Market Structure and Innovation: A Dynamic Analysis of the Global Automobile Industry*

Aamir Rafique Hashmi†
and
Johannes Van Biesebroeck‡

April 22, 2009

Abstract

We study the relationship between market structure and innovation in the global automobile industry for the 1982-2004 period. We use the dynamic industry framework of Ericson and Pakes [1995] and estimate the parameters of the model using a two-step procedure proposed by Bajari et al. [2007]. Since the industry has seen a lot of consolidation since 1982, mergers are an important ingredient of our model. We show that our estimated model predicts the data well. We study the implications of the estimated model for the relationship between the market structue and innovation and also for welfare. Our findings are the following: (1) There is no clear relationship between competition and innovation at the firm level; (2) The effect of market structure on innovation at the industry level depends on the initial state of the industry. If the industry is not very concentrated, as it was in 1982, some consolidation may increase the innovative activity. However, if the industry is already concentrated, as in 2004, further consolidation may reduce the incentives to innovate; (3) Innovation is costly and despite being an optimal response to market conditions, it hurts producers more than it benefits consumers.

Key words: Competition and Innovation; Automobile Industry; Dynamic Games
JEL Classification Codes: C73; L13; L62; O31

*We thank participants at the conferences of the CEA (Montreal and Halifax), EEA (Vienna), IIOC (Savannah), SED (Prague), and at the seminars at the University of Toronto, National University of Singapore, Victoria University (Wellington, New Zealand) and London School of Economics for their comments. We also thanks Bart Van Looy and Xiaoyan Song for help with patent data. Financial support from AUTO21, CFI, and SSHRC is greatly appreciated.

†Department of Economics, National University of Singapore. E-mail: aamir@nus.edu.sg

‡Department of Economics, University of Toronto and NBER. E-mail: johannes.vanbiesebroeck@utoronto.ca

1 Introduction

industry. Innovation is also likely to advance a firm's competitive position through higher product quality and reliability, the introduction of desirable features in its products, or lower production cost. A large number of analysts makes a living measuring these benefits for consumers and investors alike.[3]

Unlike most of the studies that use reduced form regressions to study the relationship between market structure and innovation, we do so in a strategic and dynamic environment. There are at least two reasons for this. First, innovation is inherently a dynamic activity. Firms make R&D investments today expecting uncertain future rewards through better products or more efficient production. The magnitude of these benefits are intricately linked to the future market structure and the level of innovation of competitors.[4] Merging or forming strategic alliances with rivals is also best analyzed in a dynamic framework. Mergers interact with innovation through their influence on market structure as well as through the consolidation of the knowledge stock in the industry. The global automobile industry seems an ideal place to study these two forces and their interaction.

Second, we now have computationally tractable methods to estimate models of dynamic competition. In this study, we employ a recently developed technique in the estimation of dynamic games that does not require one to solve for the equilibrium of the game, see Bajari et al. [2007]. Our study is one of the first to put these new techniques to a practical test, demonstrate their usefulness and highlight some practical difficulties associated with their application.[5]

The principal objective of our paper is to study the response of firms in terms of their innovative activity to changes in the level of competition — a combination of market structure and the level of innovation of all market participants. We do so in a dynamic environment in which firms produce differentiated products that differ in quality. Firms invest in R&D to improve the quality of their products and to lower the cost of production. The market share of a firm depends on the relative quality of its product and the price, which is set strategically in each period. Investments in R&D increase the technological knowledge of the firm but the exact outcome of R&D is uncertain. On average, higher knowledge translates into a higher quality product and lower marginal cost. The investment in R&D is modeled as a strategic decision: a firm takes the actions of its rivals and their possible future states into account before making its R&D decision. In addition, firms incorporate potential future mergers into their expectations.

[3] To name but a few, *J.D. Power* and *Consumer Reports* measure defect rates in vehicles; numerous consumer magazines and internet sites compare the performance and discuss the features of vehicles; and *Harbour Consulting* and *KPMG* continuously compare productivity in the industry.

[4] Aghion and Griffith [2005] provide an overview of the issues involved and review some popular modeling approaches.

[5] Among a growing list of (working) papers using various approaches, we can point the interested reader to the following applications: Ryan [2006], studying the effect of environmental regulation on the cement industry; Collard-Wexler [2005], studying the effect of demand fluctuation in the ready-mix concrete industry; Aguirregabiria and Ho [2006], studying the airline industry; and Sweeting [2006] estimating switching costs for radio station formats.

The rest of the paper is organized as follows. The supply and demand side of the model as well as the Markov perfect equilibrium concept we rely upon are introduced in Section 2. Section 3 introduces the data. The two-step estimation methodology and the coefficient estimates are discussed in Section 4. In Section 5 we report our empirical findings. We examine the robustness of our estimates in Section 6 and Section 7 concludes.

2 The Model

Our modeling strategy follows Ericson and Pakes [1995]. There are n firms, each producing a differentiated vehicle. Firms differ in their technological knowledge, which is observed by all market participants as well as by the econometrician, and in a firm-specific quality index, known to the market participants but not observed by the econometrician. We denote the technological knowledge of firm j ($j = 1, 2, \ldots, n$) by $\omega_j \in \mathbb{R}^+$ and the quality index by $\xi_j \in \mathbb{R}$. For the industry as a whole, the vectors containing ω and ξ are denoted by $\mathbf{s}_{\boldsymbol{\omega}}$ and $\mathbf{s}_{\boldsymbol{\xi}}$. For later use, we define the vectors of ω and ξ excluding the firm j as $\mathbf{s}_{\boldsymbol{\omega}}{}^{-j}$ and $\mathbf{s}_{\boldsymbol{\xi}}{}^{-j}$. We also define $\mathbf{s} = \{\mathbf{s}_{\boldsymbol{\omega}}, \mathbf{s}_{\boldsymbol{\xi}}, m\}$ and $\mathbf{s}^{-j} = \{\mathbf{s}_{\boldsymbol{\omega}}{}^{-j}, \mathbf{s}_{\boldsymbol{\xi}}{}^{-j}\}$.[6]

Time is discrete. At the beginning of each period, firms observe $\mathbf{s}$ and make their pricing and investment decisions (see below). Although the pricing decisions are static, the investment decisions are dynamic and depend on the current as well as expected future states of the industry.

Firms invest to increase their technological knowledge. A higher level of knowledge may boost demand, for example, by improving vehicle quality or by introducing more innovative product features. It may also reduce marginal cost by improving the efficiency of production. Hence, the model features both product and process innovation. The effect of R&D investment on knowledge is the sum of a deterministic and a random component, capturing that innovation is a stochastic process. The technological knowledge depreciates at an exogenous rate.

There is no entry or exit in our model, reflective of the evolution of the industry over the last twenty five years or so. However, a firm may 'exit' by merging with another firm. Mergers are an important component of our model as they lead to discrete changes in market structure and force the firms to readjust their prices and investments to take new industry structure into account. In our model, mergers take place for exogenous reasons, but firms take them into account when forming expectations over future valuations. In the remainder of this section we describe the demand and supply sides in some detail and then define the Markov perfect equilibrium of the model.

[6] is the number of consumers in the market.

2.1 The Demand Side

Following Berry et al. [1995] and many others studying the automobile industry, we use a discrete choice model of individual consumer behavior to model the demand side. There are n firms in the industry producing differentiated vehicles. Vehicles differ in quality, which has two components. The first component, observable to the market participants as well as the econometrician, is positively related to the firm's technological knowledge, $g(\omega)$, with $\partial g(\omega)/\partial \omega \geq 0$. The second component ($\xi$) is unobservable to the econometrician, but firms and consumers know it and use it in their pricing and purchase decisions.

There are m consumers in the market in each period and each of them buys one vehicle. m grows at a constant rate that is exogenously given. The utility of a consumer depends on the quality of a vehicle, its price and the consumer's idiosyncratic preferences. The utility consumer i gets from buying vehicle j is

$$u_{ij} = \theta_\omega g(\omega_j) + \theta_p \log(p_j) + \xi_j + \nu_{ij}, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, n. \tag{1}$$

We assume that the observable vehicle quality equals $\log(\omega_j + 1)$ and ξ_j is the unobservable quality (to the econometrician). p_j is the price of the vehicle produced by firm j and is adjusted for vehicle characteristics such as size and performance characteristics. θ_ω and θ_p are preference parameters. θ_ω shows how quality conscious the consumers are and θ_p is a measure of their price elasticity. ν_{ij} is the idiosyncratic utility that consumer i gets from good j. Assuming that ν is i.i.d. extreme value distributed, gives the following expected market share for firm j:

$$\sigma_j(\omega_j, \xi_j, p_j, \mathbf{s}^{-j}, \mathbf{p}^{-j}) = \frac{\exp(\theta_\omega \log(\omega_j + 1) + \theta_p \log(p_j) + \xi_j)}{\sum_{k=1}^{n} \exp(\theta_\omega \log(\omega_k + 1) + \theta_p \log(p_k) + \xi_k)}, \tag{2}$$

where p_j is the price charged by firm j and $\mathbf{p}^{-j}$ is the price vector of all the other firms (excluding firm j) in the industry. The expected demand for vehicle j is simply $m\sigma_j(\cdot)$. Each firm's demand depends on the full price vector in the industry, directly through the denominator of (2) and indirectly through its own price because its equilibrium price is a function of its rivals' prices.

2.2 The Supply Side

We begin our explanation of the supply side with the period profit function. We assume that R&D investments only generate useful knowledge with a one period lag and that prices can be adjusted flexibly period by period. Hence, at the beginning of each period, after observing their individual and industry states, the firms engage in a differentiated products Bertrand-Nash game. Each firm chooses its own price to maximize profits, taking the prices of its rivals ($\mathbf{p}^{-j}$) and industry state ($\mathbf{s}$) as given.

The profit maximization problem of an individual firm j is

$$\pi_j(\omega_j, \xi_j, \mathbf{s}^{-j}, \mathbf{p}^{-j}, m) = \max_{p_j}\{(p_j - mc_j(\omega_j))m\sigma_j(\cdot) - fc_j\}, \tag{3}$$

where mc_j is the marginal cost incurred by firm j to produce a vehicle. The marginal cost is a function of the firm's knowledge, capturing cost reducing process innovations. fc_j is the fixed cost of operations faced by firm j. For now, the fixed cost does not play any role.

The first order condition for firm j, after some simplification, is

$$(p_j - mc_j(\omega_j))[1 - \sigma_j(\cdot)]\theta_p + p_j = 0. \tag{4}$$

Since there are n firms, we have to solve n such first order conditions simultaneously to obtain the equilibrium price vector $\mathbf{p}^*$.[7] Equilibrium profits are given by

$$\pi_j(\omega_j, \xi_j, \mathbf{s}^{-j}, m) = (p_j^* - mc_j(\omega_j))m\sigma_j(\omega_j, \xi_j, p_j^*, \mathbf{s}^{-j}, \mathbf{p}^{*-j}). \tag{5}$$

Once we know the functional form and parameter values of the demand and cost functions, we can evaluate (5) for any industry state $\mathbf{s}$.

The investment in R&D is a strategic and dynamic decision. Each period, firms choose their R&D investment based on the expected value of future profit streams. The problem is recursive and can be described by the following Bellman equation

$$V_j(\omega_j, \xi_j, \mathbf{s}^{-j}, m) = \max_{x_j \in \ _+}\{\pi_j(\omega_j, \xi_j, \mathbf{s}^{-j}, m) - cx_j + \beta EV_j(\omega_j', \xi_j', \mathbf{s}'^{-j}, m')\}, \tag{6}$$

where β is the discount factor, c is the R&D cost per unit of new technological knowledge and x_j is the addition to new knowledge. A prime on a variable denotes its next period value. c is the only dynamic parameter in our model that we estimate. The solution to (6) is a policy function $x_j(\omega_j, \xi_j, \mathbf{s}^{-j}, m)$.

The knowledge stock of the firm evolves as follows:

$$\omega_j' = (1 - \delta)\omega_j + x(\omega_j, \xi_j, \mathbf{s}^{-j}, m) + \epsilon_j^{\omega}, \tag{7}$$

where δ is the depreciation rate of technological knowledge and is exogenously given. ϵ^{ω} is a random shock that represents the uncertainty involved in doing R&D. A firm that spends cx on R&D will, on the average, increase its knowledge stock by x. We assume that ϵ^{ω} follows a well defined and known distribution.[8]

[7] The existence and uniqueness of equilibrium in this context has been proved by Caplin and Nalebuff [1991].

[8] This law of motion for the state variable is less general than the one in the main analysis in Ericson and Pakes [1995], as depreciation is deterministic, but more general than the example they analyze in detail, as we allow to take on a continuum of values.

We assume an exogenous AR(1) process for the evolution of the unobserved component of quality (ξ).

$$\xi'_j - \bar{\xi}_j = \rho(\xi_j - \bar{\xi}_j) + \epsilon^{\xi}_j, \tag{8}$$

where $\bar{\xi}_j$ is the average value of ξ_j over the sample period.

We assume that m

mergers as a declining function of n (see the calibration of the merger probability in the Appendix) and we impose a lower threshold on the number of active firms on competition policy grounds. This is consistent with Klepper [2002a] and Klepper [2002b], who argues that the U.S. automobile industry has settled into a stable oligopoly. We assume that the same will happen to the global automobile industry. An alternative solution would be to introduce entry, but it would require more radical changes in the empirical strategy. In practice, entry has not played an important role in the evolution of the global market structure, although this could change with the development of the Chinese and Indian automotive industries.

Even with this simple merger technology we need to specify the state variables for the newly created firm and how each firm incorporates possible mergers in the evaluation of its future value. We assume that when two firms merge, the knowledge stock of the merged firm is the sum of the individual knowledge stocks of the merging firms. Another possibility would be to allow for complementarities in the knowledge stock or, at the other extreme, assume some overlap in knowledge and discount the sum. All three assumptions are equally arbitrary, but we feel that simply adding the two knowledge stocks is closest in spirit to the state transition function for knowledge used throughout. We further assume that the unobserved quality of the vehicle produced by the merged firm is the average of the unobserved qualities of the vehicles produced by the original firms.

To see how we incorporate mergers in our model, assume that the industry is a duopoly with firms A and B. The firms' states are (ω_A, ξ_A) and (ω_B, ξ_B).[10] There is an exogenous probability p_m each period that they will merge. Firm A will incorporate this information in the calculation of its value functions as follow:

$$\begin{aligned} V_A(\omega_A, \xi_A, \omega_B, \xi_B) \quad = \quad & \max_{x_A \in \ _+} \Big\{ \pi_A(\omega_A, \xi_A, \omega_B, \xi_B) - c x_A(\omega_A, \xi_A, \omega_B, \xi_B) \\ & + \beta \Big[p_m \zeta_A(\omega'_A, \xi'_A, \omega'_B, \xi'_B) EV_A(\omega'_A + \omega'_B, (\xi'_A + \xi'_B)/2) \\ & + (1 - p_m) EV_A(\omega'_A, \xi'_A, \omega'_B, \xi'_B) \Big] \Big\}, \end{aligned} \tag{9}$$

where $\zeta_A(\omega'_A, \xi'_A, \omega'_B, \xi'_B)$ is the share of firm A in the total value of the merged firm. We assume that it is simply the ratio of the stand alone value of A to the sum of the values of both firms in absence of the merger:

$$\zeta_A(\omega'_A, \xi'_A, \omega'_B, \xi'_B) = \frac{EV(\omega'_A, \xi'_A, \omega'_B, \xi'_B)}{EV(\omega'_A, \xi'_A, \omega'_B, \xi'_B) + EV(\omega'_B, \xi'_B, \omega'_A, \xi'_A)}.$$

The same idea extends to an industry with more firms, but the computations become more involved—see the Appendix for details.

To summarize, in each period, the sequence of events is the following:

[10] The market size ( ), which is the only industry-level state variable in our model, is implicit and we suppress it for simplicity of notation.

1. Firms observe individual and industry states.

2. Pricing and investment decisions are made.

3. Profits and investment outcomes are realized.

4. Individual and industry states are updated before the mergers.

5. Mergers are drawn randomly (see Appendix for details).

6. The state variables of merged firms are adjusted and the industry states are updated accordingly.

2.4 Markov Perfect Equilibrium

The Markov Perfect Equilibrium of the model consists of $V(\omega, \xi, \mathbf{s})$, $\pi(\omega, \xi, \mathbf{s})$, $x(\omega, \xi, \mathbf{s})$ and $Q(\mathbf{s}', \mathbf{s})$ such that:

1. $V(\omega, \xi, \mathbf{s})$ satisfies (17) and $x(\omega, \xi, \mathbf{s})$ is the optimal policy function;

2. $\pi(\omega, \xi, \mathbf{s})$ maximizes profits conditional on the state of the industry;

3. $Q(\mathbf{s}', \mathbf{s})$ is the transition matrix that gives the probability of state $\mathbf{s}'$ given the current state $\mathbf{s}$.

Estimating the parameters in the period profit function—demand and supply parameters—is fairly standard. The one dynamic parameter in the model that we estimate, c the cost of R&D, poses a greater challenge. There are at least two ways to proceed. The first is to compute the Markov Perfect Equilibrium (MPE), as defined above, from a starting value of c and use maximum likelihood estimation, as in Rust [1987] or Holmes and Schmitz [1995], to fit the observed investment decisions to the predictions from the model's Euler equations. Having solved for the equilibrium one can simulate the model and study the dynamics of interest. Benkard [2004] uses a similar approach in his study of the market for wide-bodied commercial aircraft. The main disadvantage of this approach is that the numerical solution for the MPE is computationally very intensive. This remains a problem despite some recent innovations by Pakes and McGuire [2001], Doraszelski and Judd [2004] and Weintraub et al. [2005].

The second is a two-step approach proposed by Bajari et al. [2007]. Their method allows for the estimation of the policy and value functions and for the recovery of structural parameters of the model without having to compute the MPE. They get around the problem of computing the

equilibrium by assuming that the data we observe represent an MPE.[11] This assumption is not completely innocuous. For example, in the auto industry firms often undergo structural changes with adjustments spread over many years. During this transition period, firms do not behave as they would in equilibrium. A prime example would be the three-year recovery plan Renault initiated at Nissan, when it took control of the troubled Japanese automaker. Similarly GM and Ford, having lost a big chunk of their market share in recent years, are undergoing massive restructuring. Their decisions during this restructuring phase are likely to be different from their decisions in a stable equilibrium. Nevertheless, the assumption that firms always play their equilibrium strategy sounds reasonable and, more importantly, does away with the need to compute the MPE.[12]

Given this assumption, the first step proposed by Bajari et al. [2007] is to estimate the state transition probabilities and the equilibrium policy functions directly from the observed information on investments and the evolution of the two state variables. Together with estimates of the period profit function, these can be used to obtain the value functions by forward-simulation. In the second step, the value function estimates from the first step are combined with equilibrium conditions of the model to estimate the structural parameter(s). In Section 4 we elaborate on these steps in some detail and explain how we apply them to our model. First, we describe the data set.

3 Data

We choose our sample period to be 1982–2004. This period covers most of the consolidation that has taken place in the industry over the last few decades. We limit the sample to the largest thirteen firms (in terms of unit sales) that were active in 2004. The industry has seen significant consolidation over these twenty three years and working back to 1982, these thirteen groups emerged from twenty three initially independent firms. Throughout, we do not distinguish between full and partial ownership ties, e.g. Nissan and Renault are treated as a single firm after they initiated an alliance in 1998, even though Renault never obtained majority control. Figure 1 lists all firms in the initial and final year of the sample. In 2004, our sample accounted for slightly more than 95% of global automobile sales. The remaining 5% of sales were shared by a large number of small firms. Since patenting activity of these small firms is negligible and other information on them is spotty,

[11] Alternative approaches that avoid computing the MPE at each iteration include Aguirregabiria and Mira [2007] and Pakes et al. [2005]. We follow Bajari et al. [2007] because their method allows for continuous choice variable, as we have in our model.

[12] Another implication of this assumption is that one has to rule out the possibility of multiple equilibria or simply assume that multiple equilibria may exist but the firms only play one and the same equilibrium in all periods. For a review of the problem of multiple equilibria in these models and its possible solutions see Section 6 in Doraszelski and Pakes [2006].

we ignore these fringe firms.[13]

[Figure 1 approximately here]

In order to estimate the parameters we need data on the following four variables: gross additions to a firm's knowledge (x); knowledge stock (ω); market share; and prices. Our measure of gross addition to a firm's knowledge is the number of patents applied for by a firm in a calendar year.[14] We use PATSTAT (http://wiki.epfl.ch/patstat) for patent data. We combine data on the number of new patents applied for by each firm to the US and European patent offices. Since different subsidiaries of the same firm might file for patents, we searched the database using several variations of the names of each firm and manually scrolled through the results to make sure that all appropriate patents were included. We combine the US and European patent data in the following way. For each firm-year, we combine the observations as: $x_{jt} = \max(x_{jt}^{US} + \lambda x_{jt}^{EU})$, where x_{jt}^{US} is the number of new patent applications by firm j in year t to the US patent office and x_{jt}^{EU} to the European patent office. λ is the weight given to a European patent relative to a US patent. It is computed by taking the ratio of US to European patents by four big firms (Daimler, Ford, Honda and Toyota) that have significant presence in both regions. We compute this weight to be 2.2.

Our measure of the knowledge stock of a firm is its 'patent stock'. Using the number of patents applied for by each firm for the period 1982–2004, we construct the patent stock using the perpetual inventory method: $\omega_{t+1} = (1-\delta)\omega_t + \tilde{x}_t$.[15] The initial patent stock is given by $\omega_0 = \frac{x_0}{g+\delta}$, where x_0 is the number of new patents applied for, δ is the depreciation rate of the patent stock and g is the growth rate of the new patent applications. We estimate g using the data for the first five years of our sample.

Our empirical counterpart to sales is the number of vehicles sold worldwide by each firm and its affiliates. This information is obtained from Ward's Info Bank, the Ward's Automotive Yearbooks, and is supplemented by information from the online data center of Automotive News for the most recent years. Market share of a firm is computed as the ratio of the number of vehicles sold by a firm to the total number of vehicles sold in the global market.

[NOTE TO JO: PLEASE UPDATE THE FOLLOWING PARAGRAPH] To construct prices for these 'composite' models, we estimate a hedonic price regression for all available models in the market. The log of the list price is the dependent variable and a host of vehicle characteristics as

[13] The only sizeable firm in 1980 not included in our sample is Lada in the USSR. British Leyland in the U.K. was part of the Rover Group and AMC in the U.S became 46% owned by Renault shortly after 1979.

[14] Patents are a widely used as a measure of innovation output. In his survey on use of patents as a measure of technological progress, Griliches [1990] concludes: "In spite of all the difficulties, patents statistics remain a unique resource for the analysis of the process of technical change."[p. 1701]

[15] The difference between this equation and the law of motion for the knowledge stock, equation (7), is that the firm plans to obtain $_t$ patents, but the randomness in the innovation process yields the observed $\tilde{}_t = {}_t +$ new patents, with $(\tilde{}_t) = {}_t$.

explanatory variables—see Goldberg and Verboven [2001] for an example.[16] We include a full set of firm-year interaction dummies and these coefficients capture the relative price for each firm in each year. The log price is relative to the base firm—GM—exactly as needed to estimate the demand equation in (10) below. For now, we estimate the hedonic regression only for the U.S. passenger vehicle market, updating the data set in Petrin [2002] to 2004. Figure 2 illustrates the evolution of these prices for a number of firms.

[Figure 2 approximately here]

4 Estimation Methodology and Results

4.1 Step 1

We now describe the estimation methodology, which closely follows Bajari et al. [2007]. In the first step, we estimate the demand and cost of production parameters, the transition probabilities and policy functions and use them to evaluate the value functions using forward simulations. We present the results immediately following the relevant piece of the model.

4.1.1 Estimation of Demand Parameters

The demand side in our model is static and we do not need the full model to estimate the demand parameters.[17] Following Berry [1994] we can write the log of the market share of firm j relative to a base firm 0 as

$$\log[\sigma_j(\cdot)/\sigma_0(\cdot)] = \theta_\omega \log[(\omega_j + 1)/(\omega_0 + 1)] + \theta_p \log[p_j/p_0] + [\xi_j - \xi_0]. \tag{10}$$

Using the observed market shares and with data on ω's and prices, we can estimate the above equation by OLS to get the estimates for θ_ω and θ_p. However, as producers use information about the unobserved vehicle quality (ξ) in their pricing decisions, prices will be correlated with the error term and OLS estimates will be inconsistent. In particular, we expect the price coefficient to be upwardly biased. We use an IV estimator and follow the instrumenting strategy of Berry et al. [1995]: the sum of observable characteristics of rival products is an appropriate instrument for own price. In our case, this boils down to the sum of rivals' knowledge. The residuals from (10) are our empirical estimates of ξ relative to GM.

[16] Bajari and Benkard [2005] discuss the performance of hedonic pricing models when some product characteristics are unobservable.

[17] We estimate demand, cost, and policy functions pooling data across all years; time subscripts are omitted. Throughout, we use GM as the base firm.

The demand parameters using three different estimation methods—least squares (OLS), instrumental variables (IV) without and with time fixed-effects—are in Table 2. The impact of knowledge, θ_ω, is estimated similarly under all three specifications: it has a positive impact on sales, as expected, and is estimated very precisely.

The price coefficient is estimated negatively, even with OLS, which ignores the correlation between price and unobserved product characteristics. Often, OLS estimates of product level discrete choice demand systems produce a price coefficient that is positive or close to zero. A firm's patent stock seems to control for a lot of the usually unobserved quality variation that can lead to an upward bias in the price coefficient. Note that we do not observe any other characteristic; the price variable controls for any observable differences in the products offered by each firm.

NOTE TO JO: PLEASE UPDATE THIS PARAGRAPH. Instrumenting for price with the knowledge stock of competitors leads to an estimated demand curve that is more elastic; again in line with expectations (see for example Berry et al. [1995], Table III). Including time dummies increases the elasticity further. Using the estimates in the third column, the price elasticity varies between -7.2 and -1.96. Without enforcing the first order conditions for optimal price setting, all firms are estimated to price on the elastic portion of demand, consistent with oligopoly theory. The price-marginal cost markups implied by these estimates are below those obtained in other studies that estimate demand systems for car models (see, for example, Berry et al. [1995]). This is reasonable because we work at a much higher level of aggregation. At the firm level, a much larger fraction of costs will be variable than at the model level. At the same time, the residual demand for a firm will be less elastic than for an individual model.

[Table 2 approximately here]

For the results in the third column, a 1% price decrease has the same effect on demand as a 5% increase in the knowledge stock. A value of 0.44 for θ_ω means that a firm with a patent stock half of GM's, will on the average have a market share that is 30% ($= 0.44 \times \ln(0.5)$) lower than GM's, holding price constant.

4.1.2 Estimation of Cost Parameters

We do not observe marginal cost, but once we have the estimates for θ_ω, θ_p, and the vector $\boldsymbol{\xi}$ from the demand system, we use the system of first order conditions (4) to recover marginal costs. Assuming firms are setting prices optimally, we solve for the marginal costs that rationalize the observed prices and market shares. While it is possible to impose optimal price setting in the demand estimation to increase precision, we chose not to do so as it would force the price elasticity to exceed unity. We denote this new variable by mc. While the marginal costs we recover vary by firm and year, we need to be able to predict marginal costs at any possible state in order to

calculate the value function. To get some idea about the relationship between marginal cost and the state variables, we plot the marginal cost against the patent stock (ω) and unobserved product characteristics (ξ) in Figure 3.

[Figure 3 approximately here]

In part (a) of the figure, for low values of the patent stock, an increase in patent stock appears to be associated with higher marginal cost. For higher values of patent stock, the relationship is negative. This suggests a quadratic relationship between the marginal cost and the patent stock. In part (b) of the figure, there is clearly a positive relationship between the marginal cost and the unobserved product characteristics. Motivated by these pictures we specify the marginal cost function as:

$$mc_j = \gamma_0 + \gamma_1(\omega_j/\omega_0) + \gamma_2(\omega_j/\omega_0)^2 + \gamma_3\xi_j + \gamma_4\xi_j^2 + \epsilon_j^{mc}. \tag{11}$$

We also try two alternative specifications: (i) we impose that $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$, to study the case in which marginal cost is constant and does not vary with the technological knowledge of the firm; (ii) we use $\log(\omega_j + 1)$ instead of ω_j, to impose diminishing returns to knowledge in terms of cost savings. The parameter estimates are in Table 3.

[Table 3 approximately here]

The estimates in the table show that all three specifications fit the data well. However, since the marginal cost was derived using ω and ξ vectors, it is problematic to regress the marginal cost on the variables that were used to derive it. In view of this problem, we assume that the marginal cost is constant and not affected by the state variables. This shuts down innovations in the model. Later in Section 6 we study the implications of the other two marginal cost specifications for our estimates.

4.1.3 Estimation of the Policy Function

The next task is to characterize the policy function from the observed investment decisions. Ideally, the control variable x should be modeled as a completely flexible function of a firm's own state and the full vector of its rivals' states that are contained in the industry state. Bajari and Hong [2005] discuss the resulting distribution of the estimator if this step is carried out nonparametrically. Before specifying the policy function, we plot the number of new patent applications against the two state variables separately in Figure 4. There is a clear positive relationship between patent applications and patent stocks in Part (a) of the figure: firms with higher patent stocks tend to file more patent applications. Moreover, the relationship appears to be concave: as patent stock grows, the number of applications does not grow as fast. Apparently there is no clear relationship between

patent applications and unobserved product characteristics in Part (b) of the figure. However, a closer look suggests an inverted-saucer shaped relationship.

Keeping these pictures in mind, we try many different specifications for the policy function and finally adopt the following.

$$
\begin{aligned}
x_j \quad = \quad & \alpha_0 + \alpha_1\omega_j + \alpha_2\omega_{-j} + \alpha_3\xi_j + \alpha_4\omega_j^2 + \alpha_5(\omega_{-j})^2 + \alpha_6\xi_j^2 + \alpha_7\omega_{-j} \\
& +\alpha_8\omega_j\xi_j + \alpha_9\xi_j\omega_{-j} + \alpha_{10}\xi_j(\omega_{-j})^2 + \alpha_{11}\xi_j^2\omega_{-j} + \alpha_{12}\omega_j(\omega_{-j})^2 \\
& +\alpha_{13}\omega_j^2\omega_{-j} + \alpha_{14}\omega_j\xi_j^2 + \alpha_{15}\omega_j^2\xi_j + +\alpha_{16}\omega_j\xi_j\omega_{-j} + \epsilon_j^x,
\end{aligned}
\qquad (12)
$$

where $\omega_{-j} = \sum_{k\neq j}\omega_k$. Under our assumption that firms always play their equilibrium strategy, the OLS estimate of the above equation will in effect give us the equilibrium policy of the firm as a function of its own and the industry state. The error term in equation (12) captures the approximation error between the true policy function $x(\omega_j, \xi_j, \mathbf{s})$ and the one we estimate. The policy function estimates are in Table 4.

[Table 4 approximately here]

Our desire to have a flexible policy function results in most of the coefficients being statistically insignificant. Still we can explain most of the variation in the number of new patent applications ($R^2 = 0.90$). The coefficient on own patent stock (ω_j) is negative but when ω_j is interacted with ξ, ω_{-j} and ξ^2, the coefficients are positive. The coefficients on the sum of rivals' patent stocks (ω_{-j}) and the unobserved product characteristics (ξ_j) are both negative but the presence of so many interaction terms makes their interpretation difficult. Our main objective, as stated above, is to get a flexible policy function and the one reported here serves that purpose very well. This is the policy function we use for forward simulations to get the value function (see below).

An important feature, which is not so obvious from the estimation results, of this policy function is that as we forward simulate the industry, the estimated number of new patent applications does not increase without bound. Instead, there is invariably a finite upper bound for each firm, though the bound is different for different firms. This is an important feature of the policy function and ensures that patent stocks reach a finite stochastic steady state in finite time. In absence of this feature, the number of new patent applications and hence the patent stocks will increase without bound. This will make the problem of finding numerical estimates of the value function intractable.

4.1.4 Estimation of the State Transition Function

The state transition function for the knowledge stock ω is given by (7). It is measured as the accumulated stock of past patents, which decreases exogenously in value because of economic obso-

lescence of knowledge and expiration of patent rights and increases with newly acquired knowledge captured by patent applications (x).[18]

The only parameters in the state transition function are the depreciation rate (δ) and the parameters of the distribution of ϵ^{ω}. We do not estimate these parameters from the data, but assign them values that we believe are reasonable.

We assume a depreciation rate of 15% (i.e. $\delta = 0.15$).[19] This is the same depreciation rate used in Cockburn and Griliches [1988] to construct R&D stock. It captures both patent expirations and the economic obsolescence of older knowledge. The standard deviation of ϵ^{ω} is set to 10% of x. Given the assumption of a normal distribution with mean zero for ϵ^{ω}, a firm that chooses to have 100 new patents in a particular year has a 67% probability of ending up with 90 to 110 patents.[20]

An alternative approach would be to let the control x be the value of R&D and estimate a patent production function $\Delta\omega_t = f(\omega_{t-1}, x, \epsilon)$ from the data. The stock of knowledge would then evolve according to $\omega_t = (1-\delta)\omega_{t-1} + \Delta\omega_t$. In our model, the cost of innovation is estimated at cx and we will compare this with observed R&D expenditures as a reality check.

We do not observe ξ's. Instead we recover them as residuals from the demand equation. Using these recovered ξ's as data, we estimate the AR(1) model in (8). Our estimate of ρ is 0.6 with a standard error of 0.04. The estimated residuals ($\hat{\epsilon}_j^{\xi}$) are almost normally distributed with mean zero and the standard deviation is 0.44. Hence for simulation purposes we assume that each firm's ξ follows an AR(1) process with an error term draw from a normal distribution with zero mean and 0.44 standard deviation. When two firms merge, we simply assign the average of individual firm ξ's to the merged firm.

We estimate the law of motion of the only industry level state variable (m) as a linear trend. The estimated equation is:

$$
\begin{aligned}
m &= 4.14E7 + 7.81E5 \cdot t \\
& \quad\; (8.72E5) \quad (6.23E4),
\end{aligned}
\tag{13}
$$

where t is the year (e.g. 1982), and the numbers in parentheses are standard errors.

4.1.5 Computation of the Value Function

We can put the previous four building blocks together to obtain a numerical estimate of the value function starting from any state $(\omega_j, \xi_j, \mathbf{s}^{-j}, m)$, given the structural parameter vector $\boldsymbol{\theta}$. For the

[18] This idea of patent stock is similar to the one used by Cockburn and Griliches [1988].

[19] Of course, we use the same value of x to construct the patent stock from patent data.

[20] Although we choose the standard deviation of $_{\omega}$ arbitrarily, our estimate of the structural parameter is hardly affected by this choice. This is because of zero mean assumption.

simple model described above, $\boldsymbol{\theta}$ consists of a single parameter: c – the R&D cost required to obtain one new patent in expectation. For now, the evaluation of the value function will be conditional on a starting value for this parameter (c_0). In step 2 below we use the equilibrium conditions of the model to derive a minimum distance estimator of this parameter.

To evaluate the value function we use forward simulation. We first explain our forward simulation procedure for the case when there are no mergers. For an initial industry state $\mathbf{s}_0$, the estimated demand model and cost equation give us the equilibrium period profit vector over all active firms in the industry according to (5).[21] The initial state directly determines optimal innovation, according to the estimated policy function, giving the net profit vector: $\pi_0(\mathbf{s}_0) - c_0 x_0(\mathbf{s}_0)$. Next, we use the state transition function to find the next period's state and denote it by $\mathbf{s}_1$. In order to do this, we draw a vector of ϵ^{ω} values, one for each active firm, which enter equation (7). We draw a similar vector for ϵ^{ξ}. Following the same steps as above we can compute the expected net profit in period 1 as $\pi_1(\mathbf{s}_1) - c_0 x_1(\mathbf{s}_1)$, which we discount back to period 0 using an appropriate discount factor β. We continue this process for a sufficiently large number of periods T until the discount factor β^T gets arbitrarily close to zero. In other words, we evaluate the following equation:

$$V(\mathbf{s}_0|\boldsymbol{\theta}) = \mathbb{E}\left[\sum_{t=0}^{T} \beta^t [(\pi_t(\mathbf{s}_t) - c_0 x_t(\mathbf{s}_t)]\right], \tag{14}$$

where the expectation is over future states. We run these forward simulations a large number of times, using different draws on ϵ_{ω}, the only source of uncertainty in the model, and take their average as a numerical estimate of $V(\mathbf{s}_0|\boldsymbol{\theta})$.

The forward simulation of the value function becomes somewhat more complicated in the presence of mergers. At the end of each period, all firms receive a draw from the uniform distribution over the unit interval. If in some period two firms receive a draw below p (this value is calculated in the Appendix) they merge.[22] The ω of the merged firm is the sum of ω's of individual firms and ξ of the merged firm is the average of individual ξ's. This changes the state of the industry in all calculations from that point onwards.

In order to allocate the future profits of the merged firm to the value functions of each of the merging firms, we calculate the share of each firm in their combined value should they have remained independent—according to equation (18). This still requires the calculation of future patent stocks for all firms and the profits for the two merging firms as if the merger had not taken

[21] This subsumes the calculation of the equilibrium price vector by solving the system of first order conditions for all firms.

[22] If only one firm draws a merger, no merger takes place. If two firms draw a merger, they merge. If three firms draw a merger, we merge two of them randomly and the third remains unmerged. If four firms draw a merger, we merge two of them randomly and then merge the remaining two. And a similar procedure is used if more than four firms draw a merger. The probability to draw a merger (i.e.) is different from the probability of actually experiencing a merger (i.e. $_m$). We further clarify this distinction in the Appendix.

place in all future periods. As a result, the computational burden rises substantially. Starting from twenty three firms, the first merger at time τ requires the calculation of additional value functions for twenty two firms, although only $T - \tau$ future periods will be considered, and so forth. We avoid this computational problem by first estimating the values without mergers. We then assign a share to each firm according to its relative value without mergers. For example, if firms A and B merge in period τ and their estimated values in no-merger case are $V_A(\omega_A, \xi_A)$ and $V_B(\omega_B, \xi_B)$ then, for each period after τ, we split the returns between them according to the following weights: $\zeta_A = V_A/(V_A + V_B)$ and $\zeta_B = V_B/(V_A + V_B)$.

The only other parameter we need is the discount rate β, which is set at 0.92. Given the static parameter estimates, an initial state of the industry, and a starting value for the dynamic parameter c_0, we simulate forward the evolution of the industry state and calculate profits for all firms as we go along. We simulate forward for 150 periods—$\beta^{150} = 3.70E\text{-}06$—and construct the value functions as the present discounted value of the profit streams.[23]

4.2 Step 2

In step 2 we use the results from the first stage together with the equilibrium conditions on the MPE to recover the dynamic parameter(s) of the model ($\boldsymbol{\theta}$). The following steps assume that the model is identified and there is a unique true parameter vector $\boldsymbol{\theta}_0$. Bajari et al. [2007] propose a minimum distance estimator for this true parameter vector. Let $\mathbf{x}(\mathbf{s})$ be the equilibrium policy profile. For this to be a MPE policy profile, it must be true that for all firms, all states, and all alternative policy profiles $\mathbf{x}'(\mathbf{s})$

$$V_j(\mathbf{s}, \mathbf{x}(\mathbf{s}), \boldsymbol{\theta}) \geq V_j(\mathbf{s}, \mathbf{x}'(\mathbf{s}), \boldsymbol{\theta}), \tag{15}$$

where $\mathbf{x}' \neq \mathbf{x}$ only at the jth element, as all other firms play their Nash strategy. Equation (15) will hold at the true value of the parameter vector $\boldsymbol{\theta_0}$.

The minimum distance estimator for $\boldsymbol{\theta_0}$ is constructed as follows. For each firm j and each state $\mathbf{s}$ we observe in the sample, we use the forward simulation method to calculate $V_j(\mathbf{s}, \mathbf{x}(\mathbf{s}), \boldsymbol{\theta})$. We do the same calculations using a number of alternative policy profiles $\mathbf{x}'(\mathbf{s})$ and compute the difference $V_j(\mathbf{s}, \mathbf{x}(\mathbf{s}), \boldsymbol{\theta}) - V_j(\mathbf{s}, \mathbf{x}'(\mathbf{s}), \boldsymbol{\theta})$. We denote this difference by $d(j, \mathbf{s}, \mathbf{x}'|\boldsymbol{\theta})$. We then find d for all j, $\mathbf{s}$ and $\mathbf{x}'(\mathbf{s})$ for a given value of $\boldsymbol{\theta}$ and compute the sum of the squared $\min\{d(j, \mathbf{s}, \mathbf{x}'|\boldsymbol{\theta}), 0\}$ terms. This only penalizes the objective function if the alternative policy $\mathbf{x}'$ leads to a higher value function, which should not happen if $\mathbf{x}$ is the MPE profile. The $\boldsymbol{\theta}$ with the smallest sum is our

[23] The values and subsequent estimates of parameter are vitually unchanged if is increased beyond 150. Since the size of is an important contributor towards the computation burden, we use = 150. A smaller value of does affect the firm values and estimate of .

estimate of $\boldsymbol{\theta}_0$, i.e.

$$\hat{\boldsymbol{\theta}}_{\mathbf{0}} = \arg\min \sum_{j,\mathbf{s},\mathbf{x}'} \Big[\min\{d(j,\mathbf{s},\mathbf{x}'|\boldsymbol{\theta}),0\}\Big]^2. \tag{16}$$

The only dynamic parameter in the model is the (average) R&D cost of a obtaining a new patent (c). Each period, a firm chooses the number of patents it would like to add to its knowledge stock, denoted by x. The R&D expenditure this requires upfront is $c \cdot x$ and the firm will obtain $x + \epsilon^{\omega}$ new patents by the next period.

The minimum distance estimator as defined in (16) gives us an estimate of c. To evaluate the objective function we have to specify alternative policies that differ from the equilibrium Nash policy. For each firm j and each industry state $\mathbf{s}$ we specify two alternative policies as: $\mathbf{x}'(\mathbf{s}) = (\iota + ae_j)'\mathbf{x}(\mathbf{s})$, where ι is a vector of ones, e_j a vector of zeroes with a single one at position j (both vectors are of length n) and $a \in \{-0.01, 0.01\}$.

Our results (based on the average of 300 random merger sequences) show that the cost per merger is around 41.1 million dollars.[24] It takes about 43 hours to get this estimate on our desktop with Intel Core 2Duo E6550 (2.33 GHz) processor. We bootstrap our original sample to get the standard errors for c. Our initial results suggest a large standard error. We are still working on this. The reported estimate of c depends on our assumptions about demand, marginal cost and policy functions. It also depends on some other parameters like the depeciation rate of patent stock, the discount factor and the weight given to a European patent when we combine the US and the European patents to construct our patent stock. We examine the sensitivity of the estimate of c to these assumptions in Section 6.

5 Findings

Thus far, we have estimated the structural parameters of the model which gave us some insights into the importance of innovation in the automotive industry. We found evidence of product and process innovation affecting both the demand and cost sides in plausible ways. Firms' optimal innovation policy depends on the state of the industry and the model produces a plausible estimate for the R&D cost of new patents. In this section, we use the model to study the interaction between innovation and market structure. We also explore welfare implications of the model. However, first we compare some predictions of the model with data.

5.1 Model Predictions and Data

The results of this exercise are in Figure 5. We plot the actual and predicted market shares in Figure 5(a). We get the estimated market share from (2) by using equilibrium price vector from

[24] This number is based on a normalized price of $10,000 for a GM vehicle.

solution to the Bertrand-Nash game. The model does a good job of predicting the market share. The correlation between the actual and the predicted market share is 95%. The deviations from the 45° line can be traced to differences between actual and estimated markups, which we plot in Figure 5(b). The actual markup is the ratio of hedonic prices to the marginal cost that we recovered using the first order condition in (4). The estimated markup is the ratio of equilibrium price (from the Bertrand-Nash game) to the constant marginal cost assumed in the analysis. The correlation between actual and estimated markups is 0.94. The deviation between actual and estimated markups is due to the estimation error in marginal cost. Although the marginal cost implied by the first order condition is not constant we assume a constant marginal cost when we compute the spot market equilibrium.

[Figure (5) approximately here]

We examine the predictive power of the estimated policy function in Figure 5(c), where we plot the actual number of patent applications against the estimated number. The model slightly underestimates the number of patent applications, especially for large firms. However, overall the model fairs well and the correlation between the two variables is 0.94. This is reassuring because the policy function is one of the main building blocks of our model.

Finally in Figure 5(d) we plot the average actual patent stock (by year) in black and the average simulated patent stock in grey. We show the simulated patent stock for ten random merger sequences and show the simulated numbers until the year 2014 i.e. ten years out of sample. Here again the model is able to predict the data fairly well, though for most of the 1990s the simulated patent stock is higher than the actual patent stock. This is because the actual patent stock grows slowly until 1997 and then picks up during the last seven years of the sample. This is shown by an abrupt change in the slope of the black line around the year 1997. The estimated patent stock does not have the sharp change around 1997 because it is simulated using the policy function, which is estimated using the whole sample. As a result, the estimated patent stock evolves smoothly relative to the actual.

5.2 Market Structure and Innovation

The question how market structure affects innovation has been extensively studied in the literature and the evidence is mixed. Examining the evidence, Cohen and Levin [1989] (p.1075) cite several studies that find a negative relationship between the level of competition and innovative activity, several others that find a positive relationship, and one that finds an inverted-U relationship. They conclude that "the empirical results concerning how firm size and market structure relate to innovation are perhaps most accurately decribed as fragile."(p.1078)

More recently, Aghion et al. [2005] predict an inverted-U relationship between competition and innovation in a general equilibrium setting and also find empirical support for it in the manufactur-

ing sector of the U.K. In their model, competition is exogenous. An increase in competition leads to more innovation if firms are technologically close to one another and less innovation if they are technologically far apart. The net aggregate effect depends on the steady state distribution of the technology gap across industries. In their model, there is no reverse feedback from innovation to competition. Although we are interested in the same question, our model differs from theirs in two fundamental ways. First, we focus on a single industry and hence abstract from general equilibrium issues. Second, in our dynamic structural model there is a feedback from innovation to the level of competition firms face: innovation today affects the industry state in the next period. Having estimated the parameters of the model, we are now in a position to examine how market structure and innovation interact in our model of the global automobile industry.

The market structure of the global auto industry can most appropriately be classified as an oligopoly and has been so for a long time. However, the extent of competition faced by firms in this industry has changed over time as the industry consolidated. In what follows, we use two definitions of competition: the inverse of the degree of concentration in the industry and the ratio of marginal cost to price. We measure innovation by R&D intensity which we define as the ratio of R&D expenditures to sales. In our model, competition and innovation are determined simultaneously and evolve together. Random mergers reduce the extent of competition exogenously and change the incentives to innovate. Innovation, in turn, changes the (knowledge) states of the industry and influences the nature of competition.

To study this relationship we begin with the actual state of the global auto industry in 1982. Given the state of the industry, firms make their pricing and R&D investment decisions according to our estimated parameters. Firms also draw mergers and merge according to the merging technology specified in the text. Based on these decisions, the firms' draws of unobserved product quality and deterministic increase in the market size, we update the industry state. We continue this forward simulation of industry for fifty periods. For each period of simulation we record firms' decisions and compute the statistics of interest for individual firms as well as for the industry as a whole. We allow mergers until the number of firms in the industry is down to ten. Figure 6 plots the results of this exercise based on a random merger sequence, which is reported in Table 6.[25]

In Figure 6(a), the X-axis indicates firm level competition as measured by the ratio of marginal cost to price.[26] Innovation intensity, R&D expenditures to sales, is plotted on the Y-axis. The saucer-shaped curve shows the result of a quadratic OLS regression of innovation intensity on competition. The figure shows a saucer-shaped relationship between competion and innovation. Although the quadratic curve does not fit the simulated data well (R^2 is generally around 1%), this saucer-shape is quite robust to different merger sequences. When the firm level competition is high

[25] We report the results based on a single merger sequence here, but the conclusions are robust when we use different sequences.

[26] This is just one minus the Lerner's Index. The Lerner's Index is defined as $\frac{p-mc}{p}$.

(firm markups are low), innovation intensity is high because firm sales are low. When the firm level competition is low (markups are high), innovation intensity is high because firms innovate a lot. At very low levels of competition, inovation intensity is back to average but the quadratic regression does not pick it up. When we fit a third or fourth degree polynomial (not shown) to the simulated data we see that innovation intensity is close to its average when competition is very low.

We see a similar patten in Figure 6(b) where competition is measured by the cumulative market share of a firm's competitors. After examining these pictures from a large number of simulations, each using a different random sequence of mergers, we conclude that there is no strong firm-level relationship between competition and innovation in our simulated model. There is a weak saucer-shaped relationship but it is mainly driven by low sales for small firms and by our assumed quadratic relationship. When higher order curves are fitted to this data, there is no clear pattern.

The same conclusion does not hold for the industry-level data. In Figure 6(c) we average the mc/p ratios and innovation intensities over all firms for each time period. Now there is a clear inverted-U pattern. Figure 6(d), which is similar to Figure 6(c) except that the competition is measured by one minus the Herfindhal's index, shows a similar pattern.[27] These figures show that as industry gets more concentrated, innovation intensity first increases and then decreases. To understand this pattern we depict various components that generated Figure 6(c) in Figure 7. Mergers make the market more concentrated over time and we illustrate how several industry variables evolve. The price-cost margin increases (7(a)); sales continue to grow along a linear trend (7(b)); aggregate R&D in the industry first increases and then flattens out (7(c)); which is mirrored in R&D intensity (7(d)). The first two effects follow directly from our assumptions on competition and the absence of an outside good. The trend in aggregate R&D owes its shape to concave policy function and mergers. As the industry evolves, smaller firms tend to innovate a lot. They do so because, due to the steep value function, marginal addition to knowledge adds a lot to their value. However, in our estimated model there are decreasing returns to knowledge. As firms' knowledge increases, benefits from further innovation decline and they innovate less. Mergers tend to expedite this process of decreasing returns by causing discrete increases in the knowledge stock of market participants.

To sum up, when we simulate the industry forward, starting from the acutal state of the industry in 1982, we do not see a clear relationship between the level of competition that firms face and their innovation intensity. However, there is a strong inverted-U relationship between these two variables at the industry level.

The simulations, whose results are reported above, started from the initial state of the industry in 1982, when there were twenty three firms in the industry and almost half of them had a very

[27] Herfindhal's Index $= \sum_{j=1}^{n} \ _j^2$. Where $_j$ is the market share of firm .

small knowledge stock.[28] Initially, R&D grew rapidly, faster than sales, and innovation intensity increased in the earlier periods. Once these firms grew bigger and merged, the decreasing returns caused R&D to flatten while the sales continued to grow along their linear trend and R&D intensity fell. Combining both episodes in Figures 6(c) and (d) gave rise to the inverted-U relationship.

We now repeat the same experiment but start from the industry state in 2004, when mergers

power of the firms increases. This results in higher markups and higher PS.[31] When the number of firms is down to its minimum (i.e. ten), the PS reaches its steady state value. From this point on, there are small random changes in the PS mainly due to random movements in unobserved product quality. When the producers are enjoying a higher surplus, they pass a part of it on to the consumers and the CS also increases. However, in the first 25 periods or so, the CS is almost flat or decreases a little bit. This could be because the relatively smaller firms do not find it optimal to pass any of their surplus on to the consumers. However, once they become sufficiently large due to the mergers, it is optimal for them to pass some of the increase on to the consumers. Once the merger activity is complete, the CS also reaches its steady state. Given the parameter estimates of our model, we predict a steady state PS of close to \$5000 and a steady state CS of about \$4400 (recall that every thing is normalized to the price of a GM vehicle being equal to \$10000).

In the second case that we study, we allow firms to undertake their normal R&D activities. We report the results in Figure 10(b). The CS shows a trend similar to the case when no innovation was allowed. In fact, the steady state CS is around 3.6% higher in this case compared to the case without innovation. Hence innovation benefits consumers by improving the quality of the vehicles. There is no benefit coming from reduction in costs because we assume constant marginal cost. If the marginal cost were allowed to decrease with patent stock, the CS would be even higher. The PS in this case shows a very different trend. In the beginning when firms are still small, the expected returns from R&D are high because the value function is steep when plotted against the patent stock. This induces the forward looking firms to innovate a lot. This high R&D activity eats into firms profits and the resulting PS declines. However, as firms increase their patent stocks either by innovating or by merging with other firms, the concave value function dictates lower levels of R&D activity (relative to the patent stock) and the PS increases. This increase continues after the mergers stop and eventually the PS surplus reaches its steady state value, which is slightly higher than that of the CS. However, the steady state value of the PS in this case is around 2.9% less than the case without innovation. Hence, in this model, as a whole, innovation benefits consumers and hurts producers. However, when a producer is allowed to innovate, she would still optimally choose to do so because if she does not, and her rivals do, she will lose her market share. This strategic behavior is characteristic of an oligopoly market.

What about the total surplus (the sum of CS and PS)? As 10(c) shows, although the steady state total surplus (TS) is roughly the same in two cases, it is lower in the case with innovation for most of the transition period.[32] Hence in this simple model, innovation is bad for the society. This conclusion is counter intuitive and derives from the fact that the policy function is very steep when the patent stock is low and almost flat when it is high. It is also partially the result of our

[31] When we do the same exercise without mergers the PS is lower.

[32] More precisely, the steady state total surplus in the case with innovation is 0 14% higher than in the case without innovation.

constant marginal cost assumption, which shuts down the process innovation.

6 Robustness Analysis

In this section we check the robustness of our estimates under different assumptions on demand, marginal cost, policy function, depreciation of patent stock, discount factor, relative importance of European and American patents and how to combine the state variables when two firms merge.

In Bajari et al. [2007] framework, the estimated policy function plays the key role. The policy function is used in forward simulations, which are used to compute the value function. The estimates of dynamic parameters depend on the simulated value functions. We tried many different specifications for the policy function: some very simple and parsimonious; some very flexible with lots of parameters. Our preferred policy function, given in (12), is non-parametric in levels. In the following discussion we shall call it the baseline policy function. To test for the robustness of the estimate of the dynamic parameter, we tried two other policy functions. First, we tried a parsimonious policy function that had only eight parameters (excluding the constant). The parameters were the coefficients of the following eight right-hand side variables: ω_j, ω_j^2, ω_{-j}, ω_{-j}^2, ξ_j, ξ_j^2, $\omega_j\omega_{-j}$ and $\omega_j\xi_j$. This policy function predicted a higher number of patent applications for a given patent stock compared to the baselind policy function. It was also less concave than the baseline policy function. These features suggest that this policy function would be consistent with the observed patent applications if the R&D cost of a single patent were lower than what would be predicted by the baseline policy function. The reason is simple. One can only justify a higher number of patents if they are cheaper to acquire. The estimate of the cost of R&D per patent (c) under this policy function is \$24.5 million. This is much lower than the estimate of \$41.1 million under our preferred policy function.

Second, we tried a policy function similar to the baseline except that all righthand side variables were in logs. This policy function was more concave than the baseline. So much so that for low and medium levels of patent stocks it predicts a slightly higher (compared to the baseline) number of patent applications and for high levels of patent stocks it predicts a lower number of patent applications. However, the overall net difference between the predictions of this policy function and the baseline are small and the estimate of c under this policy function is \$40.3 million, which is very close to our preferred \$41.1 million.

We next compare the estimates of c under various demand assumptions. Our preferred demand specification is the one using IV estimation without time effects. If we use OLS specification (see Table 2) the parameter estimates are very similar and using this demand the estimate of c is \$39.3 million. On the other hand if we use the demand estimates from the IV regression with the time effects the elasticty of demand increases (in absolute value) from 2.31 to 3.89. This high elasticity

of demand results is lower profits and hence lower values. With the lower values we need lower c to rationalize the observed patent behavior. Indeed, with this demand specification the estimate of c drops to \$25.5 million.

We have assumed in the main analysis that marginal cost is constant. We did toy with other assumptions about the marginal cost despite the fact that making the marginal cost dependent on the very state variables that were used to derive it is problematic. In Table 3 we have reported the marginal cost function under two alternative assumptions. If we assume that the marginal cost is quadratic in state variables, the parameter estimates suggest that the marginal cost increases with patent stock when the latter is low. This reduces profits and hence values. The result is a lower estimate of c when this marginal cost function is used. The estimate of c drops from \$41.1 million to \$31.5 million. On the other hand if we assume that the marginal cost is log-linear in the state variables, the results are similar to the baseline case and the estimate of c drops marginally to \$40.6 million.

How to combine the state variables when two firms merge. This is an important question. In the main text we assume that when two firms merge, their combined knowledge stock is the sum of their individual pre-merger knowledge stocks. We also assume that the unobserved product quality of the merged firm is the average of the unobserved qualities of the individual firms. Here we keep the first assumption and drop the second. The conclusions will apply to the case when assumption one is dropped. First consider the alternative assumption that when two firms merge, the ξ of the combined firm is the maximum of the two individual ξ's. This will increase the demand of the merged firm relative to the baseline case and hence will result in higher profits and values. This implies a higher value of c. Our computations show that the estimated value of c increases to \$46.6 million in this case. We also try a third assumption that the combined ξ is a weighted average of individual ξ's, where the weights are given by the knowledge stock of the respective firms. This assumption increases the value of c only marginally to \$42.4 million.

When two firms merge during the forward simulation of the value function, we need to specify the share of each firm merging firm in the merged firm inorder to compute their separate values as viewed from the starting period. In the main text we used the adjusted values from the case without mergers as weights. Here we first examine the effect on c if the values are not adjusted. If the values are not adjusted the estimated c will slightly drop to \$40.8 million. We also tried a third assumption about weights. Here we just used the present discounted value of each firm's profits as weight ignoring the cost of R&D. This drastically reduced the estimate of c to just \$5.5 million.

When computing the patent stock and also in forward simulations we assume a depreciation rate of 15% for the patent stock. We know examine the effect of relaxing this assumption. Intuitively, if we assume a lower depreciation, it will lead to a higher patent stock. Given the concave policy function, this would imply a lower level of innovation and hence lower values. The result would

be a smaller estimate of c. This indeed is the case. When we assume 5% depreciation for patent stock, the estimate of c drops to \$10.6 million. If we assume 25% depreciation, the estimate of c increases to \$56.3 million.

Due to the nature of the estimation technique, the effect of the discount factor on the estimated c is straignt forward. As the values are computed using forward simulations of the industry, a higher discount factor will lead to higher values. Higher values would justify a higher estimate of c. For example, when we increase the discount factor from our baseline value of 0.92 to 0.94, the estimate of c increases to \$43.5 million. If we reduce the discount factor to 0.90, the estimate of c decreases to \$40.1 million.

When computing the patent stocks from observed patent applications, we treat the European and the American patents differently. We assume that a European stock is 2.2 times as valuable as an American stock. If we treat a European stock in the same way as an American stock the overall patent stock of each firm will be lower. This will have the same effect as using a higher depreciation rate. In this case the estimated c is \$48.1. On the other hand, if we assume that a European patent is 3 times as valuable as an American stock, the estimated value of c drops to \$40.0 million.

7 Concluding Remarks

We construct a dynamic game-theoretic industry model of the global automobile industry that allows for random mergers. We estimate the structural parameters of the model using a new method proposed by Bajari et al. [2007]. Our application illustrates the usefulness of their method in terms of saving in computation time and highlights some of the problems faced by us while employing their methodology.

We then confront the model with the data and find that the model does a reasonable job of capturing the trends in the data. Next, we use the model to study the interaction between market structure and innovation. Our main findings are: (1) At the firm level, there is no clear-cut relationship between competition and innovation; (2) at the industry level, the effects of market structure on innovation depend on the initial state of the industry. If the industry is very fragmented, an increase in concentration (brought about by random mergers in our model) promotes innovative activity. However, if the industry is already concentrated, a further increase in concentration will discourage innovation. A general implication of the last finding is that the observed market strcture of an industry may span only a subset of the entire feasible space, and hence we might observe a monotone positive or negative relationship between market concentration and innovation intensity when the actual relationship might be more complex.

The policy message of this study is clear. The global automobile industry has consolidated enough to provide good incentives for innovation. Any further consolidation is likely to be harmful

for innovation and diminish competition. However, mergers between small firms (or between a small and a medium-sized firm) may still be beneficial.[33]

In its present form the model has an unrealistic implication: mergers are always value destroying for the merging parties. The reason is twofold: the decreasing marginal returns to knowledge and the loss of an independent firm. However, depending on the size of merging firms, mergers can have positive or negative effects on aggregate industry value. Specifically, the bigger the size of the merging firms the better it is for rival firms' values and hence for the aggregate industry value.

Another limitation of the present study is that it does not allow fully endogenous mergers. As we argued above, in our view the idea of endogenous mergers does not fit well into the global auto indutry. However, our proposed modeling environment has the potential to be expanded into a model of compeletly endogenous mergers. We are currently working on this idea and hope to report some interesting theoretical as well as empirical results in future.

The above limitations notwithstanding, we believe that this study is a step forward in the literature on the interaction between market structure and innovation. In line with earlier surveys, we also believe that this relationship is industry specific and generalization are unlikely to be very useful for policy purposes. Nevertheless, the current framework provides a flexible and general approach to study oligopolistic-dynamic interactions and has wider applicability.

References

P. Aghion and R. Griffith. *Competition and Growth: Reconciling Theory and Evidence*. The MIT Press, Cambridge, MA, 2005.

P. Aghion, N. Bloom, R. Blundell, R. Griffith, and P. Howitt. Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120(2):701–728, 2005.

V. Aguirregabiria and C.-Y. Ho. A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments. Working paper, University of Toronto, 2006.

V. Aguirregabiria and P. Mira. Sequential Estimation of Dynamic Discrete Games. *Econometrica*, 75(1):1–53, 2007.

S. Ahn. Competition, Innovation and Productivity Growth: A Review of Theory and Evidence. OECD, Economics Department Working Papers No.317, 2002.

[33] This paper covers the auto industry only up to the year 2004. It also assumes a constant increase in market size. The recent developments in the industry, especially the drastic decline the demand for automobiles, are not built into the model. Hence the conclusion that further consolidation will reduce innovation may not hold in the present circumstances.

P. Bajari and C. L. Benkard. Hedonic Price Indexes with Unobserved Product Characteristics, and Application to Personal Computers. *Journal of Business and Economic Statistics*, 23(1): 307–347, January 2005.

P. Bajari and H. Hong. Semiparametric Estimation of a Dynamic Game Model of Incomplete Information. Working Paper, 2005.

P. Bajari, C. L. Benkard, and J. Levin. Estimating Dynamic Models of Imperfect Competition. *Econometrica*, 75(5):1331–1370, September 2007.

C. L. Benkard. A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft. *Review of Economic Studies*, 71(3):581–611, 2004.

S. Berry. Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal of Economics*, 25(2):242–262, 1994.

S. Berry, J. Levinsohn, and A. Pakes. Automobile Prices in Market Equilibrium. *Econometrica*, 63 (4):841–890, 1995.

A. Caplin and B. Nalebuff. Aggregation and Imperfect Competition: On the Existence of Equilibrium. *Econometrica*, 59(1):26–61, 1991.

I. Cockburn and Z. Griliches. Industry Effects and Approbriability Measures in the Stock Market's Valuation of R&D and Patents. *American Economic Review*, 78(2):419–423, 1988.

W. M. Cohen and R. C. Levin. Empirical Studies of Innovation and Market Structure. In R. Schmalensee and R. D. Willig, editors, *Handbook of Industrial Organization*, volume 2. Elsevier Science, Amesterdam: North-Holland, 1989.

A. Collard-Wexler. Plant Turnover and Demand Fluctuations in the Ready-Mix Concrete Industry. Working paper, NYU Stern, 2005.

U. Doraszelski and K. L. Judd. Avoiding the Curse of Dimensionality in Dynamic Stochastic Games. Working paper, Harvard University, 2004.

U. Doraszelski and A. Pakes. A Framework for Applied Dynamic Analysis in I.O. Working paper, Harvard University, 2006.

R. Ericson and A. Pakes. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *Review of Economic Studies*, 62(1):53–82, 1995.

R. Gilbert. Looking for Mr Schumpeter: Where Are We in the Competition-Innovation Debate? In A. B. Jaffe, J. Lerner, and S. Stern, editors, *Innovation Policy and the Economy (NBER)*, volume 6. The MIT Press, Cambridge: Massachusetts, 2006.

P. K. Goldberg. Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry. *Econometrica*, 63(4):891–951, 1995.

P. K. Goldberg and F. Verboven. The Evolution of Price Dispersion in the European Car Market. *Review of Economic Studies*, 68(4):811–48, 2001.

Z. Griliches. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4):1661–1707, 1990.

T. J. Holmes and J. A. J. Schmitz. On the Turnover of Business Firms and Business Managers. *Journal of Political Economy*, 103(5):1005–1038, 1995.

M. I. Kamien and N. L. Schwartz. *Market Structure and Innovation*. Cambridge University Press, Cambridge, UK, 1982.

S. Klepper. The Capabilities of New Firms and the Evolution of the US Automobile Industry. *Industrial and Corporate Change*, 11(4):645–666, 2002a.

S. Klepper. Firm Survival and the Evolution of Oligopoly. *RAND Journal of Economics*, 33(1): 37–61, 2002b.

A. Pakes and P. McGuire. Stochastic Algorithms, Symmetric Markov-perfect Equilibrium and the Curse of Dimensionality. *Econometrica*, 69(5):1261–1281, 2001.

A. Pakes, M. Ostrovsky, and S. Berry. Simple Estimators for the Parameters of Discrete Dynamic Games (with entry/exit examples). *Rand Journal of Economics*, forthcoming, 2005.

A. Petrin. Quantifying the Benefits of New Products: The Case of Minivan. *Journal of Political Economy*, 110(4):705–29, 2002.

J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987.

S. Ryan. The Cost of Environmental Regulation in a Concentrated Industry. Working paper, MIT, 2006.

J. A. Schumpeter. *Capitalism, Socialism and Democracy*. Harper, New York, 3 edition, 1950.

A. Sweeting. Estimating the Cost of Format Switching. Working paper, Northwestern University, 2006.

G. Y. Weintraub, C. L. Benkard, and B. V. Roy. Markov Perfect Industry Dynamics with Many Firms. Working paper, Stanford University, 2005.

A Appendix: The Merger Technology

When there are more than two firms in the industry and we allow for mergers, the value function of firm j is somewhat more complicated. $p_m(n)$ denotes the probability that a firm experiences a merger at the end of the current period and will be calibrated below. This probability is an increasing function of n and constant across firms. k indexes a firm that j is randomly matched with to merge and $\mathbf{s}'^{-jk}$ is the industry state excluding firms j and k after the merger. The value function of firm j can then be written as

$$\begin{aligned} V_j(\omega_j, \xi_j, \mathbf{s}^{-j}) \quad = \quad & \max_{x_j \in +} \Big\{ \pi_j(\omega_j, \xi_j, \mathbf{s}^{-j}) - c x_j(\omega_j, \xi_j, \mathbf{s}^{-j}) \\ & + \beta \Big[\frac{p_m}{n-1} \sum_{k \neq j} \zeta_j(\cdot) E V_j(\omega'_j + \omega'_k, (\xi'_j + \xi'_k)/2, \mathbf{s}'^{-jk}) \\ & + (1 - p_m) E V_j(\omega'_j, \xi'_j, \mathbf{s}'^{-j}) \Big] \Big\}, \end{aligned} \tag{17}$$

where

$$\zeta_j(\omega'_j, \xi'_j, \omega'_k, \xi'_k, \mathbf{s}'^{-jk}) = \frac{V_j(\omega'_j, \xi'_j, \mathbf{s}'^{-j})}{V_j(\omega'_j, \xi'_j, \mathbf{s}'^{-j}) + V_k(\omega'_k, \xi'_k, \mathbf{s}'^{-k})}. \tag{18}$$

The value is the sum of the period profit, the continuation value should a merger take place (summed over all rival firms k, and the continuation value in the case no merger takes place. As discussed earlier, we set a lower bound on the number of firms in the industry, denoted by $\underline{n}$. When $n = \underline{n}$, $p_m = 0$ and (17) reduces to (6).

We now show how we can impute the probability that a firm will be involved in a merger (p_m), from the observed number of mergers in the data. Let there be $n > 1$ firms in the industry and let p be the probability (same for all firms) that a firm will be up for merger this period. This probability p differs from the merger probability p_m, because mergers only take place if at least two firms are up for merger.

With only two active firm, $p_m = p^2$. With more than two active firms, the probability that two firms merge at the end of the period is the sum of the following probabilities (if more than two firms are up for merger, we pick two at random):

$$\begin{aligned} \Pr(\text{The firm is up for merger and 1 other firm is also up}) &= p P_1^{n-1} \\ \Pr(\text{The firm is up for merger and 2 other firms are up}) &= \frac{2}{3} p P_2^{n-1} \\ \Pr(\text{The firm is up for merger and 3 other firm are up}) &= p P_3^{n-1} \\ \Pr(\text{The firm is up for merger and 4 other firms are up}) &= \frac{4}{5} p P_4^{n-1} \\ \vdots &= \vdots \end{aligned}$$

where P_k^n is the binomial probability that out of n firms exactly k are up for mergers and is given by $\binom{n}{k} p^n (1-p)^{n-k}$. The last term in the above series of probabilities depends on whether n is an

odd or an even number. If n is odd, the last term would be $\frac{n-1}{n} P_{n-1}^{n-1}$ and if n is even, the last term would simply be P_{n-1}^{n-1}. Adding these terms together, we can write the sum as

$$p_m(n,p) = \Big[\sum_{i \in O, i<n} P_i^{n-1} + \sum_{i \in E, i<n} \frac{i}{i+1} P_i^{n-1} \Big] \cdot p, \tag{19}$$

where $O = \{1, 3, 5, \ldots\}$ and $E = \{2, 4, 6, \ldots\}$. We have derived an expression for p_m in terms of n and p. Our next task is to impute p from the data. In our sample we observe 10 mergers in 23 years. We shall assume for the sake of simplicity that these mergers are evenly spread over the entire sample period. Then the expected number of mergers in any period is $\frac{10}{23}$. We make this expected number of mergers depend on n in a simple way. Let $\overline{n}$ be the average number of firms in the industry per period. Then we may write the expected number of mergers as

$$E(M) = \frac{10}{23} \cdot \frac{n}{\overline{n}}. \tag{20}$$

If the actual number of firms in the industry is equal to the average over the sample period, we expect $\frac{10}{23}$ mergers to take place in that period (or we expect 1 merger in every $\frac{23}{10}$ such periods). If the actual number of firms is above the average we expect more mergers and vice versa. The simple idea behind (20) is that as n declines, we reduce the expected number of mergers that we want to match by choosing an appropriate p. In other words, we want to make p small as n declines.

So far we have been trying to get a reasonable number of expected mergers from the data. But if we know p and n, we can easily derive the expected number of mergers by using the following equation

$$E(M) = \sum_{M=1}^{\lfloor \frac{n}{2} \rfloor} M \cdot P(M), \tag{21}$$

where $P(M)$ is the probability that M mergers take place. If $M = 1$ then $P(M) = (P_2^n + P_3^n)$. In words, the probability that one merger will take place is simply the sum of the probabilities that out of n firms 2 or 3 are up for merger. If 3 firms are up for merger, we shall pick 2 at random and the third will not merge and remain independent. Equation (21) can explicitly be written as

$$E(M) = \begin{cases} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor - 1} i(P_{2i}^n + P_{2i+1}^n) + \lfloor \frac{n}{2} \rfloor (P_{2\lfloor \frac{n}{2} \rfloor}^n + P_n^n) & \text{if } n \text{ is odd} \\ \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor - 1} i(P_{2i}^n + P_{2i+1}^n) + \frac{n}{2} P_n^n & \text{if } n \text{ is even.} \end{cases}$$

The last equation implicitly defines p in terms of n and $E(M)$. But in (20) we have already defined $E(M)$ as a function of n. Hence, given n we can use (20) to solve for $E(M)$ and then use the last equation above to solve for p. Once we have p, we can use (19) to get p_m.

Table 1: R&D expenditure by industry in selected countries for 2003 (in PPP $m)

Industry	ISIC Rev.3	OECD[1]	USA	EU	Japan
Motor vehicles	**34**	**76,199**	**17,034**	**21,258**	**12,765**
		(13.2%)	(8.3%)	(16.9%)	(15.1%)
Radio, television, telecom. equipment	32	71,623	22,399	13,812	11,081
Total business sector	15-99	577,316	204,004	125,591	84,676
Pharmaceuticals	24.2	57,541	15,962	16,850	6,363
Medical, precision, optical instruments	33	38,440	20,400	6,782	3,619
Computer and related activities	72	38,189	19,854	6,990	1,814
Aircraft and spacecraft	35.1	34,065	15,731	8,518	385
Wholesale & retail trade; repairs	50-52	30,066	26,580	1,039	578
Office, accounting, and computing mach.	30	24,136	7,664	2,568	10,764
R&D[2]	73	22,417	12,460		4,951

Note: The five most R&D intensive sectors in each of the three country are included and sorted by total R&D expenditure.

[1] We only observe statistics from the following countries: USA, EU, Japan, Canada, Korea, Norway, Poland, Czech Republic.

[2] For the EU, R&D is included in other sectors.

Source: OECD ANBERD database, Version 2, 2006.

Table 2: Demand parameters

	The dependent variable is log market share		
	OLS	IV	IV
Price (ln p)	−2.42*** (0.18)	−2.31*** (0.57)	−3.89*** (1.40)
Knowledge (ln(1+ω))	0.44*** (0.01)	0.44*** (0.02)	0.49*** (0.04)
Time Fixed−Effect	No	No	Yes
R^2	0.72	0.72	−
No. of Observations	414	414	414

Notes: (1) The only instrument in columns 2 and 3 is the sum of all rivals‘ knowledge. All variables are normalized by GM.
(2) Significance at 1, 5 and 10% levels is shown by ***, ** and *.
(3) Numbers in parentheses are standard errors

Table 3: Parameter estimates for the marginal cost function

	The dependent variable is marginal cost		
	OLS	OLS	OLS
Constant	1.22*** (9.81e−3)	1.23*** (0.02)	1.31*** (0.05)
Relative Knowledge (ω_j/ω_0)		0.09* (0.05)	
$(\omega_j/\omega_0)^2$		−0.08*** (0.02)	
Unobs Product Characteristics (ξ)		−0.03** (0.02)	
ξ^2		−0.02 (0.02)	
$\log[(\omega_j+1)/(\omega_0+1)]$			0.01*** (4.32e−3)
$\log[\xi - \min(\xi) + 1]$			−0.05 (0.04)
R^2	0.00	0.05	0.03
No. of Observations	405	405	405

(1) Significance at 1, 5 and 10% levels is shown by ***, ** and *.
(2) Numbers in parentheses are standard errors.

Table 4: Policy function

The dependent variable is the number of new patent applications			
Ind. Var.	OLS	Ind. Var.	OLS
Constant	125.21 (91.26)	$\xi \cdot \omega_{-j}$	−2.11e−3 (0.01)
ω_j	−0.14** (0.06)	$\xi \cdot \omega_{-j}^2$	2.52e−8 (3.73e−7)
ω_{-j}	−0.01 (9.86e−3)	$\xi^2 \cdot \omega_{-j}$	−3.17e−3 (3.90e−3)
ξ	32.35 (116.56)	$\omega_j \cdot \omega_{-j}^2$	−7.55e−10*** (1.61e−10)
ω_j^2	−7.81e−6 (1.82e−5)	$\omega_j^2 \cdot \omega_{-j}$	−1.23e−10 (7.88e−10)
ω_{-j}^2	2.89e−7 (2.48e−7)	$\omega_j \cdot \xi^2$	0.03* (0.02)
ξ^2	42.84 (52.55)	$\omega_j^2 \cdot \xi$	−4.78e−6 (1.05e−5)
$\omega_j \cdot \omega_{-j}$	3.47e−5*** (6.21e−6)	$\omega_j \cdot \xi \cdot \omega_{-j}$	1.95e−6 (2.77e−6)
$\omega_j \cdot \xi$	−0.02 (0.04)	R^2 / Observations	0.93 / 368

Notes: (1) ω_j = Own Knowledge.
(2) ω_{-j} = Sum of Rivals' Knowledge.
(3) ξ = Unobserved Product Characteristics
(4) Significance at 1, 5 and 10% levels is shown by ***, ** and *.
(5) Numbers in parentheses are standard errors

Table 5: List of Abbreviations

Abbreviation: Full Name			
ALR:	Alfa Romeo	KIA:	Kia
BMW:	Bayerische Motoren Werke	MTB:	Mitsubishi
CHR:	Chrysler	NSN:	Nissan
DBZ:	Daimler−Benz	PSA:	Peugeot−Citrion
DSU:	Daihatsu	REN:	Renault
DWO:	Daewoo	ROV:	Rover Group
FIA:	Fiat	SAB:	Saab
FRD:	Ford	SZK:	Suzuki
GMS:	General Motors	TYO:	Toyota
HND:	Honda	VOL:	Volvo
HYU:	Hyundai	VOW:	Volkswagon
JAG:	Jaguar		

Table 6: Merger Sequence 1

Period	Merging Firms
2	VOL VOW
6	VOL VOW HND
6	FIA DSU
7	REN FRD
8	JAG GMS
11	FIA DSU TYO
11	DWO HYU
12	DBZ PSA
19	JAG GMS DBZ PSA
21	ALR DWO HYU
25	SAB BMW
29	NSN JAG GMS DBZ PSA
33	ROV SZK

Table 7: Merger Sequence 2

Period	Merging Firms
3	PSA FIA
5	HYU HND
17	MTB BMW

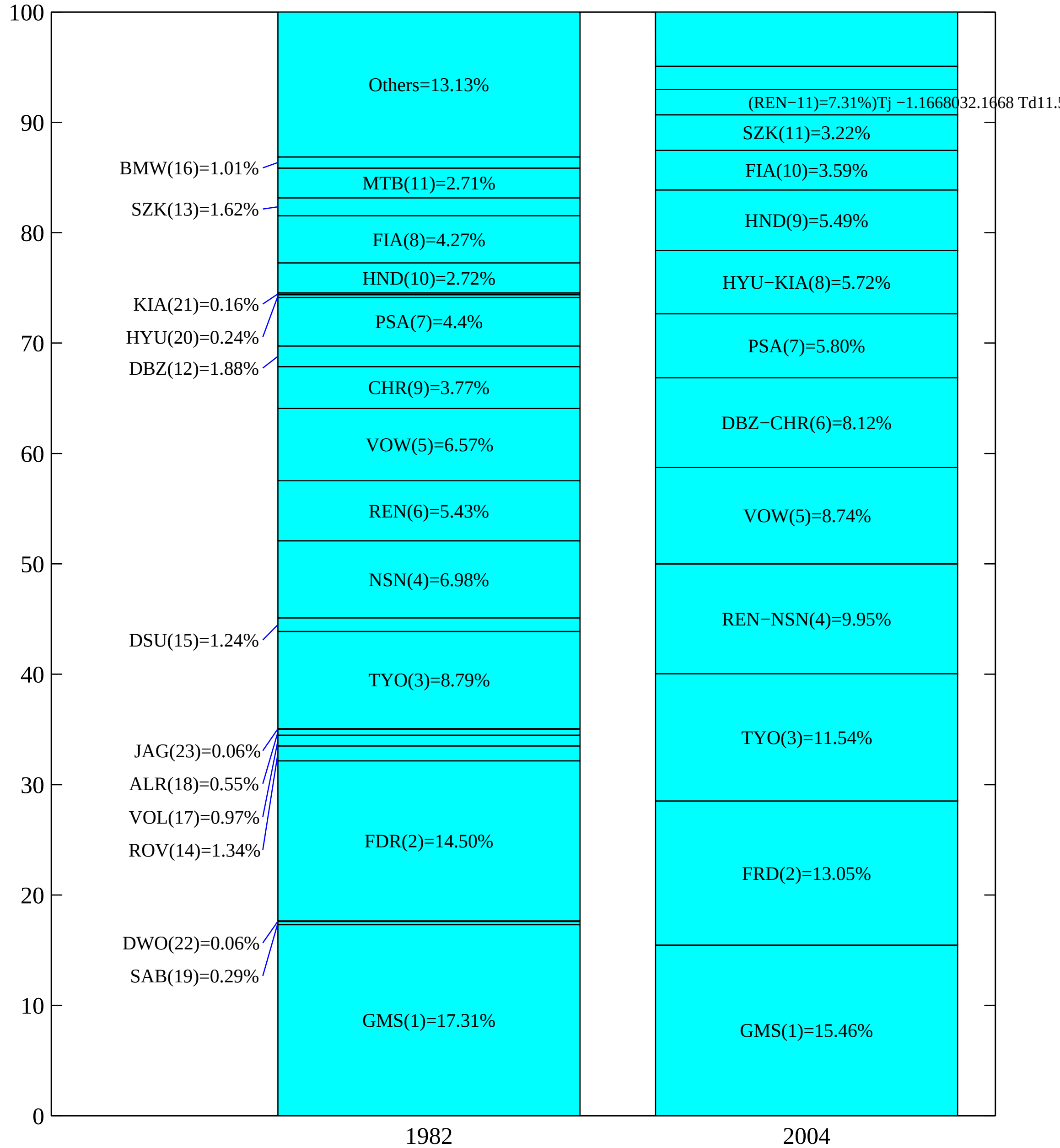

Figure 1: Market shares of the firms in the sample

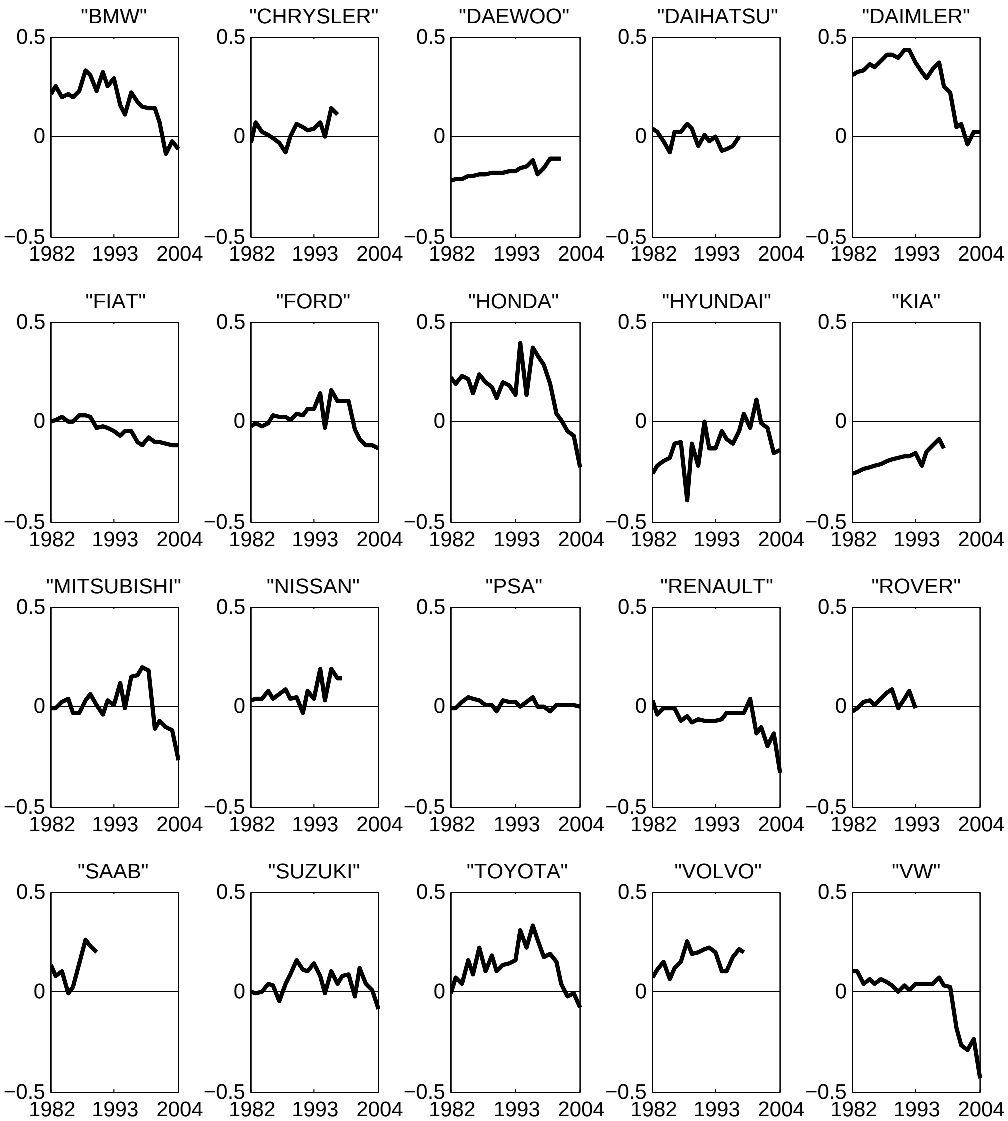

Figure 2: Hedonic prices for a number of firms (relative to GM)

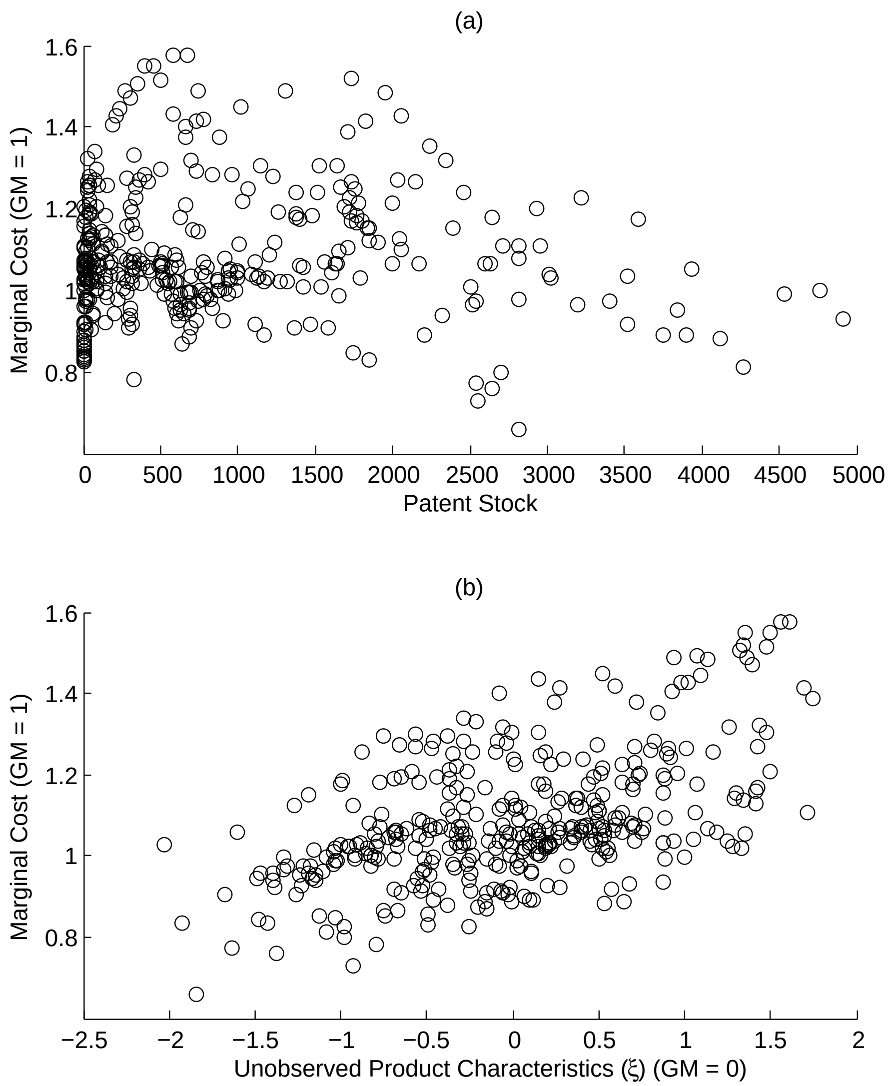

Figure 3: Marginal Cost, Patent Stock and Unobserved Product Characteristics

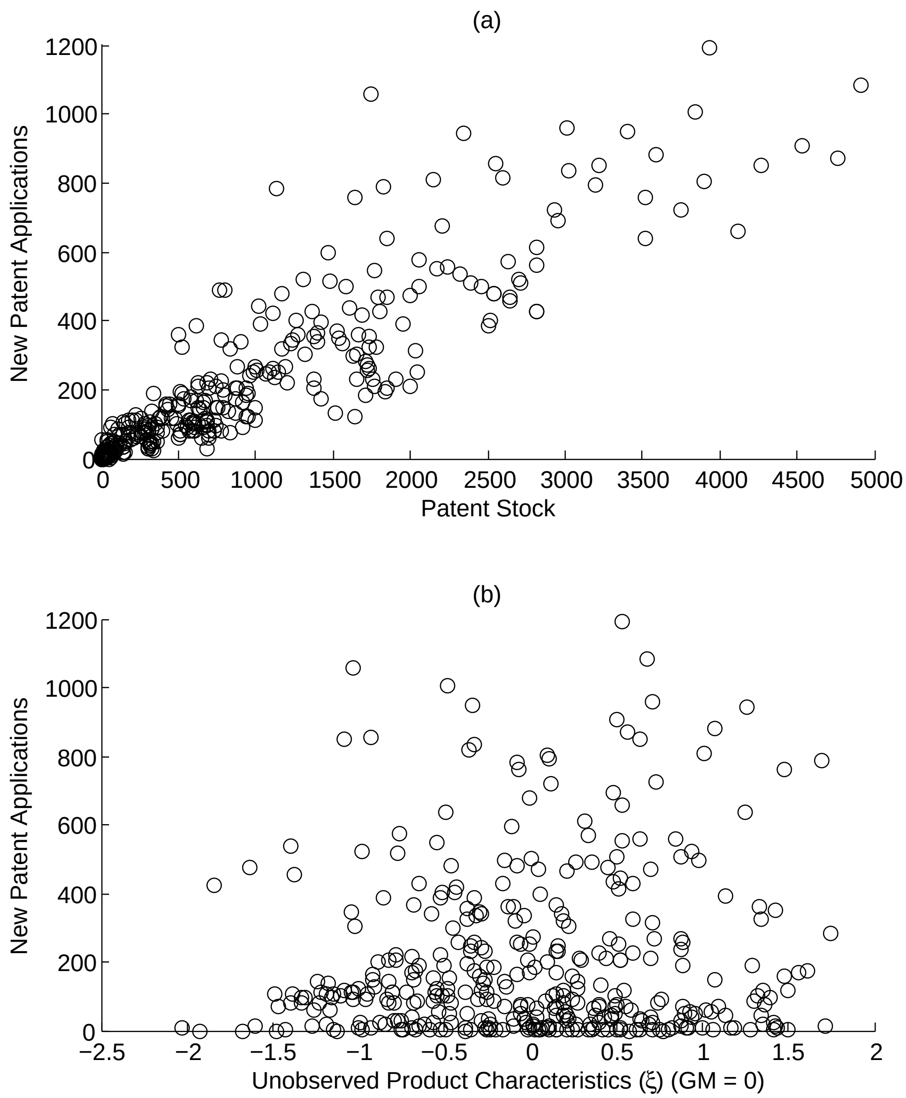

Figure 4: New patents, patent stock and unobserved product characteristics

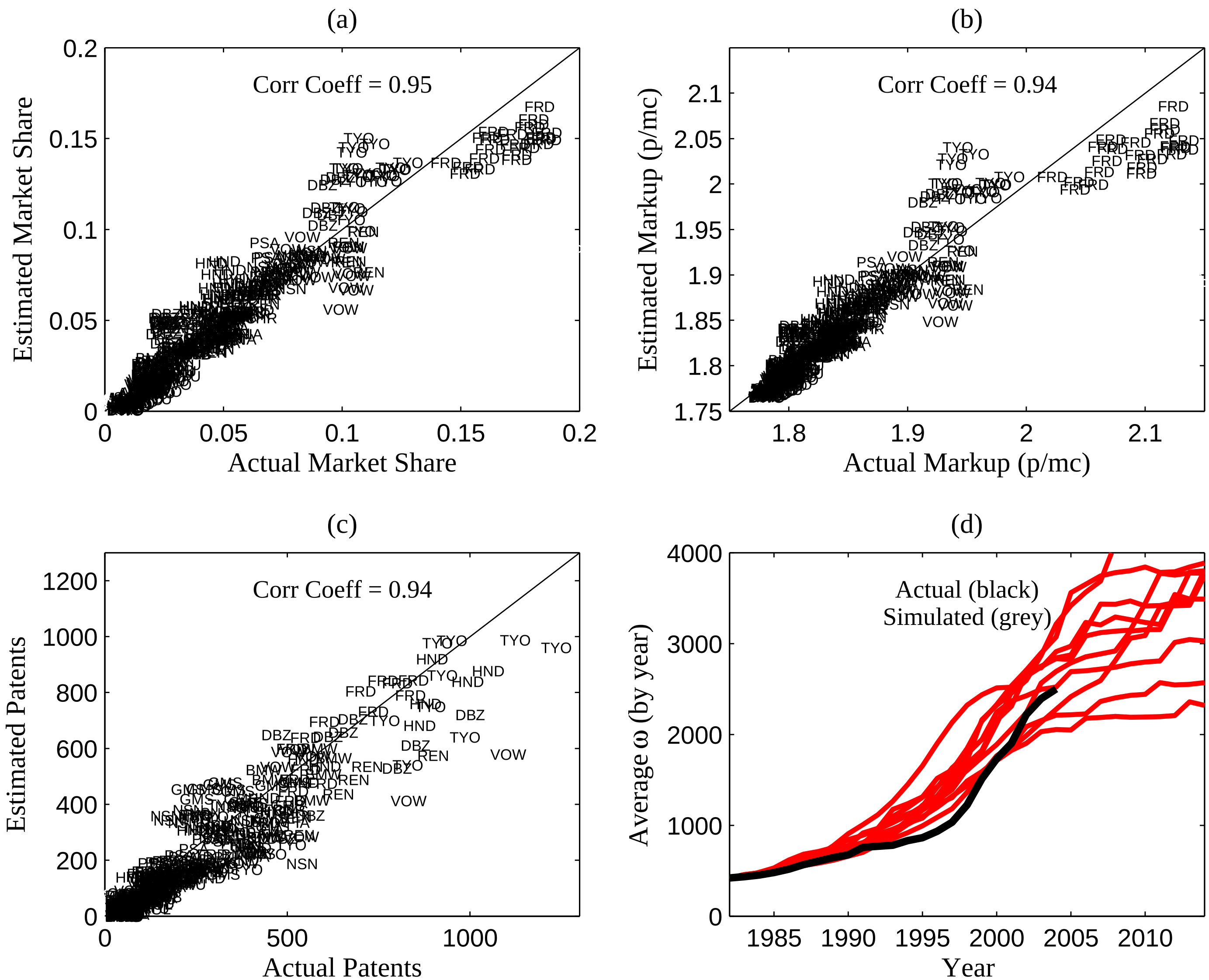

Figure 5: Model vs Data

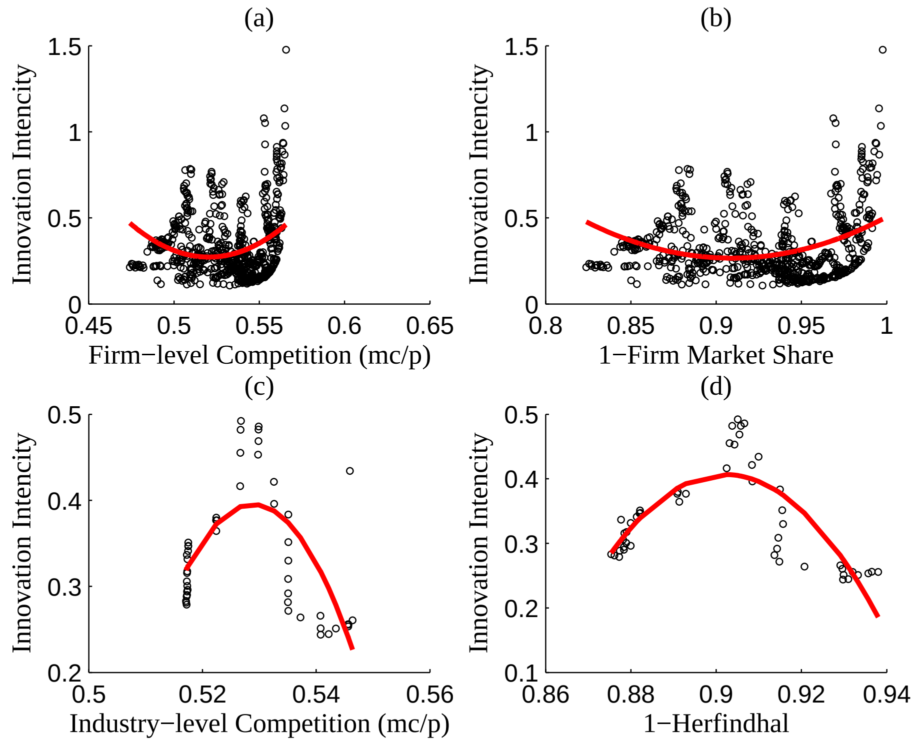

Figure 6: Competition and Innovation at firm and industry levels

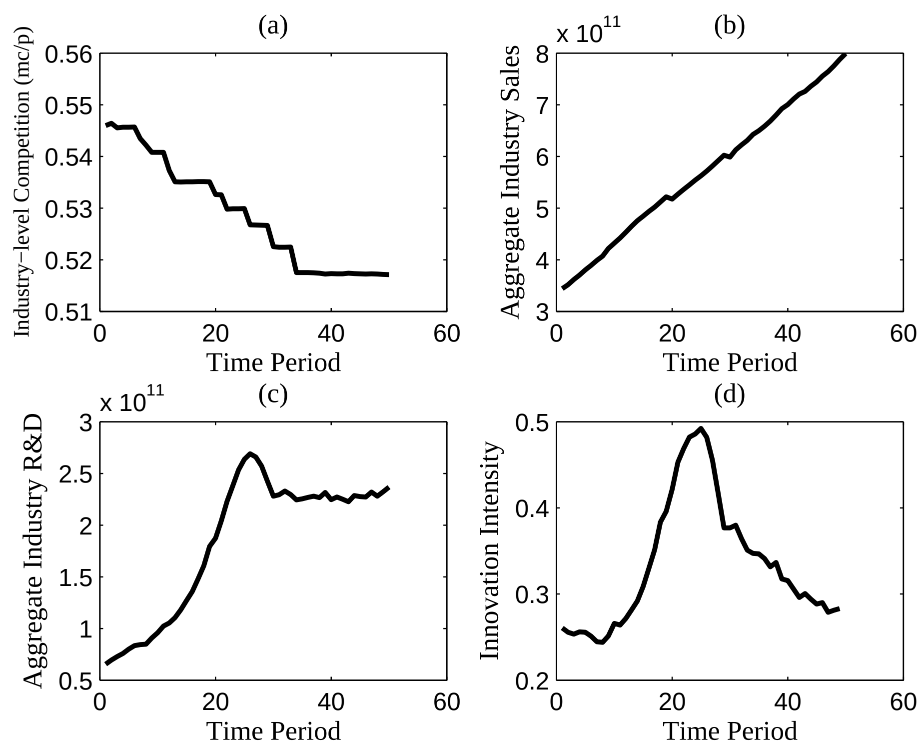

Figure 7: Competition and Innovation over time

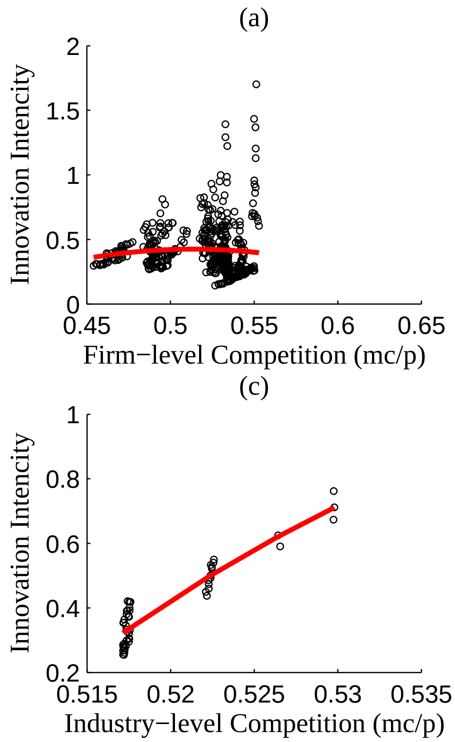

Figure 8: Competition and Innovation at firm and industry levels

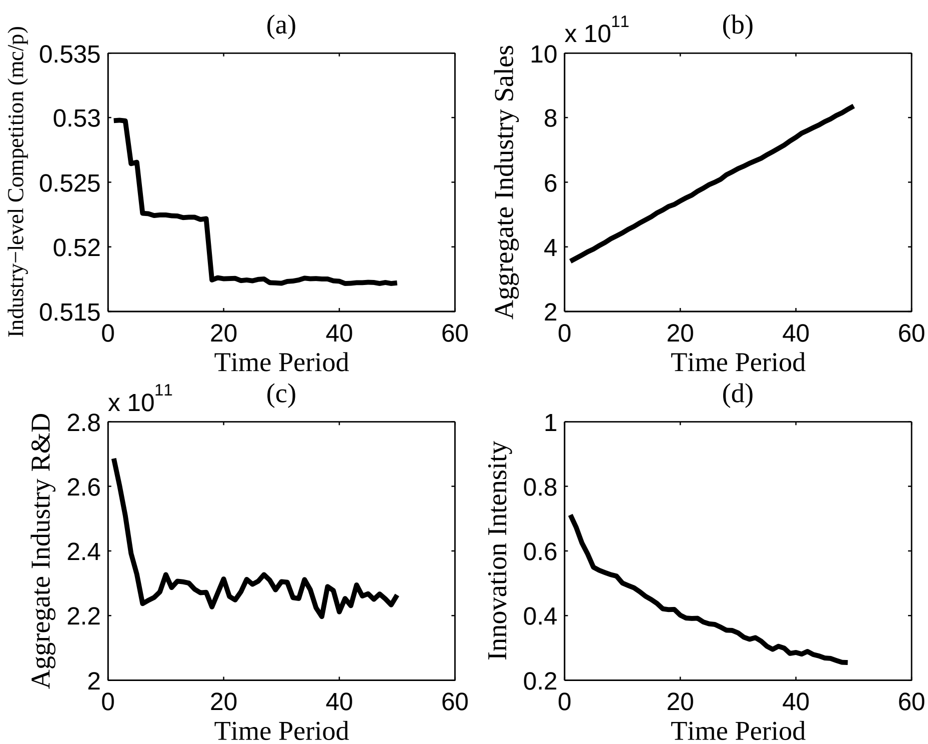

Figure 9: Competition and Innovation over time

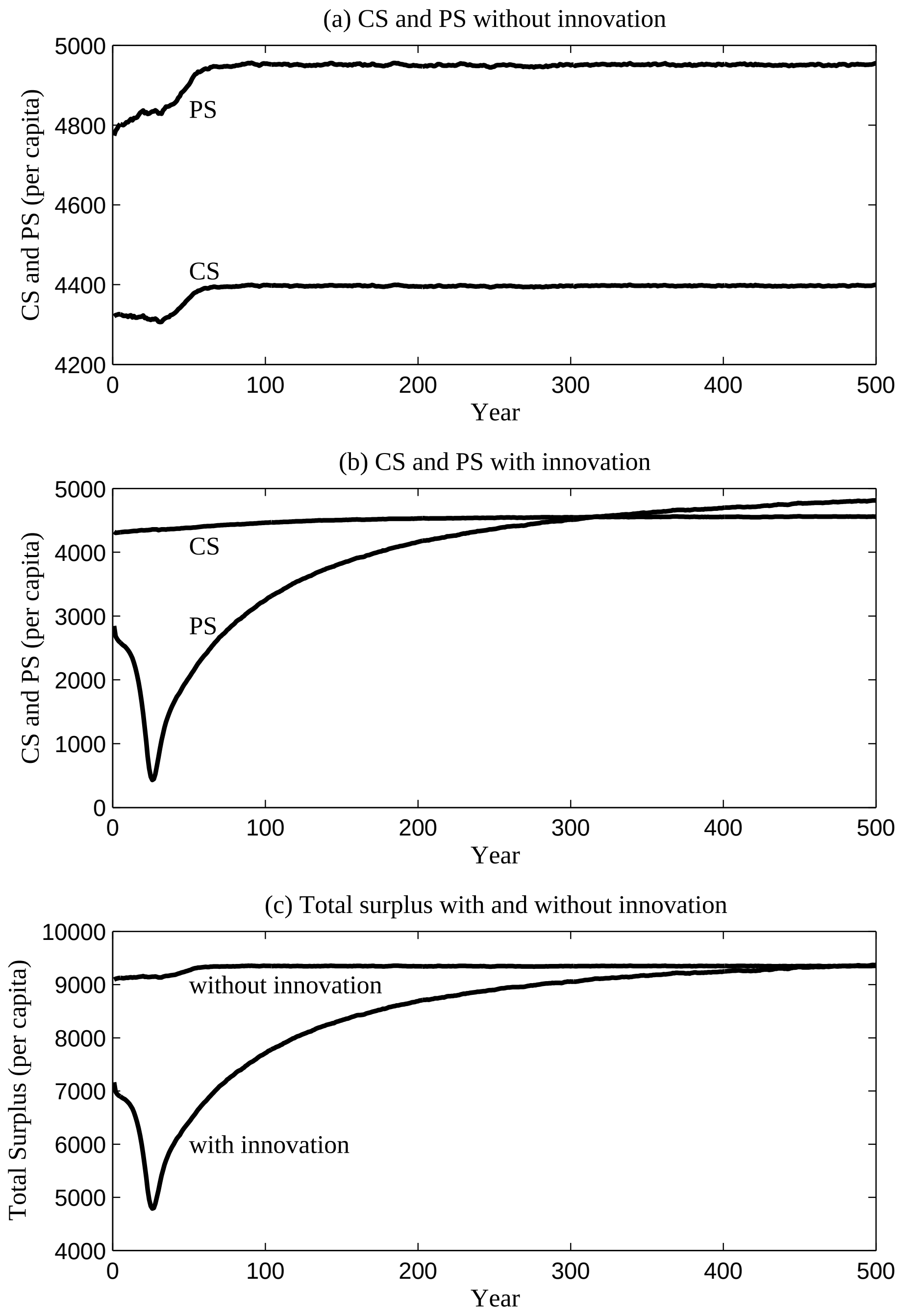

Figure 10: Consumers', producers' and total surpluses (with and without innovation