small, which is in line with the frequent product introduction observed in this industry. The estimates of c_2^A and c_3^A imply that firms on average receive a scrap value of \$13M if they decide to leave the supercomputer industry. This value is very small compared with the fixed costs of being active in this industry.

2.6. Policy experiments

The motivation for constructing a dynamic model for the supercomputer industry was its ability to simulate counterfactual policy experiments upon recovering the underlying primitives from the data. Several important questions can be answered with the model. For example, what would be the impact of a permanent demand shock on innovation rates? Would a merger between two firms increase equilibrium innovation rates? In this Chapter, I use the structural model to answer a long-standing question in the Industrial Organization literature: how does technological progress depend on competition? The strategy to obtain the answer consists in two steps. First, I solve the MPNE of the game for a given set of observable states. For simplicity, I integrate out all private information in each firm's problem. That is, taking as given rival policies, observed states and the value function estimate, each firm chooses its optimal policies to maximize the expected discounted stream of payoffs, defined as

$$\begin{aligned} V_f(\mathbf{s}) &= \underset{\sigma_f(\mathbf{s})}{Max} \int_{\varepsilon} \{\pi_f(\sigma_f(\mathbf{s}), \sigma_{-f}(\mathbf{s}, \varepsilon_{-f}), \mathbf{s}, \varepsilon_f) \\ &\quad + \beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}} \left[h_f(\mathbf{s}', \theta) | \mathbf{s}, \sigma_f(\mathbf{s}), \sigma_{-f}(\mathbf{s}, \varepsilon_{-f}) \right] \Big\} \, dF(\varepsilon) \end{aligned}$$

Second, taking as given an initial profile of firm computing speed frontiers, I simulate the maximal computing speed available in the industry using the equilibrium frontier extension rates. Since the analysis of welfare from innovation has received considerable attention in the literature (e.g., Trajtenberg (1989), Green and Scotchmer (1995), Bresnahan, Stern and Trajtenberg (1997) and Bresnahan and Greenstein (1999)), I also compute total welfare as defined below.

Ideally, one would like to perform the proposed experiment with an arbitrary number of firms in order to completely describe the effects of competition in innovation. However, I can only use the estimated integrated value function approximant if there are no changes in the structural parameters of the game. Since the purpose here is to have a parsimonious

Table 2.17. Technological Frontiers of the Top 3 firms in 1997

Firm	Frontier in Giga-FLOPS
Intel	1338
Cray/SGI	815
Hitachi	368.2

analysis of the effect of market structure on frontier innovation, I conduct the experiment under three different scenarios: monopoly, duopoly and three-firm market. As I assume that there can exist at most three firms in this market, I must compute the integrated value function in the simulations[27].

The simulation routines follow the exposition of Miranda and Fackler (2002, sections 8.5 and 9.8) for dynamic games solution and simulation. As before, I use first-kind Chebyshev polynomials up to order three to compute the approximate value function, which is again assumed to satisfy anonymity and symmetry. As an initial setup, I consider the three most advanced supercomputer producers in 1997, whose frontiers are presented on Table 2.17. For all years except 1997, the values underlying the graphs below regard averaged values across 20000 simulations.

Under the assumptions on demand, consumer welfare is given by

$$CW_t = \frac{1}{\alpha}\sum_{k=0}^{F} N_{kt} \ln \left(\sum_{d=1}^{F} \exp(-\alpha M k p_{dt} + \tau_1 \kappa_{dt} + \nu_{dt} + \psi N_{dt} + \Lambda_d \mathbf{1}\{k = d\}) \right)$$

[27]Note that here the only change in the model structure considered here is the number of firms which can operate in the market. So the estimated value function approximant could be used in simulation if all fifteen slots were considered. I exploit this property in Chapter 3 in the context of merger analysis.

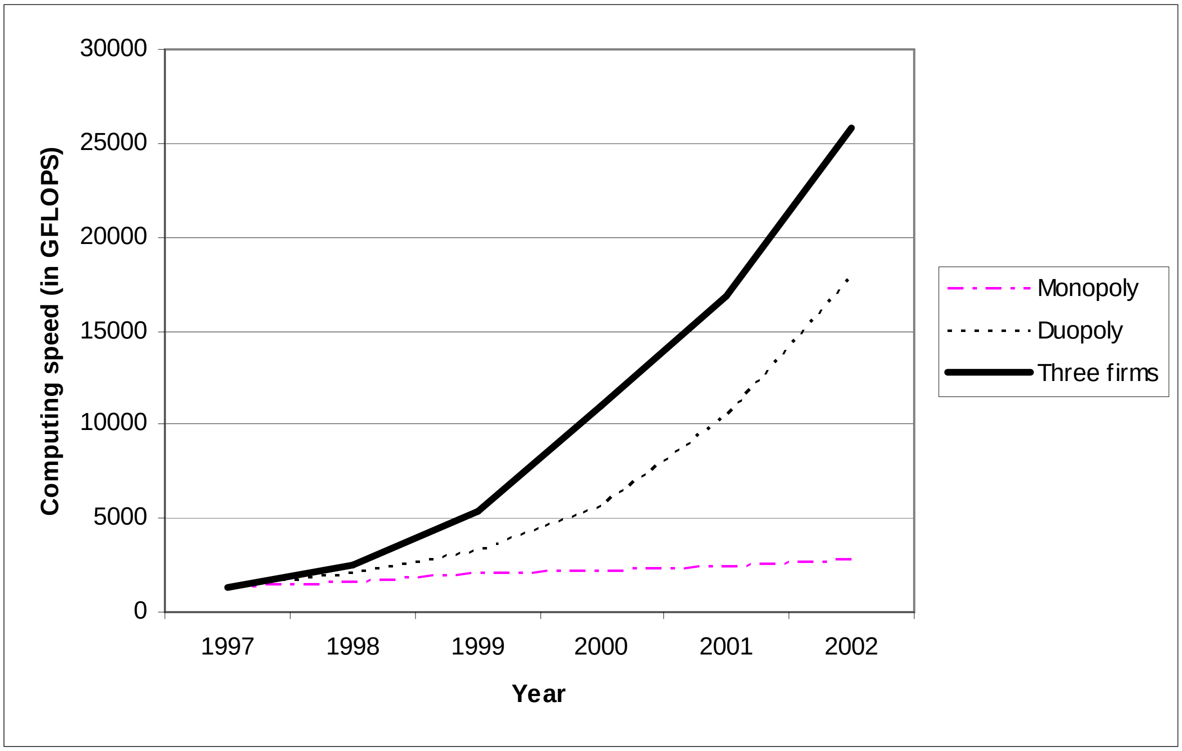

Figure 2.1. Maximum computing speed under different market structures

while total welfare is defined as the sum of consumer welfare with the aggregate flow payoff of all the firms at time t. That is,

$$TW_t = CW_t + \sum_{f=1}^{F} \left(\Pi_{ft} + \Upsilon_{ft} - C^I_{ft} - C^A_{ft}\right)$$

Results suggest that competition significantly encourages technological progress. The differential between the three-firm case and the duopoly one is not as large as the one between monopoly and duopoly maximal computing speeds. The influence of the assumption that only three firms can operate in the market is a potential explanation, which will be investigated in future research.

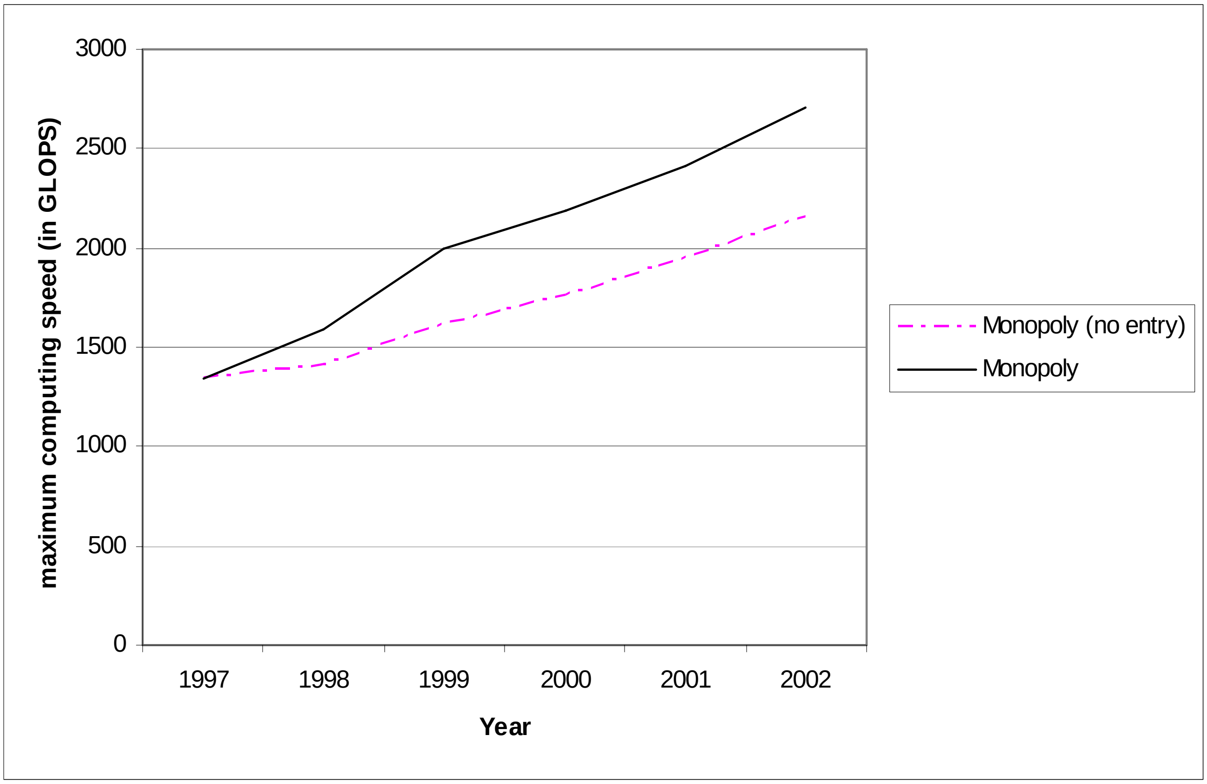

Figure 2.2. Maximum computing speed under monopoly (with and without entry)

An important question in the innovation literature is how much innovation from a monopolist is due to the threat of entry. My framework allows me to shed light on this question by simulating maximal computing speed patterns under the assumption that no more than one firm can operate in this market. That is, I compare the monopoly maximum computing speed displayed on Figure 2.1 with the simulation outcome of the case where only a multiproduct monopolist can operate in this market. The paths displayed on figure show that this gain can be substantial. A monopolist under the threat of entry will yield a maximal computing speed in the industry at least 12.8% bigger than under no entry threat. Like for the maximum computing speed evolution under three different

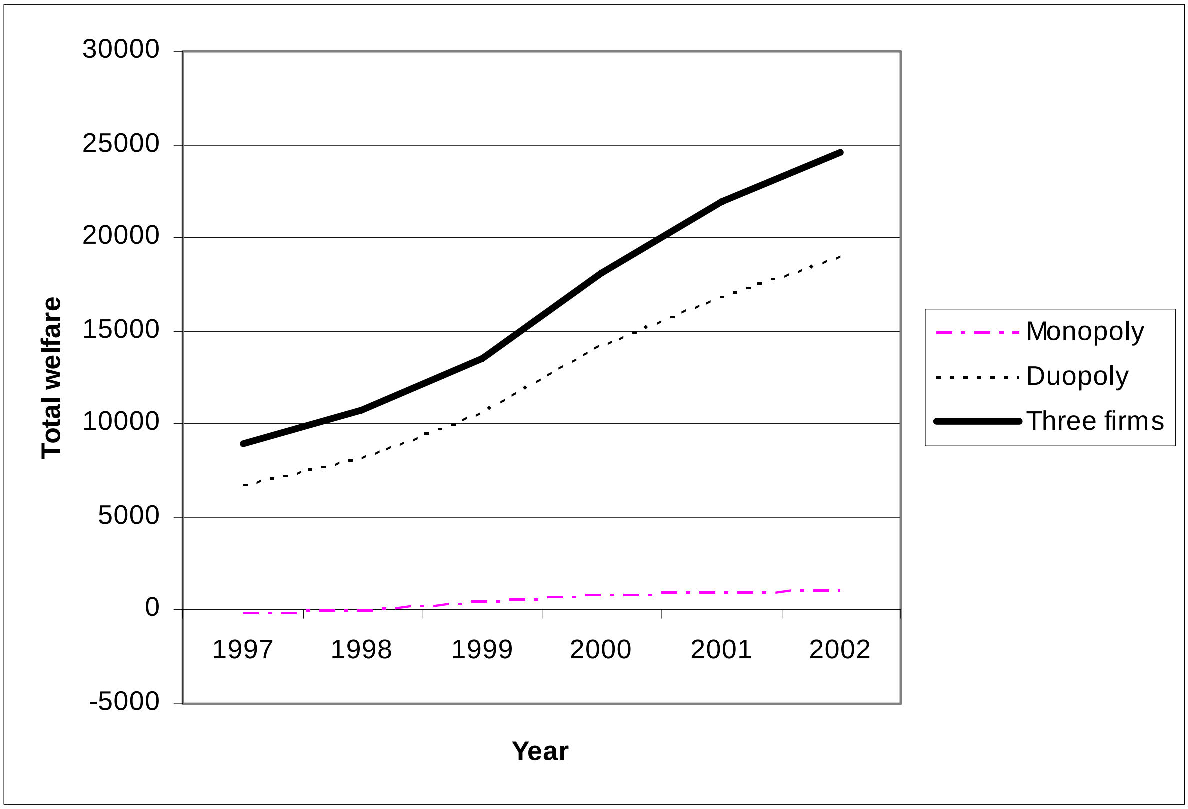

Figure 2.3. Simulated total welfare under different market structures

market structures, increases in welfare from monopoly to duopoly are considerably higher compared to the case of moving from duopoly to a three-firm market.

2.7. Conclusion

This paper proposes a structural model of competition where firms expand their technological frontier, invest in product quality and set constant absolute markups for their products. Strong assumptions imposed in other innovation models proposed in the literature are avoided by building on recent refinements to the dynamic oligopoly games literature. The additional difficulty of modeling innovation in a differentiated products industry is circumvented by using Nevo and Rossi's (2007) framework, where model primitives imply equilibria in which firms charge constant absolute markups. The model can be brought to data by using a variant of the two-step estimation method introduced in the literature by Bajari, Benkard and Levin (2006). Identification of the model parameters is based on the orthogonality of selected instruments to unobserved product quality, and optimality conditions on markups, quality investment and frontier extension rates. Estimates suggest that firms derive considerable benefits outside product market profits from improved technological frontiers. The fact that the incremental payoff from innovating is higher when a firm is in "neck-and-neck" competition with technologically similar rivals (i.e., the "selection effect" introduced in the literature by Aghion, Harris, Howitt and Vickers (2001)) is the source of identification of these benefits.

The paper also quantifies positive externalities of technological states to innovation costs, which indicate that advances on both the firm's maximal computing speed and the ones of its most technologically advanced rivals considerably reduce the fixed costs of innovation investment. The model estimates are also used to assess how does technological progress depend on competition. In line with Aghion, Harris, Howitt and Vickers (2001) and Aghion, Bloom, Blundell, Griffiths and Howitt (2003), I find that increased levels of competition have positive effects on the evolution of supercomputer technology. The fact that firms innovate primarily to escape competition with "neck-and-neck" rivals also impacts positively total welfare.

Other important questions could be addressed with the model and its estimates. For example, what market structures would yield maximal welfare and rates of product improvement? One could measure the impact of both network externalities and switching costs on competition and welfare. Another issue of particular interest is the impact of mergers on markups, welfare and product quality dynamics. Alternative applications of the results may include, for example, the effects of an exogenous shock in demand on innovation patterns, or the consequences of a given subsidy scheme for innovation behavior. In these cases, the simulation of counterfactuals can be done by making appropriate changes on either the state vector or the firms' cost structure. The study of these and other questions is left for future research.

CHAPTER 3

Dynamic merger analysis in differentiated product industries

3.1. Introduction

When competing in dimensions affecting long-run payoffs, firms account for the impact of their actions on both current and future payoffs. In these situations, however, equilibrium firm policies will differ from the ones implied by static competition models commonly used in merger analysis. Static merger evaluation methods can therefore lead to misleading conclusions about post-merger welfare and equilibrium prices when applied to dynamically competitive industries[1]. Despite the importance of these facts for antitrust analysis, much of the existing literature on merger evaluation is tied to static competition models. In this Chapter, I take a step toward filling this void. I evaluate the long-run impact of an actual merger in the supercomputer industry on consumer welfare and maximum computing speed available in this market.

I propose a framework for evaluating the long-term effects of a merger for differentiated products industries. My general strategy is to model competition so that all payoff-relevant features of firms can be encoded into a state vector. I use data from the supercomputer industry to estimate the model proposed in Chapter 2. Using both the model estimates (see Chapter 2) and a Markov-Perfect Nash Equilibrium (MPNE)

[1]See Nevo (2000) for a discussion.

assumption, I simulate the new equilibrium that would result from the merger for five consecutive years. I compare the implied maximum available computing speed and consumer welfare with the ones for equilibrium without merger.

The supercomputer industry is well suited to assess the performance of dynamic methods of merger simulation in differentiated product industries. In 1995, Hewlett-Packard (HP) purchased Convex for $150 million. In February 1996, SGI purchased Cray Inc. for $740M. Three months later, SGI sold the SPARC/Solaris part of the Cray product line to Sun Microsystems for an undisclosed amount. On March 31, 2000, SGI sold their Cray division to Tera Computer Company for $50M[2]. Shortly after this acquisition, Tera Computer Company was renamed Cray Research. In 2002, HP completes its merger with Compaq Computer for $25 billion. In this Chapter, I focus on the long-term effects of the merger betweenHewlett-Packard and Convex. Results suggest that this merger fostered technological progress in the industry in the form of increased maximal available computing speed. This merger also implied small annual consumer welfare losses in the period 1995-2000.

One of the main challenges is how to simulate mergers with several firms while keeping computational tractability. I deal with this difficulty by using the value function approximation recovered in estimation when computing equilibria. This approach has the advantage of eliminating computations of value function parameters during equilibrium calculations at the cost of introducing estimation error in the algorithm.

The rest of this chapter is organized as follows. Section 3.2 describes the merger simulation method and welfare computation details. Section 3.3 applies these methods to

---

[2]This value consists on the sale value of $35M plus one million SGI shares outstanding.

a concrete merger analysis. Section 3.4 concludes and discusses extensions. The specific details of the HP-Convex merger simulation algorithm can be found in Appendix C.

3.2. Merger effects simulation methods

3.2.1. A dynamic model of competition

In this section I present a model of dynamic competition to simulate of the long-run effects of a merger. Time is assumed discrete with an infinite horizon, and indexed by $t \in 1, 2, ..., \infty$. There are F firms in this market, denoted $f = 1, ..., F$, which choose L actions simultaneously at each period. I denote $\sigma_{ft} \equiv (\sigma_{f1t}, ..., \sigma_{fLt})$ as the vector of firm f's actions at time t. All observed payoff-relevant information at time t is summarized by a vector of state variables $\mathbf{s}_t$, which is assumed common knowledge to all firms in the industry. I assume that before choosing its actions, each firm receives a private shock vector ε_{ft}, drawn independently across agents and over time.

Before defining each firm's intertemporal optimization problem, I assume that firms restrict attention to Markovian pure strategies. That is, all players' choices in equilibrium are deterministic functions of payoff-relevant information. Formally, this corresponds to a map between the set of states observed by each firm and their strategy set. I assume that firms seek to maximize the discounted sum of payoffs conditional on a set of observed states $\mathbf{s}$ and private information ε_f. Under the assumption that firms share a common discount factor $\beta < 1$, the firm's problem is defined by the Bellman equation

$$
\begin{aligned}
V_f(\mathbf{s},\varepsilon_f) &= \int_{\varepsilon_{-f}} \underset{\sigma_f(\mathbf{s},\varepsilon_f)}{Max} \{\pi_f(\sigma_f(\mathbf{s},\varepsilon_f),\sigma_{-f}(\mathbf{s},\varepsilon_{-f}),\mathbf{s},\varepsilon_f) \\
&\quad +\beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}}\left[V_f(\mathbf{s}',\varepsilon_f')|\mathbf{s},\sigma_f(\mathbf{s},\varepsilon_f),\sigma_{-f}(\mathbf{s},\varepsilon_{-f})\right]\right\} dF(\varepsilon_{-f})
\end{aligned}
$$

where $\pi_f(.)$ is the flow payoff function, and σ_{-f} denotes the strategy profile of firm f's rivals. In what follows, it is convenient to integrate out firm's private information in each firm's problem. Assuming that private information is independent over time, players and observed states, the integrated (ex-ante) value function of firm f is given by

$$
\begin{aligned}
V_f(\mathbf{s}) &= \int_{\varepsilon} \underset{\sigma_f(\mathbf{s},\varepsilon_f)}{Max} \{\pi_f(\sigma_f(\mathbf{s},\varepsilon_f),\sigma_{-f}(\mathbf{s},\varepsilon_{-f}),\mathbf{s},\varepsilon_f) \\
&\quad +\beta E_{\mathbf{s},,\sigma_f,\sigma_{-f}}\left[V_f(\mathbf{s}')|\mathbf{s},\sigma_f(\mathbf{s},\varepsilon_f),\sigma_{-f}(\mathbf{s},\varepsilon_{-f})\right]\right\} dF(\varepsilon)
\end{aligned}
$$

A profile of Markov strategies $\sigma = (\sigma_1, ..., \sigma_F)$ is a Markov Perfect Nash Equilibrium (MPNE) if, given the opponent profile σ_{-f}, the firm's strategy σ_f yields a higher discounted sum of payoffs than any alternative Markovian strategy σ'_f. That is, σ is an MPNE if, for all firms $f = 1, ..., F$ and Markov strategies σ'_f, we have $V_f(\mathbf{s},\varepsilon_f|\sigma) \geq V_f(\mathbf{s},\varepsilon_f|\sigma'_f,\sigma_{-f})$. Existence of pure-strategy MPNE for this game follows from Theorem 2 in Jenkins, Liu, Matzkin and McFadden (2004). I assume that there is unique MPNE in the game.

3.2.2. Simulation algorithm

My general strategy in simulating the long-term effects of a merger in a differentiated products industry consists of two steps. First, using model estimates, I simulate the MPNE that would result from the merger for five consecutive years. Second, I compare the implied statistics of interest (e.g. maximum available computing speed, consumer welfare) with the ones of MPNE without merger. I assume that a merger can be completely described by changing the observed state vector. This change takes the form of a firm absorbing the states of another firm while the acquired firm leaves its slot open to a potential entrant. I denote $\mathbf{s}_t^{\Pr e}$ and $\mathbf{s}_t^{Post}$ as the pre- and post-merger observed state vectors, respectively.

If the functional form for the integrated value function were known, the computation of MPNE would follow methods similar to solving Bayesian Nash Equilibria (BNE). However, the integrated value function does not usually have a closed form and its direct computation requires solving the high-order problem defined by the Bellman equation. In practice, this implies a prohibitive computational burden. Even though improved methods of dynamic equilibrium computation have been proposed in recent literature (e.g., Rui and Miranda (1996), Judd (1998), Vedenov and Miranda (2001), Miranda and Fackler (2002) and Doraszelski (2003)), the available approaches are still too computationally burdensome when there are multiple players in the game. I deal with this problem by replacing the integrated value function with a sieves approximation, which is assumed to be recovered during model estimation.

Assuming further that the firm problem is differentiable with respect to firm controls, each firm action $\sigma_{fl} \in \sigma_f \equiv (\sigma_{f1}, ..., \sigma_{fL})$ must satisfy the first-order condition

$$\int_{\varepsilon_{-f}} \left(\frac{\partial \pi_f(\sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-ft}), \mathbf{s}, \varepsilon_{ft})}{\partial \sigma_{fl}} + \beta E \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial \sigma_{fl}} | \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t \right] \right) dF(\varepsilon_{-f}) = 0$$

where $\Delta(S)$ denotes the dimension of the observed state space, V_f is short-hand notation for the ex-ante value function, and g_k represents the k^{th} entry in the observed state vector at period $t+1$. In what follows, V_f is replaced by its estimate $\hat{h}_f$. Letting $\tilde{t}$ be the period at which the merger would take place, the merger effects simulation algorithm can be outlined as follows.

- Given the pre-merger observed state vector $\mathbf{s}_{\tilde{t}}^{\text{Pre}}$, compute the MPNE by solving the system of first-order conditions on firm strategies[3];
- Compute all statistics of interest for the analysis (e.g., consumer welfare, industry technological frontier) using the BNE solutions under both $\mathbf{s}_{\tilde{t}}^{\text{Pre}}$ and $\mathbf{s}_{\tilde{t}}^{Post}$;
- Generate an observed state vector for period $\tilde{t}+1$ by sampling a draw from the transition probabilities evaluated at the BNE solution under both pre- and post-merger states;
- Repeat the first three steps considering the random draws of the previous step as the observed state;

[3]This computation requires a value for each player's private information. Since this is not observed by the researcher, one can either (i) sample values from the privately known state distribution, should this distribution be available, or (ii) integrate out private information in the firm's problem.

- Repeat the last step $\tilde{T}$ times to generate simulated paths under merger and no-merger scenarios up to year $\tilde{t} + \tilde{T}$;
- Repeat all steps R times, and compute the expected statistics of interest (e.g. compensating variation) paths by averaging results across the R simulations.
- Redo all the previous step using instead $\mathbf{s}_{\tilde{t}}^{Post}$ as the initial observed state;

This algorithm is valid if no model primitives are changed and if the MPNE assumption holds both with and without a merger. Providing a simulation routine where these assumptions do not hold goes beyond the scope of this Chapter. In cases where any of these conditions are not met, the ex-ante value function must be computed when solving for MPNE. Nevertheless, the ex-ante value function estimate $\hat{h}_f$ can be used as an initial guess in that higher-order computational problem, or as an approximation for the true ex-ante value function under the new primitives.

One may question how realistic is the assumption that the merger only implies changes in observed state values. In the next section I provide a concrete example in the context of the model proposed in Chapter 2. I will assume no changes in model primitives. Many arguments on behalf of mergers concern cost savings, and therefore this assumption may be strong. Nevertheless, the simulation of long-run merger effects under no changes in model primitives is interesting in its own right. For example, it can provide information on what part of the merger effects is due to firm strategic interactions alone.

3.3. Evaluating a merger

In this section I apply the proposed merger simulation method to a specific merger in the supercomputer industry. I am interested on assessing the impact of this merger on the maximal computing speed available in the industry and on consumer welfare not only in 1995, but also on the years following the merger. For this purpose, I will consider the model developed in Chapter 2 and its estimates in the simulation exercise. In principle, I could also examine price and quantity effects of the merger. However, supercomputer models are typically removed from firm product portfolios two years after its introduction due to its rapid obsolescence. Due to the frequent product introduction and scrappage of products in this market, I will instead focus on the merger impacts on maximal computing speed and consumer welfare.

Table 3.1 describes the industry state in 1995, as well as firm (relative) market shares. The exponential of markup-adjusted inclusive values were computed using demand and marginal cost estimates from Chapter 2. There are a total of fifteen incumbent firms in that year, which is the maximum number of active firms per year observed in the sample. Two of those fifteen firms - Sun Microsystems and Compaq - were players who decided to enter the supercomputer market in the previous period. Intel was the technological leader, followed closely by Fujitsu and Cray. The two merging firms, HP and Convex, were among those with the lowest technological frontier in the industry. Their network sizes are also among the lowest ones in the supercomputer market. Only 23 buyers made their last purchase from Convex, while HP had an insignificant network (14 consumers). These figures are explained by the fact that these firms entered the industry few years before 1995: Convex became an active player in the industry in 1991 while HP only entered the market in 1993. The figures on markup-adjusted inclusive values indicate that HP supplied an average quality well above Convex, even though two other players (Cray and IBM) supplied higher average quality. The last column in the table displays

Table 3.1. Observed state vector and relative market shares in 1995

Firm	Firmrank ($\kappa_{f,1995}$)	Network ($N_{f,1995}$)	$\exp(\upsilon_{f,1995})$	Share (Q_f/Q)
Cray	0.701	161	0.544	16.049%
Compaq	0.035	0	0.326	0.494%
Convex	0.020	23	0.281	0.494%
IBM	0.616	82	0.698	25.432%
Sun Microsystems	0.125	0	0.216	0.247%
Hewlett-Packard	0.052	14	0.481	3.457%
SGI	0.186	115	0.547	46.667%
Thinking Machines	0.416	53	0.351	0.741%
Intel	1.000	60	0.426	1.975%
Fujitsu	0.865	63	0.341	1.235%
Hitachi	0.198	20	0.180	0.247%
Kendall Square	0.048	23	0.309	0.494%
Meiko	0.035	11	0.467	0.001%
NEC	0.423	30	0.296	1.481%
Parsytec	0.050	11	0.394	0.988%

the relative market shares of all active firms in the industry. SGI accounts for almost half of supercomputer sales in 1995, followed by IBM and Cray. The sum of the shares of HP and Convex was less than 4%, suggesting that no excessive market power concerns arise from a merger between these two firms.

I redefine the observed states under the HP-Convex merger by (i) removing Convex from the set of incumbent firms, leaving its slot open for a potential entrant (i.e., setting $\kappa_{C,1995} = 0$, $\exp(\upsilon_{C,1995}) = 0$ and $N_{C,1995} = 0$, where C denotes the slot of Convex) and (ii) setting the new states for HP as

$$\kappa_{HP,1995}^{Post} = \max\left\{\kappa_{HP,1995}^{\Pr e}, \kappa_{C,1995}^{\Pr e}\right\}$$

$$\exp(\upsilon_{HP,1995}^{Post}) = \exp(\upsilon_{HP,1995}^{\Pr e}) + \exp(\upsilon_{C,1995}^{\Pr e})$$

$$N_{HP,1995}^{Post} = N_{HP,1995}^{\Pr e} + N_{C,1995}^{\Pr e}$$

In the context of Chapter 2, this transformation means that the new firm HP-Convex will have the maximal frontier of the merging firms, the products of both firms will be owned by the single firm[4], and the network of the merged firm will be sum of all consumers whose last purchase was from either HP or Convex. I am left to simulate firm behavior under both the industry state in 1995 and its post-merger version. This is done by computing equilibrium taking these state vectors as initial conditions. The simulation algorithm for this merger closely follows the one of the previous section. For ease of exposition, its details are discussed in Appendix C.

Figure 3.1 presents the simulation results on maximal computing speed in the supercomputer industry under both pre- and post-merger scenarios. As an accuracy check on the simulation, I include the observed maximal computing speed in the industry from 1995 to 2000. The expected maximal computing speed under the merger is fairly close to the observed computing speed in the considered period. The averaged simulated paths under no merger suggest that the maximal computing speed in the supercomputer industry would be slightly lower than the one implied by the HP-Convex merger. This suggests that the HP-Convex was beneficial for promoting frontier innovation. This is

[4]Using the markup-adjusted inclusive value formula of Chapters 1 and 2, it can be easily shown that the exponential of this statistic for a merged firm is the sum of the individual markup-adjusted inclusive value exponentials of each firm.

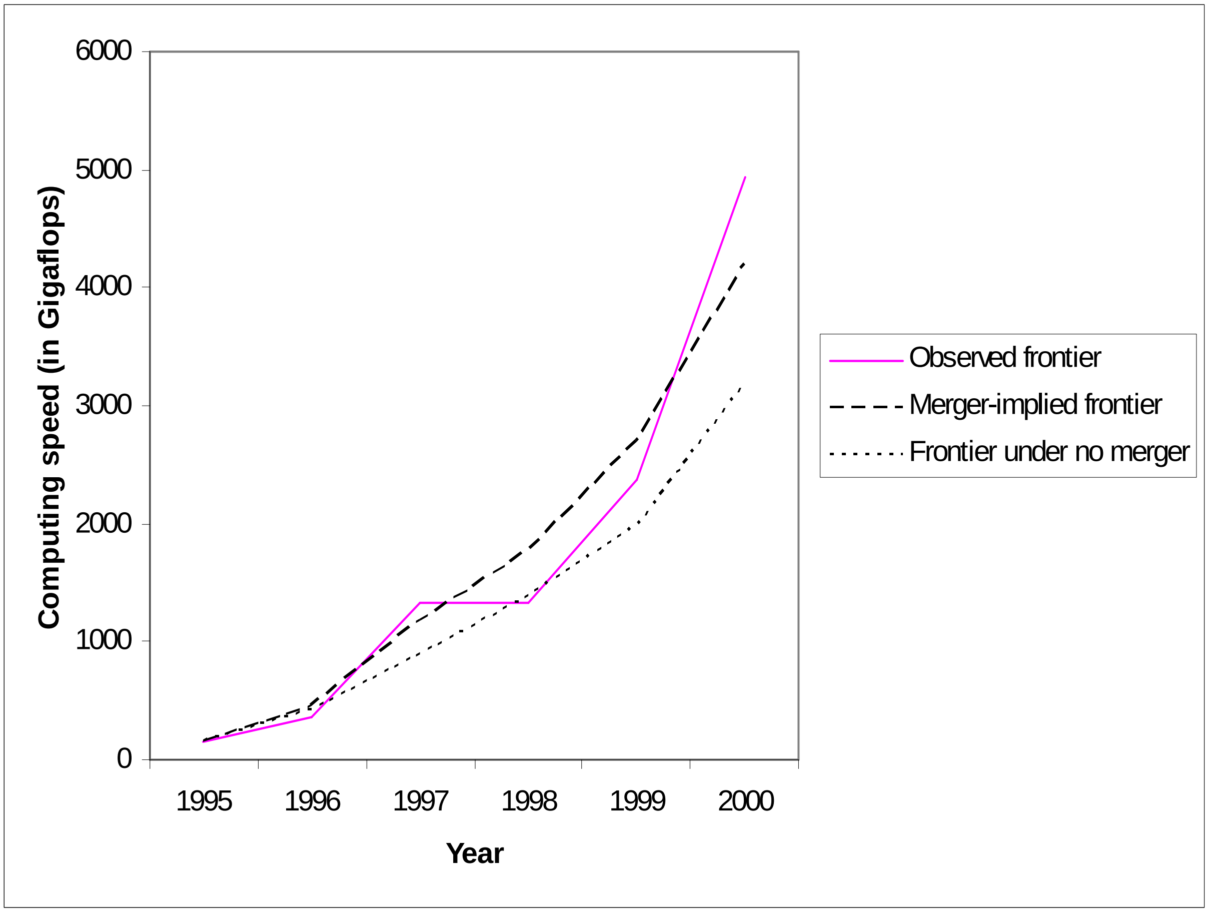

Figure 3.1. Simulated and observed maximal computing speed in the supercomputer industry.

consistent with the "selection effect" theory of Aghion, Harris, Howitt and Vickers (2001) and Aghion et al. (2003), by which the relationship between innovation and competition follows an inverted U-shape. For a reduction on the number of active firms from fifteen to fourteen, the Schumpeterian effect of more innovation via more market power seems to dominate.

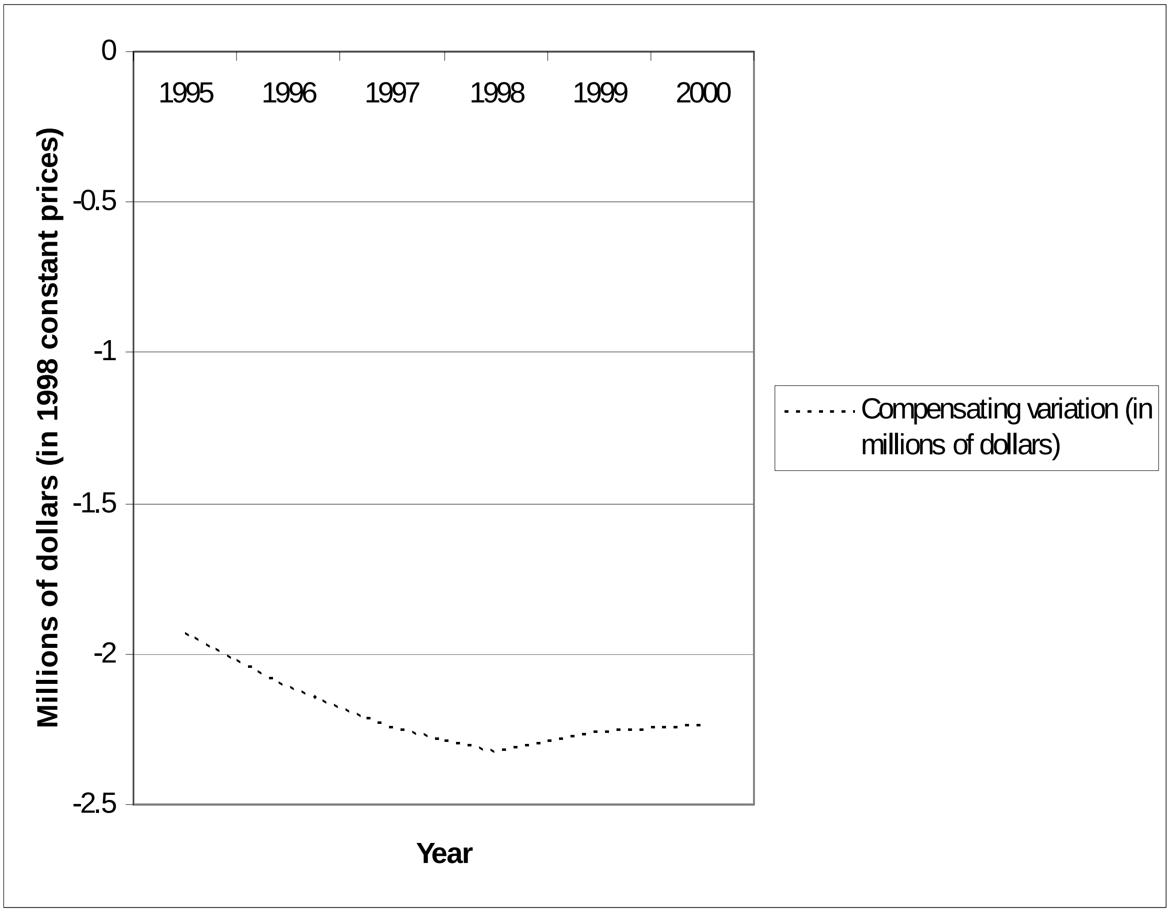

Figure 3.2. Compensating variation evolution over time.

Figure 2.2 described the simulated evolution of the compensating variation implied by the merger. This corresponds to the differential in consumer surplus implied by the merger at every year. The evidence suggests that the consumer welfare reduction implied by the merger was not particularly significant. The loss for the consumers in the supercomputer market over the period 1995-2000 was not bigger than \$2.5M per year. The loss tends to increase shortly after the merger, but it stabilizes (and is even slightly reduced) two years after the merger. There are two possible explanations for this pattern and for the relatively low burden on consumers. First, both HP and Convex were not among the dominant players in this industry in 1995. Consequently, consumer welfare losses should not be significant from increased market power. Second, consumer welfare is increasing in network size to which a consumer belongs. Hence, welfare losses from increased market power may be partially offset by higher utility from increased number of network users.

3.4. Conclusions and extentions

This chapter presented a methodology to simulate long-run merger outcomes. The approach exploits two facts. First, mergers in differentiated product industries may be described by appropriate changes in observed variables. Second, an estimate of firms' ex-ante value functions can be used in equilibrium computation if the merger implies no changes in model parameters. These properties avoid the prohibitive computational burden of simulating merger effects in dynamically competitive industries.

Simulation results are computed under the assumption that both pre- and post-merger equilibrium consist on pure-strategy MPNE. The model presented in Chapter 2 and its estimates are used to illustrate the method. Results suggest that the HP-Convex merger contributed to improved maximal computing speed in the industry at the cost of minor losses in consumer welfare. Due to frequent product introduction and destruction, the analysis excluded the computation of percentage changes in quantities and prices of specific products as a result of the merger. However, it is possible to examine price and quantity changes in industries where there at least a set of products is not likely to be scrapped. The simulated markup policies of the firms in this industry can be used for this purpose.

Other conduct models could be considered to characterize pre- and post-merger equilibria. Even though the analysis above could accommodate alternative conduct models at either stage, the MPNE concept is currently one the richest classes of dynamic equilibria that can be used in applied work. One related shortcoming is the possibility of multiple equilibria. The merger evaluation method proposed in this paper is valid under the assumption that both pre- and post-merger MPNE are unique. Unfortunately, verification ofuniqueness of MPNE is particularly difficult and still an open area of research (see Doraszelski and Satterthwaite (2007) for a discussion). The verification of this assumption for the particular case of the model considered in this Chapter is therefore left for future research.

Another shortcoming on assessing merger effects under dynamic competition is to control for the endogeneity of the merger process. Ignoring the case where the future holds the possibility of mergers can lead to misleading conclusions, as illustrated by Gowrisankaran (1999). The fact that a merger can be described by appropriate changes in the observed state vector indicates that my framework can account for the possibility of future mergers. The extent to which this eliminates any bias from no explicit modeling of merger decisions remains an open question.

The merger analysis developed in this chapter assumed away changes in the parameters on the game. Even though this allowed me to use an estimate of the firm's ex-ante value function instead of computing it, this assumption may be particularly restrictive in other settings. For example, if innovation costs reductions are plausible to occur as a result of a merger, then the value function must be computed in the simulation routine for the new set of parameters. Given the implied computational burden of that exercise when evaluating merger in industries with several firms, it would be important to assess how good of an approximation is the ex-ante value function estimate to the true value function under the new model primitives. The usefulness of this estimate as an approximation is to be assessed in further developments on dynamic merger analysis.

Even though the proposed framework does not model the merger process explicitly, the modeling of this process is important in its own right. The description of the incentives to when and with whom to merge may play a central role on explaining the frequency and wave-like pattern of mergers in some industries. However, multiplicity of equilibria is likely to be a problem in any endogenous merger process. Even if one assumes a sequential merger process in every period, which is by no means a guarantee of uniqueness or even existence of equilibrium (see Gowrisankaran (1999) for a discussion). Nevertheless, the models developed in Chapters 1 and 2 can potentially be adapted to accommodate merger decisions. The extension of these models in this direction is left for future research.

References

[1] Ackerberg, D., C. L. Benkard, S. Berry and A. Pakes (2006): "Econometric Tools for Analyzing Market Outcomes", working paper prepared for J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Volume 6.

[2] Aghion, P., C. Harris, P. Howitt and J. Vickers (2001): "Competition, Imitation and Growth with Step-by-Step Innovation", *The Review of Economic Studies*, Vol. 68, No. 3, pp. 467-492.

[3] Aghion, P., N. Bloom, R.. Blundell, R. Griffith and P. Howitt (2003): "Competition,and Innovation: an Inverted U Relationship", Harvard University, mimeo.

[4] Aguirregabiria, V. and P. Mira (2006): "Identification and Estimation of Dynamic Discrete Games", forthcoming *Econometrica.*

[5] Ai, C. and X. Chen (2003): "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions", *Econometrica*, Vol. 71, No. 6, pp. 1795-1843.

[6] Anderson, S. A. Palma and J. Thisse (1992): "Discrete choice theory of product differentiation", The MIT Press, Cambridge, Mass.

[7] Athey, S. and A. Schmutzler (2001): "Investment and Market Dominance", *The RAND Journal of Economics*, Vol. 32 pp. 1-26.

[8] Bailey, D. and Saini, S. (1996): "NAS Parallel Benchmark (Version 1.0) Results 11-96", NASA Ames Research Center, Report NAS-96-18.

[9] Bajari, P., C. L. Benkard and j. Levin (2006): "Estimating Dynamic Models of Imperfect Comeptition", forthcoming *Econometrica.*

[10] Beggs, A. and P. Klemperer (1992): "Multi-Period Competition with Switching Costs", *Econometrica*, Vol. 60, No. 3, pp. 651-666.

[11] Benitez-Silva, H., G. Hall, G. Hitsch, G. Pauletto and J. Rust (2000): "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems", Yale University, mimeo.

[12] Benkard, C.L. (2004): "A Dynamic Analysis of the Market for Wide-bodied commercial aircraft", *The Review of Economic Studies*, Vol. 71 pp. 581-611.

[13] Benkard, C.L., B. Van Roy and G. Weintraub (2006): "Markov Perfect Industry Dynamics with Many Firms", Stanford University, mimeo.

[14] Beresteanu, A. and P. Ellickson (2006): "The Dynamics of Retail Oligopoly", Duke University, mimeo.

[15] Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation", *The RAND Journal of Economics*, Vol. 25, No. 2, pp. 242-262.

[16] Berry, S., J. Levinsohn and A. Pakes (1995): "Automobile Prices in Market Equilibrium", *Econometrica*, Vol. 63, No. 4, pp. 841-890.

[17] Berry, S., M. Ostrovsky and A. Pakes (2005): "Simple Estimators for the Parameters of Dynamic Games (with Entry/Exit Examples)", NBER Working Paper No. 10506.

[18] Bresnahan and Greenstein (1999): "Technological Competition and the Structure of the Computer Industry", *The Journal of Industrial Economics*, Vol. 47, pp. 1-40.

[19] Bresnahan, T., S. Stern and M. Trajtenberg (1997): "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s", *The RAND Journal of Economics*, Vol. 28, pp. S17-S44.

[20] Bresnahan, T. and M. Trajtenberg (1995): "General Purpose Technologies 'Engines to Growth'?", *The Journal of Econometrics*, Vol. 65, pp. 83-108.

[21] Caplin, A. and B. Nalebuff (1991): "Aggregation and Imperfect Competition: On the Existence of Equilibrium", *Econometrica*, Vol. 59, No. 1, pp. 26-61.

[22] Carranza, J. (2006): "Product Innovation in Durable Goods Markets", University of Wisconsin-Madison, mimeo.

[23] Collard-Wexler, A. (2006): "Plant Turnover and Demand Fluctuations in the Ready-Mix Concrete Industry", Stern School of Business, New York University, mimeo.

[24] Dongarra, J., Luszczek and A. Petitet (2001): "The LINPACK Benchmark: Past, Present and Future", University of Tennessee, mimeo.

[25] Dongarra, J. (2006): "Performance of Various Computers Using Standard Linear Equations Software", University of Tennessee, technical report.

[26] Doraszelski, U. (2003): "An R&D Race with Knowledge Accumulation", *The RAND Journal of Economics*, Vol. 34, No. 1, pp. 20-42.

[27] Doraszelski, U. and M. Satterthwaite (2007): Foundations of Markov-Perfect Industry Dynamics: Existence, Purification and Multiplicity", Northwestern University, technical report.

[28] Doraszelski, U. and A. Pakes (2006): "A Framework for Applied Dynamic Analysis in IO", Harvard University, mimeo.

[29] Dube, J.P. (2004): "Product Differentiation and Mergers in the Carbonated Soft Drink Industry", Graduate School of Business, University of Chicago, mimeo.

[30] Dube, J.P., G. Hitsch and P. Rossi (2006): "Do Switching Costs Make Markets Less Competitive?", Graduate School of Business, University of Chicago, mimeo.

[31] Ericson, R. and A. Pakes (1995): "Markov Perfect Industry Dynamics: A Framework for Empirical Work", *The Review of Economic Studies*, Vol. 62 pp. 53-82.

[32] Evans, D. and R. Schmalensee (2002): "Some Economic Aspects of Antitrust Analysis in Dynamically Competitive Industries", in *Innovation Policy and the Economy*, by A. Jaffe, J. Lerner and S. Stern (eds), MIT Press.

[33] Farrell, J. and P. Klemperer (2006): "Coordination and Lock-In: Competition with Switching Costs and Network Effects", University of California - Berkeley, mimeo.

[34] Gallant, A. R., and D. Nychka (1987): "Semi-Nonparametric Maximum Likelihood Estimation", *Econometrica*, Vol. 52, pp. 363-390.

[35] Gallant, A. R., and G. Tauchen (1989): "Semi-Nonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications", *Econometrica*, 57, pp.1091-1120.

[36] Gallant, A. R., and G. Tauchen (1992): "A Non-Parametric Approach to Nonlinear Time Series Analysis: Estimation and Simulation", In E. Parzen, D. Brillinger, M. Rosenblatt, M. Taqqu, J. Geweke, and P. Caines (eds.), *New Directions in Time Series Analysis*, pp. 71-92, New York: Springer-Verlag.

[37] Gowrisankaran, G. (1999): "A Dynamic Model of Endogenous Horizontal Mergers", *The RAND Journal of Economics*, Vol. 30, pp. 56-83.

[38] Graham, S., M. Snir and C. Patterson (2005): *Getting up to speed: the future of Supercomputing*, the National Academies Press.

[39] Green, J. and S. Scotchmer (1995): "On the Division of Profit in Sequential Innovation", *The RAND Journal of Economics*, Vol. 32, pp. 20-33.

[40] Greenstein, S. (1993): "Did Installed Base Give an Incumbent any (Measurable) Advantages in Federal Computer Procurement?", *The RAND Journal of Economics*, Vol. 24 ,pp. 19-39.

[41] Hendel, I. (1999): "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns", *The Review of Economic Studies*, Vol. 66, pp 423-446.

[42] Hendel, I. and A. Nevo (2006): "Measuring the Implications of Sales and Consumer Inventory Behavior", forthcoming *Econometrica.*

[43] Ho, K. (2005): "Insurer-Provider Networks in the Medical Care Martket", Columbia University, mimeo.

[44] Hotz, J., R. Miller, S. Sanders and J. Smith (1994): "A Simulator Estimator for Dynamic Models of Discrete Choice", *The Review of Economic Studies*, Vol. 61, No. 2, pp 265-289.

[45] Jenkins, M., P. Liu, R. Matzkin and D. McFadden (2004): "The Browser War - Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects", University of California - Berkeley, mimeo.

[46] Judd, K. (1998): *Numerical Methods in Economics*, Cambridge, Mass: MIT Press.

[47] Khanna, T. (1995): "Racing Behavior Technological Evolution in the High-end Computer Industry", *Research Policy*, Vol. 24, pp. 933-958.

[48] Lucas, R. and N. Stokey (1989): "Recursive Methods in Economic Dynamics", Harvard Univeristy Press.

[49] Mazzeo, M. (2002): "Product Choice and Oligopoly Market Structure", *The RAND Journal of Economics*, Vol. 33, No. 2, pp. 1-22.

[50] Miranda, M. and P. Fackler (2002): *Applied Computational Economics and Finance*, Cambridge, Mass: MIT Press.

[51] McCallum, J. (2002): "Price-Performance of Computer Technology", in *The Computer Engineering Handbook*, by V. Oklobdzija (ed), CRC Press.

[52] McFadden, D. (1981): "Econometric Models of Probabilistic Choice", in *Structural Analysis of Discrete Choice with Econometric Applications*, by D. McFadden and C. Manski (eds), MIT Press.

[53] Melnikov, O. (2001): "Demand for Differentiated Durable Products: The Case of the U.S. Computer Printer Market", Cornell University, mimeo.

[54] Murphy, K. and R. Topel (1985): "Estimation and Inference in Two-Step Econometric Models", *Journal of Business and Economic Statistics*, Vol. 3, No. 4, pp. 370-379.

[55] Nevo, A. (2000): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry", *The RAND Journal of Economics*, Vol. 31, pp. 395-421.

[56] Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica*, Vol. 69, No. 2, pp. 307-342.

[57] Nevo, A. and F. Rossi (2007): "An approach for extending dynamic models to settings with multi-product firms", Northwestern University, mimeo.

[58] Pakes, A. and P. McGuire (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model", *The RAND Journal of Economics*, Vol. 25, pp. 555-589.

[59] Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", *Econometrica*, Vol. 55, No. 5, pp. 999-1033.

[60] Ryan, S. (2006): "The Costs of Environmental Regulation in a Concentrated Industry", MIT, mimeo.

[61] Rysman, M. (2004): "Competition Between Networks: A Study of the Market for Yellow Pages", *The Review of Economic Studies*, Vol. 71, pp. 483-512.

[62] Rossi, F. (2006): "An Empirical Model of Dynamic Attribute-Space Competition", Kellogg School of Management, Northwestern University, mimeo.

[63] Scotchmer, S. (1996): "Protecting Early Innovators: Should Second-Generation Products be Patentable?", *The RAND Journal of Economics*, Vol. 27, pp. 322-331.

[64] Rui, X. and M. Miranda (1996): "Solving Dynamic Games via Orthogonal Collocation: An Application to International Commodity Markets", *Annals of Operations Research*, Vol 68, pp. 87-108.

[65] Segal, I. and M. Whinston (2004): "Antitrust in Innovative Industries", Stanford and Northwestern Universities, mimeo.

[66] Seim, K. (2005): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices", forthcoming *The RAND Journal of Economics.*

[67] Song, Minjae (2006): "A Dynamic Analysis of Cooperative Research in the Semiconductor Industry", School of Economics, Georgia Institute of Technology, mimeo.

[68] Sweeting, A. (2007): "The Costs of Product Repositioning: the Case of Format Switching in the Commercial Radio Industry", Northwestern University, mimeo.

[69] Trajtenberg, M. (1989): "The Wellfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners", *The Journal of Political Economy*, Vol. 97, pp. 444-479.

[70] Vedenov, D. and M. Miranda (2001), "Numerical Solution of Dynamic Oligopoly Games with Capital Investment", *Economic Theory*, Vol. 18, pp. 237-261.

??

APPENDIX A

The Data

For the empirical part of the paper, I collect data on the supercomputer industry from 1990 to 2005 using several sources. The first consists on the TOP500 organization, which collects data on supercomputer worldwide installations twice a year on June and November. The data from this source is publicly available at www.top500.org. The available database contains buyer-specific information about the supercomputers being purchased, except for confidential purchases from selected defense and intelligence agencies. This censoring of data is not likely to be significant, since these purchases seem to represent a low percentage of the overall quantity[1].

One limitation of the TOP500 data is that information is restricted to the 500 installed machines with the highest computing speed according to the Linpack Benchmark at the time of each survey. However, this potential source of selection bias can be neglected for two reasons. The first consists on the definition of supercomputer itself. A computer can only be classified as an HPC if its computing speed is close to the best available one. This time-dependent classification allows me to consider the TOP500 surveys representative of the whole supercomputer market. The second reason consists on the obsolescence of a supercomputer after five years from its introduction. Advances in computing speed

[1]The top500 data contains many purchases of classified buyers labeled as "Government Classified", "DoD Classified" and "Defense Classified" along with reported purchases from NSA and FBI. The little available information on confidential purchases indicates that these represent a small percentage of overall supercomputer acquisitions.

become quickly incorporated in other types of computer (mainframes, workstations and eventually laptop and desktop computers), making a supercomputer unlikely to be sold few years after its introduction.

Despite the fact that the TOP500 surveys started on June 1993, I am able to recover information on supercomputer installations since 1990. At the time of that survey, the supercomputer market was yet to experience the remarkable growth in installations during the 1990s. Therefore, the earliest available survey of the TOP500 database contains installations dated as far back as 1984. However, there is only a reasonable number of reported installations from 1990 on[2]. The TOP500 dataset includes the identity of the buyer, the type of buyer (Academic, Industry, Government, Classified and Vendor), the name of the supercomputer model, the country and continent where the HPC is installed, the year of installation, quantities purchased per buyer[3], the area of application for the supercomputer, and the specific attributes of the machine. I used the consumer-specific purchase data to compute firm networks as the number of consumers whose last purchase was from a given firm.

Among those characteristics, one can find two observed quality measures. The first is Rmax, which corresponds to the computing speed of the supercomputer according to the Linpack Benchmark measured in Giga FLOPS. The second measure corresponds to the maximum possible GFLOPS that the HPC can ever process, which is denoted Rpeak. These two computing speed measures are strongly correlated. I found a correlation coefficient of about 0.99 between Rmax and Rpeak. I chose Rmax as the observed quality

[2]Before 1990, the maximal number of installations per year was 21, which is clearly insufficient for obtaining precise estimates.

[3]This information can be inferred directly from the data, since information is provided on an individual-specific basis)

measure for supercomputers for the empirical work. The other specific details of each supercomputer in the data are system family, number of processors, processor speed, processor brand, interconnect brand, architecture and operating system. Processor speed is measured in Mega-Heartz (MHz), and it determines how many instructions per second the processor can execute. Despite the fact that the top500 surveys take place twice a year on June and November, I decided to consider an yearly frequency for the data. I took this procedure since most supercomputers being installed on year t only appeared on the surveys of year $t+1$ and above. Given that the sample only provides information about the installation year, I pooled the information on all the supercomputers installed on a given year from all surveys. This guarantees a truthful description of the moment of installation of the supercomputer[4].

Unfortunately, the TOP500 dataset provides no reference on the price paid for each supercomputer. Nonetheless, it was possible to collect this information from several different sources. The Transaction Processing Performance Council (TPC), a non-profit organization, publishes detailed information submitted by vendors about prices, performance rating and characteristics of several high-end computers at their website www.tpc.org. Several prices of supercomputers introduced since 1994 were recovered from the TPC Benchmark series TPC-A, TPC-B, TPC-C, TPC-H, TPC-R and TPC-W. Another source where supercomputer prices were found was John McCallum's series on CPU performance.

[4]Moreover, the top500 surveys have started in June 1993. They contain information on supercomputers installed in previous years, but provide no information on the semester of installation. Again, there is only reference to the installation year.

McCallum (2002) uses this series to evaluate price and performance information on several computers from 1944 to 2003[5].

Yearly list price data for several supercomputers was also obtained from the HPC database compiled by the International Data Corporation (IDC)[6]. Even though all these sources only partially covered the price information for the TOP500 supercomputers, it was possible to recover the missing information from searching for the supercomputer models at several other sources. The most representative sources were articles published at Government Computer News (GCN) archives (www.gcn.com), press releases from both manufacturers and buyers, technical reports from NASA[7], Roy Longbottom's Computer Claims data from 1980 to 1996[8], the Federal Procurement Data System, the Government Business Opportunities and the Commerce Business Daily websites[9]. All prices were converted into 1998 constant dollars by using the CPI series from the Bureau of Labor Statistics. All price information found in foreign currencies was converted into US dollars by using the average exchange rate for the year in question[10].

Market size at a given year is defined as being the number of consumers potentially interested on purchasing a supercomputer. Since supercomputers are durable goods, the

---

[5]This series is available for download at www.jcmit.com, and the details of its construction can be found in McCallum (2002).

[6]The full database is freely available for download at http://www.hpcuserforum.com/benchmark/

[7]A representative report containing list price information is the technical report by Saini and Bailey (1996), available for download at http://www.nas.nasa.gov/News/Techreports/1996/PDF/nas-96-018.pdf

[8]Available at http://homepage.virgin.net/roy.longbottom/mips.htm#anchorStart

[9]FPDS procurement data is publicly available for consultancy upon account creation at www.fpdc.gov. Data from the Government Opportunities and from Commerce Business Daily can be obtained by searching on archived awards in their websites. I would like to thank the General Services Administration (GSA) staff of the FOIA and Federal Business Opportunities offices for their helpful explanations on data extraction from these sources.

[10]In particular, I used the annual average exchange rate information available at www.triacom.com.

intertemporal substitution effects on demand for durables highlighted by Melnikov (2001) must not be neglected on defining the number of potential consumers. As the discrete choice framework proposed for demand does not account for these effects, I followed the approach of Benkard (2004) by assuming that each buyer in this market optimally reallocates her supercomputer stocks every period, considering both used and new high-performance computers in her choice set. Just like in Benkard's analysis of the market for wide-embodied commercial aircraft, this assumption is acceptable for the case of the supercomputer market. This is because it corresponds to treat supercomputer purchases as rentals, and a considerable number of high-performance computers are installed under rental contracts. Implicit supercomputer rental prices and list prices can be considered proportional, since the maintenance costs for hardware and software for at least three years are included in the latter. Consequently, I defined M_t as the number of new and used supercomputers in use in year t, which is assumed to be the sum of all supercomputers sold between year $t-5$ and t[11].

Supercomputers which have equal characteristics and are produced by the same firm but sold under a different name were considered to be a single model. Moreover, I assume that two observations in adjacent years represent the same model if (a) they have the same name; (b) their Rmax score, number of processors and processor speed does not change by more than ten percent; and (c) the remaining characteristics (e.g., operative system, interconnect bandwidth type) are the same. In addition, I excluded both custom supercomputers and vendor systems (i.e., supercomputers that manufacturers produce for their own usage), although these were considered in the computation of technological

[11]This choice was motivated by the fact that the useful lifetime of a supercomputer is five years.

frontiers and investment rates. Custom systems are not in the same choice set as off-the-shelf models, implying that the logit demand model is only reasonable to describe off-the-shelf purchases. Moreover, the benefits-outside-profits function imposed in the model controls for custom revenues, since the production of custom systems usually implies an expansion of the firm's technological frontier. These procedures yield an unbalanced panel of 2204 model/year observations. Yearly market shares of each supercomputer model are computed by dividing the quantity sold by the market size.

APPENDIX B

Derivations for Chapter 2

B.1. Ex-ante value function approximation

Under the assumption that private shocks are independent across time and players, it suffices to provide an approximant to the integrated value function of each player in the procedures outlined below. The fact that the flow payoff function satisfies symmetry and anonymity suggests that a function satisfying the same properties will be a suitable approximant for the integrated value function. Therefore I define the approximant for this function as

$$\begin{aligned} h_f(\mathbf{s}_t, \theta) \quad = \quad & \sum_{k_1=0}^{K_1} ... \sum_{k_7=0}^{K_7} c_{f,k_1,...,k_7}(\theta) T_{k_1}(N_{0t}) \times T_{k_2}(\kappa_{ft}) \times T_{k_3,}(N_{ft}) \times T_{k_4}(\exp(\nu_{ft})) \\ & \times T_{k_5}(\kappa_{-ft}) \times T_{k_6,}(N_{-ft}) \times T_{k_7}(\exp(\nu_{-ft})) \end{aligned}$$

where, as in the previous section, $T_k(.)$ is an univariate first-kind Chebyshev polynomial of order k, $\kappa_{-ft} \equiv \sum_{d \neq f} \kappa_{dt}$, $\exp(\nu_{-ft}) \equiv \sum_{d \neq f} \exp(\nu_{dt})$, and $N_{-ft} \equiv \sum_{d \neq f} N_{dt}$. Note that the symmetry assumption for the approximant reduced the state space dimension to seven variables, which represents a considerable computational gain[1]. Like in other work where polynomial approximations to the value function are considered (e.g., Doraszelski

[1] Under the assumption that no more than F firms can operate in this market, my model would require a total of $3F + 1$ observed states.

(2003)), I assume that the degree of approximation for each polynomial is the same. That is, $K_i = K$, $\forall i = 1, ..., 7$, $\forall f = 1, ..., F$. Following Miranda and Fackler (2002) It can be compactly written as

$$h_f(\mathbf{s}_t, \theta) = T_f(\mathbf{s}_t)c_f(\theta)$$

where $T_f(\mathbf{s}_t) = [T_7(\exp(\nu_{-ft}) \otimes T_6(N_{-ft}) \otimes ... \otimes T_1(N_{0t})]$ is a $1 \times K^7$ row vector, $c_f(\theta)$ is a $K^7 \times 1$ column vector of coefficients conditional on the dynamic parameters θ, and $T_i(.)$ is a $1 \times K$ row vector where each entry is a first-kind Chebyshev polynomial of the first kind.

Given θ, the vector $c_f(\theta)$ can be estimated after constructing a grid of points on the observed state space. That is, after defining the $1 \times K^7$ row vector function

$$U_f(\mathbf{s}) = \left[T_f(\mathbf{s}) - \beta \int_{\varepsilon} E\left[T_f(\mathbf{s}')|\sigma_f(\mathbf{s}, \varepsilon_f), \sigma_{-f}(\mathbf{s}, \varepsilon_f), \mathbf{s}\right] dF(\varepsilon) \right]$$

and letting $H_f(\mathbf{s}) = \int_{\varepsilon} \pi(\sigma_f(\mathbf{s}, \varepsilon_f), \sigma_{-f}(\mathbf{s}, \varepsilon_f), \mathbf{s}; \theta) dF(\varepsilon)$, the column vector $c_f(\theta)$ can be computed as the unique solution to

$$c_f(\theta) = U_f^{-1} H_f$$

where U_f is a $K^7 \times K^7$ matrix where each row evaluates $U_f(\mathbf{s})$ at a given of grid point $\mathbf{s}$, and H_f is a column vector $K^7 \times 1$ which evaluates $H_f(\mathbf{s})$ at the same grid point $\mathbf{s}$. In principle, one could consider more than K^7 grid points and obtain an estimate for $c_f(\theta)$ by minimizing the distance between H_f and $U_f c_f(\theta)$ (e.g., an OLS solution). However, the fact that U_f can be written as a Kronecker product when K^7 grid points are considered makes this latter case specially attractive from a computational viewpoint[2]. In this paper, I followed the latter approach.

B.2. Auxiliary derivations to first-order conditions

The derivative of $C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft})$ with respect to frontier investment is linear on the shock ε_{1ft}, and is given by

$$\begin{aligned}\frac{\partial C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta)}{\partial I_{ft}} &= Dirac(I_{ft}) C(I_{ft}, \mathbf{s}_t; \theta) + \mathbf{1}\{I_{ft} > 0\}(2c_6 I_{ft} + c_5) \\ &\quad + (\mathbf{1}\{I_{ft} > 0\} + Dirac(I_{ft})(c_5 + \varepsilon_{1ft}) I_{ft})\varepsilon_{1ft}\end{aligned}$$

where $C(I_{ft}, \mathbf{s}_t; \theta) \equiv C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta) - (c_5 + \varepsilon_{1ft}) I_{ft}$ is a term that, by construction, does not depend on ε_{1ft} and $Dirac(I_{ft})$ denotes Dirac's delta function (i.e., the derivative of $\mathbf{1}\{I_{ft} > 0\}$). Using the properties of the Dirac function we have

$$\frac{\partial C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta)}{\partial I_{ft}} = \mathbf{1}\{I_{ft} > 0\}(2c_6 I_{ft} + c_5) + \mathbf{1}\{I_{ft} > 0\}\varepsilon_{1ft}$$

The derivative of $C_A(A_{ft}, \mathbf{s}_t, \varepsilon_{2ft}; \theta)$ with respect to A_{ft} is linear in ε_{ft}. Hence, the first-order condition on quality investment simplifies to

[2] As described in Miranda and Fackler (2002, pp130-132), U_f^{-1} can be computed after inverting seven matrices $K \times K$ rather than the more computationally intensive inversion of a $K^7 \times K^7$ matrix.

$$\int_{\varepsilon_{-f}} \left(-C_1(A_{ft}, \mathbf{s}_t; \theta) + \beta E_t \left[\sum_{k=1}^{\Delta(S)} \frac{\partial V_f}{\partial g_k} \frac{\partial g_k}{\partial A_{ft}} \right] \right) dF(\varepsilon_{-f}) - \mu_{ft}^A \mathbf{1}\{A_{ft} = -1\} = 0$$

where

$$\begin{aligned} C_1(A_{ft}, \mathbf{s}_t; \theta) &= Dirac(A_{ft})(c_0^A A_{ft} + c_1^A A_{ft}^2) + \mathbf{1}\{A_{ft} > 0\}(c_0^A + 2c_1^A A_{ft}) \\ &\quad + (1 - Dirac(A_{ft}))(c_2^A A_{ft} + c_3^A A_{ft}^2) + \mathbf{1}\{A_{ft} < 0\}(c_2^A + 2c_3^A A_{ft}) \end{aligned}$$

Again using the fact that $A_{ft} \times Dirac(A_{ft}) = 0$, $\forall A_{ft}$, $C_1(A_{ft}, \mathbf{s}_t; \theta)$ simplifies to

$$C_1(A_{ft}, \mathbf{s}_t; \theta) = \mathbf{1}\{A_{ft} > 0\}(c_0^A + 2c_1^A A_{ft}) + \mathbf{1}\{A_{ft} < 0\}(c_2^A + 2c_3^A A_{ft})$$

Fixed costs of frontier innovations are recovered by using the fact that the firm will only extend the frontier if the value of doing so exceeds the value of not doing technological innovation. That is, frontier investment takes place only if the following *positive investment condition* holds

$$-C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta) + \beta \int\limits_{\varepsilon_{-f}} E\left[V_f(\mathbf{s}_{t+1}; \theta) | I_{ft} > 0, \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t\right] dF(\varepsilon_{-f})$$

$$> \beta \int\limits_{\varepsilon_{-f}} E\left[V_f(\mathbf{s}_{t+1}; \theta) | I_{ft} = 0, \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t\right] dF(\varepsilon_{-f})$$

Using the fact that $C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta)$ is linear in ε_{1ft}, this condition can be written as

$$\varepsilon_{1ft} < \frac{\beta \int\limits_{\varepsilon_{-f}} E\left[V_f(\mathbf{s}_{t+1}; \theta) | I_{ft} > 0, \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t\right] dF(\varepsilon_{-f})}{\mathbf{1}\{I_{ft} > 0\} I_{ft}}$$
$$- \frac{\beta \int\limits_{\varepsilon_{-f}} E\left[V_f(\mathbf{s}_{t+1}; \theta) | I_{ft} = 0, \sigma_f, \sigma_{-f}(\mathbf{s}_t, \varepsilon_{-f}), \mathbf{s}_t\right] dF(\varepsilon_{-f})}{\mathbf{1}\{I_{ft} > 0\} I_{ft}} + \frac{C(I_{ft}, \mathbf{s}_t; \theta)}{\mathbf{1}\{I_{ft} > 0\} I_{ft}}$$

where $C(I_{ft}, \mathbf{s}_t; \theta)$ is the investment cost net of its stochastic component, i.e.,

$$C(I_{ft}, \mathbf{s}_t; \theta) \equiv C(I_{ft}, \mathbf{s}_t, \varepsilon_{1ft}; \theta) - c_2 I_{ft} \mathbf{1}\{I_{ft} > 0\}$$

APPENDIX C

Simulation methods for Chapter 3

C.1. HP-Convex merger simulation

Given estimates of the model parameters, state transitions and integrated value functions recovered in the previous Chapter, the simulation algorithm starting at either of those two initial states can be outlined as follows:

- Given the observed state vector in 1995, compute the MPNE by solving the system of first-order conditions on markups, quality investments and frontier extension rates of all firms. Following Ryan (2006), I integrate the system with respect to private information. I proceed analogously for the state vector implied by the merger;
- Compute consumer welfare under both pre- and post-merger scenarios using the metric described in Chapter 2, i.e.

$$CW_t = \frac{1}{\alpha}\sum_{k=0}^{F} N_{kt} \ln\left(\sum_{d=1}^{F} \exp(-\alpha Mkp_{dt} + \tau_1 \kappa_{dt} + \nu_{dt} + \psi N_{dt} + \Lambda_d \mathbf{1}\{k=d\})\right)$$

- Compute the compensating variation implied by the merger, i.e.,

$$CV_t = CW_t^{Post} - CW_t^{\Pr e}$$

- For both pre- and post-merger cases, generate the maximal computing speed available in the industry for the following period by computing maximal firm frontier times $(1 + I_{ft})$ across all firms which will remain in the industry;
- Generate an observed state vector for 1996 under pre- and post-merger scenarios by sampling a draw from the transition probabilities evaluated at the system solution and the observed state in 1995;
- Repeat the first and second steps considering the random draw of the previous step as the observed state.
- Repeat the last step $\tilde{T}$ times, so that one has a simulated path up to year $t = 1995 + \tilde{T}$. In my application, I set $\tilde{T} = 5$;
- Repeat last four steps R times, and compute the expected markup and welfare paths by averaging across the R simulations. In my application I have set $R = 5000$.