CENTRE for ECONOMIC PERFORMANCE

CEP Discussion Paper No 958

November 2009

Recovering the Sunk Costs of R&D: the Moulds Industry Case

Carlos Daniel Santos

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Abstract
Sunk costs for R&D are an important determinant of the level of innovation in the economy. In this paper I recover them using a Markov equilibrium framework. The contribution is twofold. First, a model of industry dynamics which accounts for selection into R&D, capital accumulation and entry/exit is proposed. The industry state is summarized by an aggregate state with the advantage that it avoids the "curse of dimensionality". Second, the estimated sunk costs of R&D for the Portuguese moulds industry are shown to be important (3.4 million Euros). They become particularly relevant since the industry is mostly populated by small firms. Institutional changes in the early 1990s generated an increase in demand from European car makers and created the incentives for firms to pay the costs of investment. Trade-induced innovation reinforced the selection effect by which international trade leads to productivity growth. Finally, using the estimated parameters, simulations evaluate the effects of changes in market size, sunk costs and entry costs.

Keywords: Aggregate state, industry dynamics, Markov equilibrium, moulds industry, R&D, structural estimation, sunk costs
JEL Classification: C61, D21, D92, L11, L22, O31

This paper was produced as part of the Centre's Productivity and Innovation Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

Acknowledgements
I would like to thank John Van Reenen and Philipp Schmidt-Dengler for all their guidance and support. I would also like to thank Jérôme Adda, Stephen Bond, Ulrich Doraszelski, Cristian Huse, Sorin Krammer, Vadym Lepetyuk, Mark Roberts, Bernard Salanié, Gabriel Weintraub and participants at several meetings and conferences for very useful comments and suggestions. This manuscript is part of my Ph.D. dissertation at The London School of Economics. Research supported by the Portuguese Science and Technology Foundation grant SFRH/BD/12092/2003.

Carlos Daniel Santos is an Associate of the Centre for Economic Performance, London School of Economics. He is also Assistant Professor of Economics, Department of Economics, University of Alicante, Spain. Contact email: csantos@ua.es

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

In this paper I document one effect of trade on innovation, by focusing on sunk costs of R&D. Reducing trade barriers allows firms to access larger markets. In the presence of sunk costs, this can give firms enough scale and create the right incentives for them to devote more resources towards Research and Development. Innovations can then generate future productivity growth, further increasing the gains from trade.[1] I structurally estimate the size of the sunk costs for the Portuguese moulds industry after joining the EU in 1986. Using the estimated "deep" parameters I perform counterfactual policy changes in market size that illustrate this mechanism. If the market size was exogenously reduced to the beginning of the sample level, the model predicts a reduction in R&D performance, average productivity, capital stock and number of firms. A second contribution is the estimation of a dynamic game with many players. I hope this opens the avenue for an increase in the use of structural estimation methods with microdata to studies on industry dynamics.

The data normally reveals large productivity and size differences between R&D and non R&D firms. Firms decide on R&D start-ups depending on the costs and benefits. However, these observed differences can be driven directly by R&D expenditures and/or selection into R&D, therefore leading to the question of which came first. This creates a problem for identifying the costs and benefits of R&D. To understand why, notice that large differences between R&D and non-R&D firms can signal large benefits for R&D (observed) and can be rationalized by the large costs of R&D. In this case R&D expenditures cause observed differences. Alternatively, observed differences can be due to a very strong selection effect into R&D, motivated by (heterogeneous) benefits being larger (or costs being lower) for firms with larger productivity and/or capital stocks. These two (and other) possibilities are observationally equivalent and therefore indistinguishable without closer inspection. Furthermore, it is most likely that both play a role. One way to address this question is by directly modelling these effects and using a structural model (that gives us the unobserved counterfactual) to estimate the costs and benefits. It is important to stress that the selection of larger and more productive firms into R&D is directly addressed in the model.

In this paper the role of sunk costs of R&D is investigated as a main force driving the discrete decision to become an R&D firm. Estimating sunk costs of R&D is important because it will determine R&D performance and in particular the effects market size can have on industry innovation and productivity, topics of extreme relevance for policy makers. In the presence of sunk costs, market growth will trigger R&D start-ups. For most industries, only a fraction of the firms actually perform R&D. The reason must be that: either the returns from R&D are too low or; the costs involved (not necessarily sunk) are very high and prevent firms from engaging in R&D. The evidence suggests significant returns from R&D which gives support to the latter explanation. Recovering the dynamic cost parameters, gives the opportunity to perform counterfactual analysis on the impact of changes in the sunk costs to the amount of

[1] Pavcnik (2002) estimates these gains for the Chilean industry and De Loecker (2007) for the Belgian textile industry.

R&D expenditures and industry structure.[2]

The Portuguese moulds industry has been a heavy exporter since the beginning of its existence and grew substantially during the 1994-2003 period (almost three fold increase in total sales). One of the reasons for this growth was the increase in demand from European car-makers after Portugal joined the EEC in 1986. This can be seen by the decrease in the share of exports going to the US (traditionally the larger market) in favor of Europe (mostly Germany, France and Spain). During this period there was also an increase in R&D and innovation with the strategy adopted by some players being to reinforce strong links with clients, develop new materials (product innovation) and minimize waste (process innovation). It has been documented (Beira et al, 2003; IAPMEI, 2006) that the close cooperation with car makers was a strong push towards the development of new processes and products. This type of "demand driven" innovation is not uncommon in industries where products are non-standardized and there is normally very close collaboration between supply and demand as in the case of moulds. Therefore, due to the lack of a national market, access to large foreign clients was a strong driver for the success of the Portuguese industry. The reason why car makers preferred Portuguese moulds in the first place was their recognized competence, technical skills and price competitiveness. A report from the US international trade commission (USITC, 2002) emphasizes the fast delivery, technology, quality and competitive price as the main strengths of the Portuguese moulds industry. For example, CAD/CAM technology was first introduced in the 1980's and this was a requirement from car-makers in order to ensure compatibility of the design of moulds (Beira and Menezes, 2003).

The contributions in this paper are twofold. First, a model of industry dynamics which can be used empirically is proposed. This is done by assuming that firms' individual states are private information and that the industry state is summarized by a commonly observed aggregate state. As a result there are two advantages for estimation: (i) it avoids the "curse of dimensionality", typical in dynamic industry games and; (ii) it deals with unobserved firms in the data, a problem that arises if one wants to estimate using equilibrium conditions. Second, sunk costs of R&D are estimated for the Portuguese moulds industry and found to be large (more than one year worth of sales for an R&D firm). This is done using the method developed by Bajari, Benkard and Levin (2007), henceforth BBL, to estimate dynamic games.

Sunk costs become particularly relevant because this industry is populated by many small firms. Institutional changes (joining the EEC in 1986 and the Single European Market in

[2] One hotly debated (and unsolved) issue is the link between competition and R&D performance. Aghion et al. (2005) provide a theoretical explanation and some empirical evidence arguing that there is an inverted U-shape relationship between the two, whereby innovation is higher for mid levels of competition but lower for either very competitive or weakly competitive industries. Blundell, Griffith and Van Reenen (1999), by contrast, find that the pre-innovation effect dominates. However, since both market structure and R&D performance are jointly determined in equilibrium, it is not easy to disentangle these effects without a dynamic model that addresses the market structure endogeneity issue.

1993) caused an increase in demand from European car makers. Access to a large market gave incentives for firms to invest in R&D and physical capital which later translated into increases in labor productivity. This mechanism of *trade-induced innovation* (i.e., gaining access to large markets makes it profitable to sink money into R&D projects) seems to reinforce the selection effect by which international trade can lead to productivity growth.

Regarding unobserved players, most firm level datasets[3] contain information on financial variables (balance sheets, profits and losses, number of workers) for a subset of the total population of firms in the industry. To estimate game theoretic type of models where players strategies depend on the state of all competitors, requires observing all players in the industry (if only the distribution of states is relevant, as happens when imposing symmetry and anonymity (Doraszelski and Satterthwaite, 2007), only data for the distribution of states is needed which can potentially be recovered from the sample). To see this, imagine that we want to estimate a policy function as a function of the state of all (N) competitors in the industry, $\sigma(s_1, ...s_N)$. If there is data on actions and individual states, this can easily be done non-parametrically. However, if some players are not observed we immediately face a problem of unobserved heterogeneity since some important variables are unobserved. So, either we control for this unobserved heterogeneity in some way or we face problems in estimating the policy functions.

Most studies in empirical Industrial Organization have then focused on oligopolies or regulated industries where good information for all players in the market is available, but this leaves aside a large number of industries which are relevant and interesting cases to study. In this paper the proposed framework allows us to estimate a structural model without facing these problems. Furthermore, for questions like the sunk costs of R&D, oligopolist markets are less attractive since in most cases all firms are sufficiently large and the sunk cost of R&D might not bind.

Aw, Roberts and Xu (2009) document a set of findings regarding the interaction of R&D and export status, namely the self selection into exporting and R&D of high productivity plants and the impact of this in reinforcing their productivity advantage. They develop a single agent framework where firms decide on entering the export market and doing R&D. Their model is very similar in spirit to the one proposed here. The differences are that on the one side it allows endogenous entry decision into the export market, something that I do not model due to problems with little variation in observed export status. On the other side, they do not model capital accumulation or the dynamic industry equilibrium. The second is an important disadvantage because it is hard to evaluate policy changes within a single agent framework.

[3]Examples of these are Standard & Poor's COMPUSTAT for US firms, Bureau Van Dijk's FAME (UK) and AMADEUS (Europe) or Thomson Financial's DATASTREAM (UK). Only census data would contain observations for all firms present in the industry and even in this case smaller firms are sometimes sampled.

The framework developed in this paper is a dynamic equilibrium model with productivity, physical capital accumulation and entry and exit within a monopolistic competition setting. There are both linear and quadratic costs with total irreversibility for physical capital investment. Productivity follows a first order Markov process which depends on whether the firm is an R&D performer or not. Finally, firms compete in the market where demand is modeled by a representative consumer with constant elasticity of substitution utility.

The reason to introduce both (total factor) productivity and capital stock is to account for two important characteristics behind firm and industry dynamics, size and productivity. In this way, using total factor (as opposed to labor) productivity is important since labor productivity is normally not scale free (i.e. firms with larger capital stocks have, *ceteris paribus*, higher labor productivity).

One hypothesis that could explain observed behavior is unobserved heterogeneity in the returns from (and costs of) R&D. In principle this could be relaxed by letting either benefits or costs depend on unobserved heterogeneity, but not both. However, this can only be done with a sufficiently large time series so that firm fixed effects can be consistently estimated. Notice however that part of the unobserved heterogeneity in returns is still accounted for by recovering total factor productivity estimates via production function.

There are considerable costs of becoming an R&D firm that besides producing moulds, is also able to supply its clients with conception and design skills, moulds testing and development of new materials, all at a competitive price. A successful innovative firm is able to produce not only the mould itself but also deliver all the pre and post production services required by their clients. The costs of R&D we consider can range from the training and hiring of new employees, investment in new machinery or even the establishment of links with universities and public research agencies. These costs motivate the idea of sunkness.

Sunk costs have for a long time been regarded as one potential source of inefficiency in the economy. The earlier literature emphasizes the failure of the contestability theory in the presence of sunk entry costs, which results in market failures because the industry will not be competitive and firms can maintain some degree of market power (Baumol and Willig, 1981; Stiglitz, 1987). The issue is of great importance for policy makers and regulators since their existence results in a market failure which induces the need for policy intervention.

Sunk costs of R&D, in particular, have been widely studied in the industrial organization literature, especially following the work by Sutton (1991, 1998). The main objective of this research was to explore the relationship between R&D and market structure. In particular, firms can use R&D as a strategic tool to increase barriers to entry and maintain a dominant position even for large market size. One question raised by Schmalensee (1992) is how will the incumbent maintain a dominant position, in the cases where R&D does not have a "forever lasting" effect and therefore does not create a "forever lasting" advantage/barrier. However, the study of more complex dynamics for the outcome of R&D requires a fully dynamic model that goes beyond the two period approach and this type of framework was at the time in an early development stage. Dixit (1988) acknowledges this in his work

> *"Perhaps the most important aspect ignored here is the possibility of partial progress (state variables) in the R&D race. That has so far proved intractable at any reasonably general level, but remains an important problem for future research".* Dixit (1988: 326)

As explained above, the incomplete information assumption whereby players only observe an aggregate state addresses two problems both avoiding the "curse of dimensionality"[4] by reducing the dimensionality of the state space and dealing with unobserved firms in the data by only requiring the aggregate industry state to be observed. Industry state can be summarized by the (payoff relevant) aggregate state and the notion of equilibrium is then very intuitive. Given the beliefs about the evolution of the aggregate state, agents behave optimally. Evolution of the industry state resulting from agents' optimal decisions is consistent with their (rational) beliefs. Notice that restricting the strategies to be functions of the payoff relevant variables (in our case the aggregate state) is common in the theoretical literature (Maskin and Tirole, 2001). The main problem is to guarantee that the equilibrium transition for the aggregate state is Markovian. In a sense this is close to macro-style models similar to Krusell and Smith (1998).

In related research Weintraub, Benkard and Van Roy (2008) propose the use of a different equilibrium concept, the "Oblivious Equilibrium". In this type of equilibrium firms disregard the current state of the industry and base their decisions solely upon the long-run industry state. As the number of firms in the industry grows, this converges to the MPE provided the industry state distribution satisfies a "light tail" condition. This result resembles Hopenhayn (1992) and when the number of firms grows large, with no aggregate shocks, the equilibrium is deterministic.

Introducing this form of incomplete information has some potential drawbacks by implicitly restricting strategic interactions since firms now react to the "average" competitor (i.e. firm A's reaction to a market structure where both competitors B and C are very similar will be the same as when B is very large and C is very small). How well this approximates actual competition in the industry will vary from case to case. It is more likely that the assumption is not valid in oligopolistic industries with large players where strategic interactions are very important. In other industries, competition might be well summarized by the aggregate variables. Some good examples are industries where there is a large number of players, no market

[4] The "curse of dimensionality" is not only a computational problem but will also arise in the estimation. As we will see ahead, since this industry state is very large, if one tries to estimate a flexible policy function on the whole state like proposed by Bajari, Benkard and Levin (2007), it will require a large amount of data (not available on most firm level dataset). The best one can do then is estimate the policies for some aggregation of the state space as implemented in Ryan (2006).

leaders and products are differentiated like, for instance, Industrial Machinery Manufacturing or Metalworking Machinery Manufacturing (moulds, dies, machine tools). What these industries share in common is the fact that each firm sells specialized products, prices are contract specific and information is not publicly available.

Earlier dynamic models only accounted for the effects of entry and exit and did not allowed for investment or R&D (Jovanovic, 1982; Hopenhayn, 1992). Ericson and Pakes (1995) develop an attractive dynamic framework for modeling investment decisions where players use Markovian strategies leading to a Markov Perfect Equilibrium (MPE) as later defined in Maskin and Tirole (2001).

However, solving the MPE brings with it two complications. One was the possibility of non-existence of equilibrium in pure strategies, addressed by Doraszelski and Satterthwaite (2007) with the introduction of privately observed independent and identically distributed shocks. These shocks "smooth out" reaction functions reestablishing the existence of equilibria. The second, is the "curse of dimensionality" and the computational burden attached to solving the model. Recent algorithms (e.g. Pakes and McGuire, 2001) are successful in minimizing this second problem and (depending on the size of the recursive class) can solve the model for up to 20 firms, by using techniques borrowed from the artificial intelligence literature. However, they might not solve problems where there is a larger number of firms in the market. These are exactly the kind of industries we might consider will adapt particularly well to the assumptions introduced here.

Other theoretical models exist that study the R&D decisions in an equilibrium framework. Vives (2004) for example, does this in a static setting, but it does not incorporate any heterogeneity, so that it cannot explain coexistence of R&D and non-R&D firms. Klette and Kortum (2004) use a dynamic framework with the advantage of providing an analytical solution. However, the simplification that allows the elegance of an analytical solution is also the constraint which prevents extensions to account for R&D sunk costs and aggregate uncertainty.

There has been a recent surge in the estimation of dynamic games[5] after the development of estimation methods (Aguirregabiria and Mira, 2007; Bajari Benkard and Levin, 2007; Pakes, Ostrovsky and Berry, 2007; Pesendorfer and Schmidt-Dengler, 2008). The method used is an extension of Hotz et al. (1994) as proposed by Bajari, Benkard and Levin (2007) which allows for both continuous and discrete actions.

Estimation is done in three steps. In the first step productivity is recovered using production function estimation methods. In the second step policy and transition functions are estimated. By assumption, estimated policies are profit maximizing conditional on the equilibrium being played, i.e. the equilibrium observed in the data. Continuation values are then estimated by

[5] A non-exhaustive list includes Aguirregabiria and Ho, 2009; Collard-Wexler, 2006; Hashmin and Van Biesebroeck, 2008; Ryan, 2006; Santos, 2008; Santos and Van Reenen, 2008; Schmidt-Dengler, 2006; Varela-Irimia, 2008; Xu, 2008.

simulating industry paths far enough in the future using the estimated policies and transitions. Non-optimal policies are constructed by slightly perturbing the estimated policy functions and simulating alternative (non-profit maximizing) continuation values. With these optimal and non-optimal continuation values and exploring the property that the value function is linear in the dynamic parameters, the dynamic parameters are estimated by imposing the equilibrium condition, i.e., that optimal continuation values must be larger than non-optimal continuation values. The linearity of the value function in the dynamic parameters greatly simplifies the problem since we do not need to recalculate (re-simulate) continuation values for each set of parameters.

Finally, I evaluate the impact on investment, productivity and market structure of several counterfactuals: a reduction in the sunk costs of R&D, an increase/decrease in market size to assess the impact of trade opening and an increase in entry costs. The results show that a 10% reduction in the sunk cost of R&D results in a 7% increase in average productivity and 50% increase in average capital stock. Furthermore, a decrease in market size to the equivalent of the early 1990's leads to a reduction both in R&D performance and productivity. An increase in entry barriers has a negative effect on exit by less productive firms while virtually no effect on R&D performance therefore leading to a reduction in average productivity.

The rest of the paper is organized as follows. Section 2 introduces the moulds industry, section 3 gives an outline of the model, section 4 provides details of the application, section 5 describes the estimation, section 6 summarizes the data, section 7 contains the results and section 8 the policy experiments, and finally, section 9 concludes.

2 The moulds industry

The Portuguese moulds industry is an interesting case of success. With an almost nonexistent internal market, from its early years the industry developed by exporting almost all production. Currently it exports 90% of its production and supplies 72% of its production to the very competitive car manufacturing industry accounting for more than 1% of total Portuguese Exports (CEFAMOL, 2008). The main advantage is the ability to produce complex moulds which require advanced technology at a low cost and high quality (USITC, 2002).

> "Despite Portugal's small size, it has emerged as one of the world's leading exporters of industrial molds. In 2001, despite limited production of dies, Portugal was the eighth largest producer of dies and molds in the world and it exports to more than 70 countries. The Portuguese TDM industry's success in exporting, and in adoption of the latest computer technologies, has occurred despite the fact that Portugal has a small industrial base on which the TDM industry can depend. **Since joining the EU in 1986, Portugal has focused on serving customers in the common market**." (USITC, 2002)

The 1990's saw a strong sales expansion, mostly to the automotive industry. Over the period 1994-2003 total sales and value added more than doubled (Table B.I). The reason was that joining the EEC in 1986 opened the doors to the European market causing the increase in exports mostly to France, Germany and Spain. This was later reinforced when the Single European Market came into place in 1993 and by the opening of a large car manufacturing plant (joint venture between Ford and VW) which became a large player in the economy. Tables B.II and B.III illustrate the change in export destinations from 1985 to 1990 and the strong export growth in France, Germany and Spain. The industry has stabilized since 2001 (CEFAMOL, 2008).

These changes gave a strong push in the demand for Portuguese mouldmakers and caused a shift (as requested by clients) from the production of very simple to more complex moulds. To achieve this, the development of competencies and technical skills that made the Portuguese industry a regular supplier to big European car-makers like Renault, VW, Mercedes or Saab was necessary.

It was the opening to EU countries (1986 joining the EEC and 1993 the Single Market) that gave these firms sufficient size to introduce innovations which later translated into productivity increases. Some firms used this surge in demand for their products as an opportunity to invest and innovate. This allowed them to maintain and improve their reputation. At the same time, labor productivity increased by roughly 40% over the period. However, this performance was not homogenous across firms. As documented in Table B.IV, R&D firms export more, are on average almost three times as large and 20% more productive (labor productivity).

The mechanism behind this change is interesting in itself because it is what can be called, *trade-induced innovation*. In particular, in the presence of large costs of performing R&D, the access to a larger market (EU) induced by favorable trade agreements, makes it more attractive to incur these costs. There is also strong selection since larger and more productive firms are more likely to develop R&D projects.

Another important industry characteristic is its fragmentation. The number of firms is very large and there are no dominant firms (largest market share is below 10% in any given year). This motivates the use of a monopolistic competition framework where strategic interactions are negligible and the equilibrium effects of entry, investment and pricing can be summarized in the aggregate state.

Notice also that each mould is (quasi) unique, prices depend on the mould specification and are typically contract specific, agreed between the producer and the client. Individual prices are therefore, either unobserved or difficult to compare due to product specificity. Firms tend to specialize in a particular type of mould and therefore potential clients approach firms with the expertise in their product. Portuguese mouldmakers mostly produce moulds for plastics and rubber (very little production of dies or moulds for metal). Within each type of mould, the technology is sufficiently flexible and allows producers to satisfy most needs (depending on the technology they control).

Industry history The history of the industry dates back to the 1930's and 1940's when the development of plastics created a great demand for plastics' moulds. The Portuguese moulds industry started to fill this need in the late 1950's as a major producer of moulds for the glass (where it inherited some of its expertise) and especially for the toy manufacturing industry. In the late 1980's the production started shifting from toy manufacturing towards the growing industries of automobiles and packaging as can be seen by export composition (share of exports by main client/product type) in Figure B.1. During the 1990's the biggest export markets started shifting from the US towards France, Germany and Spain. (IAPMEI, 2006)

During this long period the industry suffered several changes both in terms of the number of firms with a big increase in the early 1980's and a shift towards other main clients due to the boom of the plastics and packaging sectors. This increased pressure for the introduction of new technologies (e.g. CAD, CAM, Complex process, In-mould Assembling) and an increasingly importance of innovation and R&D. For example, current computer operated machines for building moulds use radically different techniques from the ones in the 1970's and 1980's. This state of the art machinery allows flexibility at a low cost alongside a close collaboration with the client in the pre-mould construction phase, which is crucial for car-makers. The design teams can work closely with the clients' engineers and produce 3D virtual versions of the mould which are then programed into the machine to start production. This was in fact a major requirement for car-manufacturers and one of the big advantages of these producers.

Given the industry structure, it is important to take into account the following facts that will be directly addressed:

- High investment rates in physical capital;
- Large growth in sales and productivity;
- Existence of important aggregate industry wide shocks;
- Endogeneity of the R&D start-up decision (larger and more productive firms select to R&D).

3 The aggregate state dynamic model

3.1 States and actions

This section describes the elements of the general model. Time is discrete and in every period, $t = 1, 2, ..., \infty$, there are N firms in the market (N_t incumbents and $N_t^* = N - N_t$ potential entrants) where a firm is denoted by $i \in \{1, ..., N\}$

States Agents are endowed with a continuous state $s_{it} \in \mathfrak{s}_i$[6] and a vector of payoff shocks $\varphi_{it} \in \mathfrak{J}$ both belonging to some compact set. Both the state and the payoff are privately observed by the players. The econometrician observes the states, s_{it}, but not the payoff shocks, φ_{it}.

The industry state is $\mathbf{s}_t = (s_{1t}, ..., s_{Nt}) \in \mathfrak{s}_i^N$. The vector of payoff shocks is independent and identically distributed with distribution F^{φ} and can depend on the actions of the players. This satisfies the conditional independence assumption and allows the value function to be written as a function of the state variables which keeps the number of payoff relevant state variables small (Rust, 1987).

Actions Incumbents choose $l = l^c + l^d$ actions that can be continuous $a_{it}^c \in \mathfrak{A}^c \subset \mathbb{R}^{l^c}$ or discrete (exit, R&D start-up) $a_{it}^d \in \{0,1\}^{l^d}$ and $a_{it} = \{a_{it}^c, a_{it}^d\} \in \mathfrak{A} \subset \mathbb{R}^{l^c} \times \{0,1\}^{l^d}$. For expositional purposes, throughout the rest of the analysis, discrete actions are restricted to be binary and there is only one continuous variable (investment) and one discrete variable (entry/exit). For example, if $a_{it}^d = 1$ represents incumbency and firms decide to exit the industry they set $a_{it}^d = 0$ and collect a "scrap" value, $e + \varphi_i^{scrap}$. Potential (short lived) entrants may choose to pay a privately observed entry cost $(ent + \varphi_i^{entry})$ and enter the industry.

State transition

Assumption 3.1 *(No spillovers) Conditional on current state and actions, own state evolves with transition function*

$$p(s_{it+1}|s_{it}, a_{it})$$

This assumption excludes the cases where the opponents' states or actions directly affect (i.e. not through own actions) the evolution for the state. An example which violates this is knowledge spillovers. This assumption is not necessary but it considerably simplifies the problem. In principle we could allow for $p(s_{it+1}|s_{it}, \mathbf{a}_t, \mathbf{s}_t)$.

Per period payoff Time is discrete and firms collect per period returns which depend on the state of the industry, current actions and shocks $(\pi(a_{it}, \mathbf{s}_t, \varphi_{it}))$ where the period returns are assumed continuous and bounded.

Assumption 3.2 *(a) There exists a function ($S : \mathfrak{s}^N \to \mathfrak{S} \in \mathbb{R}$) that maps the vector of individual states ($\mathbf{s}_t$) into an aggregate index ($S(s_{1t}, s_{2t}, ..., s_{Nt})$). This Aggregate State is observed with noise ($S_t = S(s_{1t}, s_{2t}, ..., s_{Nt}) + \varepsilon_t$, where ε_t is independent and identically distributed with cumulative function F_ε and bounded support).*

(b) Per period returns can be written as

$$\pi(a_{it}, \mathbf{s}_t, \varphi_{it}) = \pi(a_{it}, s_{it}, S_t, \varphi_{it})$$

[6] The model can be extended to discrete states. The focus in the continuous case is to keep notation simple and easy to follow.

Under this assumption, S_t is the payoff relevant variable commonly observed by all agents. The random shock, ε_t, guarantees that there is no perfectly informative state S_t from which agents could recover $(s_{1t}, ..., s_{Nt})$ exactly.[7] This is to prevent degeneracy of beliefs about the current industry state $g(\mathbf{s}_t|S_t, ...S_0)$ (and the resulting equilibrium transition).

Note that the payoff relevant shocks (φ_{it}) do not enter the aggregate index (for example, have no impact on the stage game pricing). One type of demand which meets this assumption is the CES utility where the aggregate industry state is aggregate industry deflated sales.

Assumption 3.3 *(a) Individual states and actions are private information and;*
(b) $g(\mathbf{s}_t|S_t, ..., S_0) = g(\mathbf{s}_t|S_t)$
where $g(\mathbf{s}_t|S_t)$ *is the density function for the industry state,* $\mathbf{s}_t$*, conditional on the aggregate state* S_t*.*

Assumption 3.3 states that the only common information is the aggregate state. Moreover it implies that everything agents can learn about the state of the industry, $\mathbf{s}_t$, is contained in S_t and history $(S_{t-1}, ..., S_0)$ adds no more extra information. This is a fundamental assumption to guarantee a Markovian aggregate state.

The timing is the following:

1. States (s_{it}) and shocks (φ_{it}) are observed by firms;
2. Firms compete in the market and collect period returns $(\pi(.))$;
3. Actions $(\mathbf{a}_t = (a_{1t}, ..., a_{Nt}))$ are chosen simultaneously;
4. New state is formed $(s_{t+1}, S_{t+1}, \varphi_{t+1} \in \mathfrak{s}^N \times \mathfrak{S} \times \mathfrak{J}^N)$;

3.2 Strategies

The aggregate state For each state firms can take actions in some space $a_{it} \in \mathfrak{A}$. Players are restricted to use Symmetric Markovian Pure Strategies.[8] The strategies map the set of states into the action space, $\sigma : \mathfrak{s} \times \mathfrak{S} \times \mathfrak{J} \rightarrow \mathfrak{A}$ $\;(\sigma_{it}(s_{it}, S_t, \varphi_{it}) = (\sigma^c_{it}(s_{it}, S_t, \varphi_{it}), \sigma^d_{it}(s_{it}, S_t, \varphi_{it})))$ where the action space is defined by $\mathfrak{A}(s_{it}, S_t, \varphi_{it}) \subset \mathfrak{s} \times \mathfrak{S} \times \mathfrak{J} \times \mathbb{R}^{l^c} \times \{0, 1\}^{l^d}$. Using symmetry we can drop the i subscript and imposing stationarity we can drop the t subscript:

$$\sigma_{it}(s_{it}, S_t, \varphi_{it}) = \sigma(s_{it}, S_t, \varphi_{it}) \tag{1}$$

[7] The intuition for this error term is the following, imagine s_{it} is marginal cost which affects pricing in the stage game so that the price is a function of the state $p(s_{it}, S_t)$. If players make pricing mistakes, imagine the actual price they set is $p(s_{it}, S_t) + \varepsilon^i_t$, where ε^i_t is independent and identically distributed, the aggregate state (average price) is then $S_t = \frac{1}{N_t}\sum_{i=1}^{N_t} p(s_{it}, S_t) + \varepsilon^i_t = \frac{1}{N_t}\sum_{i=1}^{N_t} p_{it} + \varepsilon_t$, where $\varepsilon_t = \frac{1}{N_t}\sum_{i=1}^{N_t} \varepsilon^i_t$.

[8] Anonymity as defined in Ericson and Pakes is imposed by assuming that firms do not observe each other's state.

Proposition 1 *If players use strategies of the form [1], under assumptions 3.1 to 3.3 the industry aggregate state conditional distribution takes the form* $q^{\sigma}(S_{t+1}|S_t)$.

Proof. See appendix. ■

So while the industry state is a vector $\mathbf{s}_t = (s_{1t}, s_{2t}, ..., s_{Nt})$, S_t is a scalar variable that maps all the industry state into an aggregate state $S_t = S(s_{1t}, s_{2t}, ..., s_{Nt}) + \varepsilon_t$. This result critically depends on the validity of the Assumptions, in particular the restriction on learning in 3.3 (b). As explained above this assumption can be tested and in the empirical section this is done by directly evaluating whether the transition for the aggregate state is a first-order Markov process by testing the significance of previous lags.

When some actions and states are not observed, the firm has to condition its strategies on the expected actions and state of the competitors. When nothing is observed about the competitors, the firm will have the same expectation about the state and actions for all competitors.

To understand the implications of this incomplete information assumption, recall that in the Ericson and Pakes framework with the symmetry and anonymity assumption firms "keep track" of the industry state distribution and not the whole industry state vector as would be the case with no anonymity. This is because under anonymity, the industry state distribution is a sufficient statistic for the industry state vector. In the proposed incomplete information case what matters is just one moment of this same distribution so this imposes slightly stronger conditions than the usual symmetry and anonymity. It implicitly imposes more structure in the type of strategic interactions since firms now react to the "average" competitor (i.e., *ceteris paribus,* firm A's reaction to a market structure where both competitors B and C are very similar will be the same as when B is very large and C is very small provided the aggregate state is the same). Notice that it is implicitly assumed that firms are infinitesimally small with respect to the aggregate state and, knowledge about own state is considered to have no impact on the evolution of the aggregate state conditional on knowing the current state so that $q^{\sigma}(S_{t+1}|s_{it}, S_t) = q^{\sigma}(S_{t+1}|S_t)$. If this was violated, at each point in time a firm would have own individual beliefs about the evolution for the aggregate state. Although this can be allowed it would significantly complicate the structure of the game and remove most of the gains of using the aggregate state model.

Corollary 2 *In the case* $S_t = \sum_{i=1}^{N} h(s_{it}) + \varepsilon_t$ *and under assumptions 3.1 to 3.3, as* N *becomes large* $q^{\sigma}(S_{t+1}|S_t)$ *is approximately normally distributed with conditional mean* $\mu^{\sigma}_{S_{t+1}|S_t} = (1 - \rho^{\sigma}_s)\mu_S + \rho^{\sigma}_s S$ *and standard deviation* $\sigma^{\sigma}_{S_{t+1}|S_t} = \sigma^{\sigma}_S(1 - (\rho^{\sigma}_s)^2)^{1/2}$. *Where* $\mu^{\sigma}_S, (\sigma^{\sigma}_S)^2, \rho^{\sigma}_s$ *are respectively the unconditional mean, variance and autocorrelation for the* S_t *process when players use strategies [1].*

Proof. We can write the distribution $f^{\sigma}(s_{it+1}|S_t) = \int_{\varphi}\int_{s_{it}} p(s_{it+1}|s_{it}, \sigma(s_{it}, S_t, \varphi_{it}))g(ds_{it}|S_t)dF^{\varphi}$ which is independent and identically distributed across $i = 1, ..., N$. The result then follows from the central limit theorem. ■

Corollary 3 *As N becomes large, three moments of the aggregate state distribution, $(\mu_S, \sigma_S, \rho_S)$ fully characterize $q^\sigma(S_{t+1}|S_t)$.*

Proof. Follows directly from Corollary 2. ■

3.3 Equilibrium

The equilibrium concept is Markov Perfect Bayesian Equilibrium in the sense of Maskin and Tirole (1988, 2001). Since strategies are restricted to be Markovian pure strategies the problem can be represented as:

$$V(s_{it}, S_t, \varphi_{it}; q^\sigma) = \sup_{a \in \mathfrak{A}} \pi(s_{it}, S_t, \varphi_{it}, a_{it}) + \beta E\{V(s_{it+1}, S_{t+1}, \varphi_{it+1})|s_{it}, S_t, a_{it}, \varphi_{it}; q^\sigma\} \quad (2)$$

where

$$E\left[V_{i,t+1}|s_{it}, S_t, \varphi_{it}, a_{it}\right] = \int_{s \in \mathfrak{s}, S \in \mathfrak{S}, \varphi \in \mathfrak{J}} V_{i,t+1} \tilde{q}^\sigma(ds_{it+1}, dS_{t+1}, d\varphi_{it+1}|s_{it}, S_t, \varphi_{it}, a_{it})$$

$$\tilde{q}^\sigma(s_{it+1}, S_{t+1}, \varphi_{it+1}|s_{it}, S_t, \varphi_{it}, a_{it}) = q^\sigma(S_{t+1}|S_t) p(s_{it+1}|s_{it}, a_{it}) \phi(\varphi_{it+1})$$

This value function depends on the beliefs about the transition of the aggregate state, $q^\sigma(S_{t+1}|S_t)$. These beliefs depend on the equilibrium strategies played by all players. Notice that since firm i does not observe $s_{jt}, \forall j \neq i$, it can only form an expectation of its rivals' actions conditional on the information available. This has a similar effect to the introduction of private information in Doraszelski and Satterthwaite (2007) which smooths out the continuation value and guarantees existence of equilibria (as if players used mixed strategies).[9]

Definition 4 *A collection of Markovian strategies and beliefs $(\sigma, q^\sigma())$ constitute a Markov perfect equilibrium if:*

(i) Conditional on beliefs about industry evolution (q^σ) firms' strategies ($\sigma_{it} = \sigma^*(s_{it}, S_t, \varphi_{it}; q)$ maximize the value function $V(s_{it}, S_t, \varphi_{it}; q)$.

(ii) The industry transition ($q^*(S_{t+1}|S_t; \sigma^*(s_{it}, S_t, \varphi_{it}|q)$) resulting from optimal behavior (σ^*_{it}) defined above is consistent with beliefs $q^\sigma(S_{t+1}|S_t)$

[9] Doraszelski and Satterthwaite (2007) have shown that in some cases the original Ericson and Pakes framework did not have an equilibrium in pure strategies.

The solution to the dynamic programming problem conditional on q is the optimal strategy $\sigma^*(.|q)$ and a solution exists, under Blackwell's regularity conditions. These strategies will then characterize the industry conditional distribution $q^\sigma(S_{t+1}|S_t;\sigma^*)$ and the equilibrium is the fixed point to a mapping from the beliefs used to obtain the strategies onto this industry state transition

$$\Upsilon(q)(S_{t+1}|S_t) = q^\sigma(S_{t+1}|S_t;\sigma^*(.|q^\sigma))$$

where firm's follow optimal strategies $\sigma^*(.)$. An equilibrium exists when there is a fixed point to the mapping $\Upsilon(q) : \mathfrak{Q} \rightarrow \mathfrak{Q}$

A proof of existence for a similar class of models is provided in Doraszelski and Satterthwaite (2007). The idea explored is that stochastic privately observed shocks "smooth" out reaction functions guaranteeing continuity of own strategies on opponent's actions. This ensures the existence of at least one fixed point (there can be many). For the interested reader in the technical appendix I provide a sketch for the existence proof using Schauder's fixed point theorem. I also discuss uniqueness which is not guaranteed but, there is some reason to believe the problem becomes less severe in the aggregate state model.

3.4 Discussion

Reducing the industry state into an aggregate by introducing incomplete information avoids the "curse of dimensionality". As noted before, this imposes more structure on the type of strategic interactions by making policy functions identical to all industry structures that result in the same aggregate state. In a sense this condition imposes slightly stronger restrictions than the usual anonymity and symmetry assumptions which are also fundamental to reducing the dimensionality of the state space. Symmetry and anonymity are a restriction that allows the state space to be characterized more compactly as a set of "counting measures" (i.e. the industry state distribution).[10]

Krusell and Smith (1998) explore a similar idea whereby the evolution of the aggregate variables in the economy is well approximated by some summary statistics even in the presence of substantial heterogeneity in the population.

Empirical methods like BBL can avoid equilibrium calculation and its computational burden. However, they cannot avoid equilibrium calculations when producing counterfactuals. This is one important reason for using structural econometrics models so that alternative policies can be evaluated in the absence of good experimental data.

[10] Notice that the aggregate state is the payoff relevant variable and individual states are only informationally relevant, i.e., to help forecasting what is the most likely aggregate state in the future.

Assumptions 3.2 and 3.3 might be seen as restrictive in some settings.[11] The first is satisfied by most reduced form profit functions whenever S is payoff relevant. For example, the model is flexible enough to allow different demand structures provided the aggregate state is the payoff relevant variable.

The second assumption is more restrictive as it requires that firms do not observe each other's states (and actions) and more importantly, that history of the aggregate state is irrelevant conditional on the current state. Imagine the case where the state variable is price. It states that firms observe industry aggregate prices (e.g. published by some entity) but they do not observe individual prices for the competitors (e.g. this would involve prohibitive costs in market research). For the moulds industry, each product is individual and therefore prices are product specific. Furthermore, there are many firms in the industry and no significant players. In this sense, the assumption of incomplete information is not a severe restriction.

In industries where there are market leaders, Assumption 3.3 will not hold. A possible extension in this case is to increase the state space to include the state of the market leaders (notice this state is still only informationally relevant if the aggregate state is the payoff relevant variable). There are now two dynamic problems to solve, one for the leader and one for all other firms. State space increases to (s_{it}, S_t, s_{Lt}) where s_{Lt} is the state of the leader. However, it is still to be checked what the equilibrium resulting from players using these strategies looks like. This is not a trivial extension of the work presented here.

Once equilibrium transition, $q^*(S_{t+1}|S_t)$, is known the problem can be represented as a standard dynamic programming problem which can be estimated with available techniques for single agent models (Rust (1987), Hotz and Miller (1993), Aguirregabiria and Mira (2002)) or using estimators developed for dynamic games (Aguirregabiria and Mira, 2007; Bajari, Benkard and Levin, 2007; Pakes, Ostrovsky and Berry, 2007; Pesendorfer and Schmidt-Dengler, 2008).

4 Recovering the Sunk Costs

The framework is the following: firms sell differentiated products and they can invest both in physical capital and decide to engage in R&D for which they have to pay a sunk cost. These sunk costs can range from building an R&D lab to the costs of internal reorganization or even credit constraints. Finally, potential entrants can enter and incumbents can exit. The variables and parametrization will now be defined.

4.1 State and action space

The state space s_{it} for firm i at time t is represented by four variables: Physical capital (K), productivity (ω), R&D status (R, where $R = 1$ denotes that the firm has built the R&D lab and $R = 0$ otherwise) and operating status (χ, where $\chi = 1$ denotes incumbency).

[11] Assumption 3.1 ("no spillover") is standard in the literature and it allows us to write down the transition for the individual state conditional on the firms' actions independently of the other firms' action/states.

$$s_{it} = (K_{it}, \omega_{it}, R_{it}, \chi_{it})$$

where $\omega_{it} \in \Omega$, a compact set on the real line and $K_{it} \in \Xi$, a compact set on the non-negative reals. For the discrete decisions, $R_{it} \in \{0,1\}, \chi_{it} \in \{0,1\}$.

The aggregate state S_t is average deflated industry sales ($\tilde{Y}_t$ is average industry sales and $\tilde{P}_t$ is the average industry price):

$$S_t = \frac{\tilde{Y}_t}{\tilde{P}_t}$$

There are also stochastic shocks (privately observed by the firm and unobserved by the econometrician) including shocks to investment φ_{it}^{inv} , to the sunk cost of R&D φ_{it}^{RD}, and to the scrap value φ_{it}^{scrap}. The vector of payoff shocks $\varphi_{it} = (\varphi^{inv}, \varphi^{RD}, \varphi^{entry}, \varphi^{scrap})$ is an independent and identically distributed standard normal random variable.

After entering the industry, firms can invest in physical capital, pay a sunk cost and engage in R&D and finally decide to exit from the industry. Denote the action space as a, where a superscript denotes either a continuous decision (c) such as investment levels or a discrete decision (d) such as starting an R&D lab or exiting the industry.

$$a_{it} = (a_{it}^c, a_{it}^d) = (I_{it}, R_{it+1}, \chi_{it+1})$$

Investment, $I_{it} \in \Im$ can take any non-negative value and the shape of the profit function guarantees that this is always finite.

The law of motion for the state variables depends on the previous state and actions with density function. This law of motion will be stochastic for productivity and deterministic for all other state variables.

4.2 Parametrization

Using a demand and production function we can solve the static pricing game to get the reduced form profits. This profit function satisfies Rust's (1987) conditional independence and additive separability assumptions

$$\pi(s_{it}, S_t, a_{it}, \varphi_{it}) = \tilde{\pi}(s_{it}, S_t, a_{it}) + \varphi_{it}(a_{it})$$

4.2.1 Demand

Using the Dixit-Stiglitz monopolistic competition framework in discrete varieties, there are N_t available goods, each supplied by a different firm so there are N_t firms in the market. Consumers choose quantities of each good Q_i to consume at a price P_i.

Solving the representative consumer problem, firm's demand is [see the technical appendix]

$$Q_{it} = \frac{\tilde{Y}_t}{\tilde{P}_t}\left(\frac{P_i}{\tilde{P}_t}\right)^{-\eta} \tag{3}$$

Where the price index is $\tilde{P}_t = \left[\frac{1}{N}\left(\sum_{i=1}^{N_t} P_{it}^{-(\eta-1)}\right)\right]^{-1/(\eta-1)}$ and $\left(\frac{\tilde{Y}}{\tilde{P}}\right) = \frac{\frac{1}{N}\sum_{i=1}^{N_t} P_{it}Q_{it}}{\tilde{P}}$ is average industry deflated revenues. If the goods were perfect substitutes (η is infinite), there could be no variations in adjusted prices across firms, $P_i = \tilde{P}$ and $\frac{\tilde{Y}}{\tilde{P}} = Q_i$ for all firms. Notice that the representative consumer buys a fraction of each of the available goods. While this could be relaxed using an alternative demand function, the benefits of it would only arise if there was individual price and quantity data available so that more flexible elasticities of substitution could be estimated.

4.2.2 Production function

The production technology is assumed to be Cobb-Douglas where L is labor input:

$$Q_i = e^{\omega_i} L_i^{\alpha_l} K_i^{\alpha_k} \tag{4}$$

4.2.3 Static pricing game

Firms compete in the market by setting prices simultaneously in a static demand framework. Since in the short run the only flexible input into production is labor and gross profits are $\tilde{\pi} = [P(Q_{it})Q_{it} - wL_{it}]$ (w is the wage rate), the resulting solution to the short run static pricing game is: [12]

$$\tilde{\pi}(\omega_{it}, K_{it}, S_t; \eta, \alpha_k) = \frac{1}{\gamma}\left(\frac{\eta-1}{\eta}\right)\left(\frac{\eta-1}{\eta}\frac{\alpha_l}{w/\tilde{P}}\right)^{\alpha_l\gamma} S_t^{\gamma/(\eta-1)}\left(e^{\omega_{it}}K_{it}^{\alpha_k}\right)^{\gamma} \tag{5}$$

where $\gamma = (\eta-1)/(\eta-\alpha_l(\eta-1))$, $\tilde{\pi}$ is gross profit and

$$S_t = \frac{\tilde{Y}_t}{\tilde{P}_t} \propto \frac{\tilde{Y}_t}{\frac{1}{N}\sum_j \left[\omega_{jt}K_{jt}^{\alpha_k}\right]^{\gamma}}$$

Productivity Productivity evolves stochastically with a different transition for R&D and non-R&D firms. In general, product and process innovation are difficult to disentangle from each other unless one observes firm level price data (e.g. Foster, Haltiwanger and Syverson, 2008). Since the dataset for the moulds industry does not contain firm level price data, they are considered to be indistinguishable in the model and productivity is broadly defined.[13]

[12] See derivation in the technical appendix.

[13] The model can however be extended to allow for quality in the demand specification (see Melitz, 2000). This distinction would be important to model other type of phenomena like dynamic pricing, where for example the effects of product and process innovation would be qualitatively different.

The 'internal' source of uncertainty distinguishes R&D investment from other decisions as capital investment, labor hiring, entry and exit which have deterministic outcomes and where the only source of uncertainty is 'external' to the company (e.g. due to the environment, to competition, to demand, etc.). This distinction is important since the stochastic R&D outcome will determine (together with entry and exit) the stochastic nature of the equilibrium.

Productivity is assumed to follow a controlled first order Markov process.

$$\omega_{it+1} = E(\omega_{it+1}|\omega_{it}, R_{it}) + \nu_{it}$$

where v_{it} is independently and identically distributed across firms and time.

4.2.4 Cost function

Investment cost Investment adjustment costs are quadratic (Hayashi 1982) and totally irreversible with the following parametrization:

$$C^K(I_t, K_{it-1}) = \left[\mu_1 I_{it} + \mu_2 \frac{I_{it}^2}{K_{it-1}}\right] + \varphi_{it}^I I_{it} \quad \text{if} \quad I_{it} > 0 \tag{6}$$

where $\mu_2 > 0$ indexes the degree of convexity and the 'price' of investment is $\mu_1 +$ $\varphi_{it}^I > 0$.

R&D technology The firm has the choice of building an R&D lab at a sunk cost of $\lambda +$ φ_{it}^R where φ_{it}^R is a standard normal random variable. As mentioned above the continuous R&D decision of how much to spend each period is not directly modeled. This is done mainly for simplicity since otherwise there would be an extra policy function to consider. The empirical results from the literature suggest that R&D intensity (R&D to sales ratio) is highly autocorrelated. For example, Klette and Kortum (2004) take this as a stylized fact that they fit with their model. This simplification could lead to overestimating the sunk costs of R&D because the estimates do not take into account the period by period expenditures and would therefore overestimate the benefits of R&D. The alternative followed in this paper is to estimate the static profit function using observed profits and allow for a fixed cost (or benefit) for R&D firms. In this way, any R&D costs which are not incorporated in physical capital would in fact be accounted for (see equation 11). Furthermore, physical capital decisions are directly modeled so that any physical capital costs incurred due to R&D period by period expenditure, would "show up" in the policy function for investment. The same applies for R&D labor or materials expenditures (for accounting reasons some R&D costs can show up as normal labor or physical capital in the data). For these two reasons, abstracting from the continuous R&D decision should not cause a severe bias in the estimate of the sunk costs.

Entry cost Potential entrants, denoted by $\chi_{it} = 0$, are short lived and cannot delay entry. Upon entry, firms must pay a (privately observed) sunk entry fee of $Ent + \varphi_{it}^{entry}$ to get a draw of (ω, K) with distribution $p(\omega_{t+1}, K_{t+1}|\chi_t = 0)$. Since entry effects are captured in the equilibrium transition for the aggregate state, entry does not need to be modeled for estimation purposes. However, to produce counterfactuals, we need to recalculate equilibrium transitions for the aggregate state and the entry process will then be explicitly modeled.

Exit value Every period the firm has the option of exiting the industry and collecting a scrap exit value of $e + \varphi_{it}^{scrap}$.

Payoff shocks The vector of payoff shocks $\varphi = (\varphi^{inv}, \varphi^{RD}, \varphi^{entry}, \varphi^{scrap})$ are independent and identically distributed standard normal.

4.2.5 State transition

As explained above productivity follows a controlled Markov process. The capital stock depreciates at rate δ and investment adds to the stock:

$$K_{i,t+1} = (1-\delta)K_{it} + I_{it}$$

If a firm decides to start R&D, the sunk cost is paid only once at start-up:

$$R_{i,t+1} = \begin{cases} 1 & \text{if } R_{it} = 1 \text{ or firm pays sunk cost} \\ 0 & \text{otherwise} \end{cases}$$

If a firm exits it sets $\chi_{i,t+1} = 0$ and if it enters it sets $\chi_{i,t+1} = 1$

$$\chi_{i,t+1} = \begin{cases} 1 & \begin{array}{c} \text{if } \chi_{it} = 0 \text{ and firm } i \text{ enters OR,} \\ \chi_{it} = 1 \text{ and firm } i \text{ stays} \end{array} \\ 0 & \begin{array}{c} \text{if } \chi_{it} = 0 \text{ and firm } i \text{ does not enter OR,} \\ \chi_{it} = 1 \text{ and firm } i \text{ exits} \end{array} \end{cases}$$

4.2.6 Period Returns

Using the specified cost structure the per period return function of an incumbent is

$$\pi(\omega_{it}, K_{it}, R_{it}, R_{it+1}, \chi_{it+1}, I_{it}, S_t) =$$
$$= \left\{ \begin{array}{c} \tilde{\pi}(\omega_{it}, K_{it}, S_t) - \mu_1 I_{it} - \mu_2 \frac{I_{it}^2}{K_{it-1}} - \varphi_{it}^{inv} I_{it} \\ -(\lambda + \varphi_{it}^{RD})(R_{it+1} - R_{it})R_{it+1} + (1-\chi_{it+1})(e + \varphi_{it}^{scrap}) \end{array} \right\}$$

The demand system specified in equation [3] illustrates the two aggregate variables which affect a company's revenues. One is market size $(\tilde{Y})$ which evolves exogenously and the other is average industry price ($\tilde{P}$) which is determined endogenously. Since individual prices are determined in the static pricing game by productivity and physical capital ($P_i^* = P(\omega_i, K_i, \tilde{P}, \tilde{Y})$, see the technical appendix), the price index is a mapping from individual firms' productivity and capital stock onto a pricing strategy so that the aggregate state variable is

$$S_t = \frac{\tilde{Y}_t}{\tilde{P}_t(\omega, K, R, \tilde{Y})}$$

5 The estimation procedure

There are several alternatives to estimate dynamic games (see for example Pesendorfer and Schmidt-Dengler, 2008; Aguirregabiria and Mira, 2007; Bajari, Benkard and Levin, 2007; and Pakes, Ostrovsky and Berry, 2007). The method used in this paper follows Bajari, Benkard and Levin (2007) since this allows for both discrete and continuous choices and is easily applicable to the model outlined above. This estimation procedure has been applied by Ryan (2006) to study the impact of environmental regulation changes on capacity investment for the cement industry in the US. Ryan (2006) also considers Markovian strategies on individual states and an aggregate state. Since he models an investment capacity game, the industry state is the sum of competitors' capacities. The difference from the framework proposed here is that industry state aggregation is only used for estimation of the policy functions and not in calculating equilibrium transitions, i.e. the equilibrium transition is still the individual state by state transition.

The estimation proceeds in three steps. In the first step, unobserved productivity (ω_{it}) is recovered by estimating a production function. A number of ways for estimating the production function are considered (Olley and Pakes, 1996; Levinsohn and Petrin, 2002; Ackerberg et al, 2007, and Bond and Soderbom, 2005). The different methods give broadly similar estimates.[14] The second step recovers the profit function ($\tilde{\pi}(\omega_{it}, K_{it}, R_{it}, S_t)$), the firm-level and industry-level state transitions, ($p(\omega_{it+1}|\omega_{it}, R_{it}, \chi_{it})$ and $q(S_{t+1}|S_t)$) as well as the equilibrium policy functions for investment, R&D and exit. Finally, in the third step, the dynamic parameters (μ_1, μ_2, λ, e) are estimated using the equilibrium conditions.

5.1 Step 1: Productivity

Productivity is not directly observed and there are methods [15] to estimate it as the residual from a production function (Olley and Pakes, 1996; Levinshon and Petrin, 2003; De Loecker, 2007). De Loecker (2007) proposes an estimator for both the production function parameters and the demand elasticity under imperfect competition when one uses deflated sales instead of quantities (see also Klette and Griliches, 1995). This method, however, cannot be directly applied to the

[14] See a companion paper (Santos, 2008) which compares several alternative production function methods to recover productivity using the same dataset.

[15] Ackerberg et al. (2007) provide a survey on the literature for estimating production functions.

model outlined above.[16] The reason for this arises from the fact that input demand function depends on the industry state, more precisely in our case, the aggregate industry state. To see this, notice that for example investment functions depend on equilibrium beliefs about industry evolution. Specifying the investment policy as originally proposed by Olley and Pakes (1996), $i(\omega)$, is a misspecification since the equilibrium policy for investment in an Ericson and Pakes (1995) style game depends on the state of all players in the industry $i^*(\omega_1, .., \omega_N)$.

Applying this to our case, elasticity of demand cannot be recovered in the first step since the input demand is also a function of aggregate sales.[17] Taking the log of total sales and using equations (3) and (4):

$$y_{it} - \tilde{p}_t = q_{it} + p_{it} - \tilde{p}_t = \frac{1}{\eta}(\tilde{y}_t - \tilde{p}_t) + \frac{\eta - 1}{\eta}(\omega_{it} + \alpha_k k_{it} + \alpha_l l_{it}) \tag{7}$$

Using materials to control for the unobservable as in Levinsohn and Petrin (2003), the input demand is a function of the state at time t (individual and industry states)

$$m_{it} = m(\omega_{it}, k_{it}, R_{it}, S_t) \tag{8}$$

Assuming invertibility this can be expressed as[18]

$$\omega_{it} = \omega(k_{it}, R_{it}, S_t, m_{it}) \tag{9}$$

and the unobservable is now a function of observables. Note however that since productivity is also a function of market conditions ($S_t = \frac{\tilde{Y}_t}{\tilde{P}_t}$) in eq. (8), demand elasticity (η) cannot be recovered in the first stage, because it enters nonlinearly in the control function (9). This is the main difference from a single agent framework as in De Loecker (2007) where input demand depends solely on individual state variables ($m_{it} = m(\omega_{it}, k_{it}, R_{it})$).

Using the controlled first order Markov process assumption for productivity

$$\omega_{it} = E[\omega_{it} | \omega_{it-1}, R_{it-1}] + \nu_{it}$$

where ν_{it} is an independent and identically distributed random shock to productivity and is assumed to be additively separable.

[16] This applies to most literature on production function methods following approaches similar to Olley and Pakes (1996) which look at single agent problems and disregard the possibility that policy functions (investment or materials) are equilibrium solutions to dynamic games. Therefore, the policy functions will be misspecified and the productivity control function will not include some important industry level variables leading to bias in the parameter estimates due to unobserved heterogeneity.

[17] There is also the selection problem due to exit as explained in Olley and Pakes (1996). In this paper I will abstract from this problem since the number of exits observed in the data is very small and the correction does not affect my estimates. However, if there were more observations on exits, this problem could be more carefully addressed.

[18] A slight concern with invertibility and imperfect competition is the fact that with imperfect competition an increase in productivity might not lead to a direct increase in output and therefore in materials usage. For the demand system specified, an increase in productivity is equivalent to a decrease in costs and it translates directly into a decrease in prices (see the technical appendix). This means total output goes up and therefore also does materials usage.

Stage I From eq. (7) rewrite the production function using deflated sales as variables $y_{it}^p = y_{it} - \tilde{p}_t$ and $\tilde{y}_t^p = \tilde{y}_t - \tilde{p}_t = \ln(S_t)$

$$y_{it}^p = \frac{\eta-1}{\eta}\alpha_l l_{it} + \phi(k_{it}, R_{it}, \tilde{y}_t^p, m_{it}) + \varepsilon_{it}^y$$

where ε_{it}^y is measurement error in y_{it}^p and

$$\phi(k_{it}, R_{it}, \tilde{y}_t^p, m_{it}) = \frac{1}{\eta}\tilde{y}_t^p + \frac{\eta-1}{\eta}\alpha_k k_{it} + \frac{\eta-1}{\eta}\omega(k_{it}, R_{it}, \tilde{y}_t^p, m_{it})$$

This can be estimated non-parametrically or using an nth-order polynomial approximation. This provides estimates of $\widehat{\frac{\eta-1}{\eta}\alpha_l}$ and $\hat{\phi}$.

Stage II For the second stage use the estimated values to construct$\hat{\phi}_{it} = \hat{y}_{it}^p - \widehat{\frac{\eta-1}{\eta}\alpha}_l l_{it}$ and with this get an estimate of $\frac{\eta-1}{\eta}\omega_{it}$ for a given $\widetilde{\frac{\eta-1}{\eta}\alpha_k}$ and $\widetilde{\frac{1}{\eta}}$

$$\widehat{\frac{\eta-1}{\eta}}\omega_{it} = \hat{\phi}_{it} - \widetilde{\frac{1}{\eta}}\tilde{y}_t^p - \widetilde{\frac{\eta-1}{\eta}}\alpha_k k_{it}$$

Approximate non-parametrically $E[\omega_{it}|\omega_{it-1}, R_{it-1}]$. Several approximations are used in the empirical section and in practice a cubic polynomial fits the data well for the productivity transition without large unreasonable behavior at the tails of the observations[19]

$$\begin{aligned}
\hat{y}_{it}^p - \widehat{\frac{\eta-1}{\eta}}\alpha_l l_{it} \quad = \quad & \frac{1}{\eta}\tilde{y}_t^p + \frac{\eta-1}{\eta}\alpha_k k_{it} + \\
& + \left[\begin{array}{l} \gamma_0^0 + \gamma_1^0\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right) \\ +\gamma_2^0\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right)^2 \\ +\gamma_3^0\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right)^3 \end{array}\right] \times 1\left[R_{it-1} = 0\right] \\
& + \left[\begin{array}{l} \gamma_0^1 + \gamma_1^1\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right) \\ +\gamma_2^1\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right)^2 \\ +\gamma_3^1\left(-\hat{\phi}_{it-1} + \frac{1}{\eta}\tilde{y}_{t-1}^p + \frac{\eta-1}{\eta}\alpha_k k_{it-1}\right)^3 \end{array}\right] \times 1\left[R_{it-1} = 1\right] \\
& + \nu_{it}
\end{aligned} \tag{10}$$

Finally, use eq. (10) to estimate $\frac{1}{\eta}$ and α_k by nonlinear least squares.

[19] Instead of using a cubic polynomial, also reported are the results with a sigmoidal function which preserves monotonicity: $E(\omega_{it}|\omega_{it-1}) = \frac{\gamma_0^0}{(1+\gamma_1^0\exp(-\omega_{it-1}))}$

Robustness In the second stage, the error term in equation (10), ν_{it}, must be uncorrelated with k_{it} and $\tilde{y}_t^p$. While this might be a reasonable assumption for k_{it} due to the timing of investment that makes k_{it} independent from "news" at period t, the same is not necessarily true for $\tilde{y}_t^p$ if there is an aggregate time component ν_t in the productivity shock ν_{it} . One potential instrument is the use of lagged $\tilde{y}_{t-1}^p$.

The potential multicollinearity problem between l_{it} and $(k_{it}, R_{it}, \tilde{y}_t^p, m_{it})$ as mentioned by Ackerberg, Caves and Frazer (2006) is acknowledged by estimating the production function and recovering the labor coefficient in the second step. An alternative to address this multicollinearity problem is to use the method proposed by Bond and Soderbom (2005).

Finally, there can be sample selection due to exit. The selection problem arises if smaller firms are more likely to exit upon a negative shock generating a negative correlation between productivity and capital stock for the firms which remain in the industry. This is likely to be relevant in industries with severe exit behavior, but less likely to be true for industries with little exit.

Most of these problems are addressed in a companion paper where several methods for estimating production functions are compared using data for the moulds industry (Santos, 2008).

5.2 Step 2: Policies and transitions

5.2.1 Static Profits

There are two options to parametrize the profit function. First we can use the estimated parameters of the production function $(\alpha_l, \alpha_k, \sigma)$ and use the parametric form eq. (5). Alternatively, since profits are reported in accounting data, we can directly estimate the parameters of the profit function. Since reported accounting profits can be negative, this is only possible if there are fixed production costs. To model the existence of fixed costs (which might depend on firm size) we can specify the profit function as

$$\tilde{\pi}_{it} = \alpha_0^\pi e^{\alpha_1^\pi \omega_{it}} K_{it}^{\alpha_2^\pi} S_t^{\alpha_3^\pi} + \alpha_4^\pi + \alpha_5^\pi K_{it} + \alpha_6^\pi R_{it} + \xi_{it} \tag{11}$$

where ξ_{it} is measurement error in observed profits, $\tilde{\pi}_{it}$. This can be estimated by nonlinear least squares.

5.2.2 Policies

Using the observations for all state variables (ω, K, S, R), the policy functions can be easily estimated. The investment function which results as the solution to the dynamic problem is

$$I_{it} = \frac{1}{2\mu_2}\left(\beta \frac{\partial E(V(s_{it+1}, S_{t+1}|s_{it}, S_t))}{\partial I_{it}} - \mu_1\right) - \frac{1}{2\mu_2}\varphi_{it}^{inv} \tag{12}$$

which can be estimated separately for R&D and non-R&D firms as

$$I_{it} = P^{n,i}(\omega_{it}, K_{it}, S_t, R_{it}) + \tilde{\varphi}_{it}^{inv} \tag{13}$$

where $P^{n,i}(.)$ is a n^{th}-order polynomial. After trying with several polynomial degrees, the evidence favors polynomials with smaller degrees because they produce policy functions which are more robust at the tails of the observations. The reason is because even if higher order polynomials produce a better fit in the range of data with more observations, it might create large distortions outside this interval of data by imposing highly nonlinear and unreasonable functional forms, similar to Runge's phenomenon. Since the intervals with less observations are normally in the extremes (tails), this can create large distortions in the estimates at the extremes which are points that can significantly drive the average results. Since errors in the policy functions enter nonlinearly in the second step, this can significantly bias the estimates in small samples. All results have been checked to avoid this by looking at the predictions from the policy functions.

The R&D equation is estimated with a probit model where firms will decide to start doing R&D if the costs $(\lambda + \varphi_{it}^{R})$ are smaller than the benefits $\beta[E(V_{it+1}|R_{it+1} = 1) - E(V_{it+1}|R_{it+1} = 0)]$ and the probability that the firm starts performing R&D is

$$\begin{aligned} \Pr(R_{it+1} &= 1|R_{it} = 0, s_{it}, S_t) = \\ & \Phi\left(-\lambda + \beta \left[\begin{array}{c} E\{V(s_{it+1}, S_{t+1})|R_{it+1} = 1\} \\ -E\{V(s_{it+1}, S_{t+1})|R_{it+1} = 0\} \end{array} \right]\right) \end{aligned} \tag{14}$$

since φ_{it}^{RD} is assumed to be a standard normal random variable. This can be approximated by

$$\Pr(R_{it+1} = 1|R_{it} = 0) = \Phi\left(P^{n,rd}(\omega_{it}, K_{it}, S_t, R_{it} = 0)\right) \tag{15}$$

where again $P^{n,rd}(.)$ is an n^{th}order polynomial (the same argument in favor of lower degree polynomials is in place here).

The exit function can be treated in a similar fashion resulting in

$$\Pr(\chi_{it+1} = 0|\chi_{it} = 1) = \Phi\left(P^{n,\chi}(\omega_{it}, K_{it}, S_t, R_{it})\right)$$

5.2.3 The transition function

Aggregate state From Corollary 3 the observed aggregate state has a conditional normal distribution with mean $\mu_{S_{t+1}|S_t} = (1-\rho_S)\mu_S + \rho_S S$ and variance $\sigma_{S_{t+1}|S_t} = \sigma_S(1-\rho_S^2)^{1/2}$. Where $(\mu_S, \sigma_S, \rho_S)$ are respectively the unconditional mean, variance and autocorrelation for the S process and are easily estimated using the sample moments. Alternatively, $q(S'|S)$ can be estimated non-parametrically.

Productivity The transition for individual productivity is estimated separately for R&D and non-R&D firms using a cubic polynomial on lagged productivity $(g^{RD}(\omega_{i,t-1}), g^{NRD}(\omega_{i,t-1}))$.

$$\begin{aligned}
\omega_{i,t+1} &= E(\omega_{i,t+1}|\omega_{it}, R_{it}) + \nu_{it+1} \\
&= (\alpha_0^{\omega,R} + \alpha_1^{\omega,R}\omega_{it-1} + \alpha_2^{\omega,R}\omega_{it-1}^2 + \alpha_3^{\omega,R}\omega_{it-1}^3)R_{it} \\
&\quad +(\alpha_0^{\omega,NR} + \alpha_1^{\omega,NR}\omega_{it-1} + \alpha_2^{\omega,NR}\omega_{it-1}^2 + \alpha_3^{\omega,NR}\omega_{it-1}^3)(1-R_{it}) + \nu_{it+1}
\end{aligned} \tag{16}$$

Alternative functional forms (lower order polynomials) are also reported.

5.3 Step 3: Minimum distance estimator

Once the policy (investment, R&D and exit) and transition functions (productivity and aggregate state) have been recovered, we can proceed as follows:

1. Starting from some state (s_{it}, S_t) at $t = 0$, draw a random vector of payoff shocks $\varphi_{it} = (\varphi_{it}^{inv}, \varphi_{it}^{RD}, \varphi_{it}^{scrap})$; transition function shocks to productivity (ν_{it}) and to the aggregate state (ε_t). Use n_s different starting values so that $(s_{it}, S_t) = [(s_{1,it}, S_{1,t}), ..., (s_{n_s,it}, S_{n_s,t})]$ which can be equal to the states for each observation in the dataset;

2. Simulate actions (a_{it}) by reading off the estimated policy functions and using the payoff shocks;

3. Update states (s_{it+1}, S_{t+1}) by reading off the transition functions and using transition shocks $(\nu_{it}, \varepsilon_t)$;

4. Repeat 2-3 for several periods (each path simulated for $\bar{T}$ periods), and construct a sequence of actions and states $\{a_{it}(s_{i0}, S_0), s_{it}(s_{i0}, S_0), S_t(S_{i0})\}_{t=1}^{\bar{T}}$ from each of the n_s starting configurations;

5. Using the sequence of actions and states, compute the discounted stream of profits for a given parameter vector θ:
$\sum_{t=0}^{\bar{T}} \beta^t \pi(a_{it}, s_{it}, S_t, \varphi_{it}; \tilde{\alpha}, \widetilde{P}^n, \theta)$;[20]

6. Repeat steps 1-5 n_J times to produce an average estimate at each of the n_s states. This gives an estimate of the expected value from a starting configuration, $(s_{it}, S_t)_{t=0}$:

$$\widehat{EV}(s_{i0}, S_0; \tilde{\alpha}, \theta) = \frac{1}{n_J}\sum_{j=1}^{n_J}\sum_{t=0}^{\bar{T}} \beta^t \pi(a_{it}^j, s_{it}^j, S_t^j, \varphi_{it}^j; \tilde{\alpha}, \widetilde{P}^n, \theta)$$

[20] The discount factor is set at $\beta = 0.96$.

In order for a strategy, σ, to be an equilibrium, at equilibrium beliefs, $q^*(.)$, for all $\sigma' \neq \sigma$ the following condition holds

$$V(s_{i0}, S_0; \sigma, q^*(S_{t+1}|S_t); \theta) \geq V(s_{i0}, S_0; \sigma', q^*(S_{t+1}|S_t); \theta)$$

Given the linearity of the value function in the dynamic parameters this can be written as

$$V(s_{i0}, S_0; \sigma, q^*(S_{t+1}|S_t); \theta) = W(s_{i0}, S_0; \sigma, q^*(S_{t+1}|S_t)) * \theta$$

where $W(s_{i0}, S_0; \sigma, q^*(S_{t+1}|S_t)) = E_{\sigma|s_{i0},S_0} \sum_{s=0}^{\infty} \beta^s w_{is}$ and $\theta = [1, \mu_1, \mu_2, \lambda, e]$, $w_{is} = \left[\widetilde{\pi}(s_{is}, S_s; \eta), I_{is}, I_{is}^2, 1(R_{is+1} = 1, R_{is} = 0), 1(\chi_{is+1} = 0, \chi_{is} = 1)\right]$;

7. Construct alternative investment, R&D and exit policies (σ'), for example, by drawing a mean-zero normal error and adding it to the estimated second step policies. With these non-optimal policies, construct alternative expected values following steps 1-6 to get $W(s_0, S_0; \sigma', q^*(.))$. Do this for n_σ alternative policies;

8. Finally, compute the differences between the optimal and non-optimal value functions for several (X_k) policies and states ($X_k, k = 1, ...n_I$), where X_k is a given pair of (σ', s_{i0}, S_0) so that we have $n_I = n_\sigma * n_s$ of them;

$$\hat{g}(X_k; \theta, \tilde{\alpha}, \widetilde{P}^n) = \left[\hat{W}(s_{i0}, S_0; \hat{\sigma}, \hat{q}(S_{t+1}|S_t)) - \hat{W}(s_{i0}, S_0; \hat{\sigma}', \hat{q}(S_{t+1}|S_t))\right] * \theta$$

Since the estimated policies should be optimal, the expected value when using σ cannot be smaller than using alternative σ'. The empirical minimum difference estimator minimizes[21] squared equilibrium condition violations, $g(X_k, \theta, \tilde{\alpha}, \widetilde{P}^n) < 0$

$$\hat{J}(\theta; \tilde{\alpha}) = \frac{1}{n_I} \sum_{k=1}^{n_I} \left(\min\left\{\hat{g}(X_k; \theta, \tilde{\alpha}, \widetilde{P}^n), 0\right\}\right)^2$$

and

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \frac{1}{n_I} \sum_{k=1}^{n_I} \left(\min\left\{g(X_k; \theta, \tilde{\alpha}, \widetilde{P}^n), 0\right\}\right)^2$$

[21] When the objective functions lacks smoothness (e.g. problems with discontinuous, non-differentiable, or stochastic objective functions) using derivative based methods might produce inaccurate solutions. Using derivative free methods (for example, Nelder-Mead) to minimize the empirical minimum distance ($\hat{J}$) helps to circumvent these problems. Non-smoothness might occur with finite n_I, because of the min operator in the empirical objective function, $\hat{J}$, which takes only the negative values of $g(.)$ and this creates discontinuities even if $g()$ is continuous in θ. All results are not affected by the optimization method used.

These methods are known to be slow and sometimes innacurate. Speed is not a problem because the computational burden rests in the simulations and not in the optimization. To deal with possible innacuracy, an option is to rescale the parameters so that they all lie within a small interval (e.g. $[-1, 1]$) and restart the optimization algorithm at different starting values to check the answer is correct.

The time of each path is set at $\bar{T} = 75$, the number of starting configurations $n_s = 1,017$ which is the total number of observations, the number of simulations for each configuration $n_J = 200$ and the number of alternative policies $n_\sigma = 125$, so that the total number of differences is $n_I = 127,125$.

A note on alternative policies and set identification The vector of dynamic parameters θ, must rationalize the observed strategy profile, σ. In general θ can be point or set identified depending on the model, data available and alternative policies. Bajari et al. (2007) also propose a method for (bounds) set identification on θ.

The objective function, $J(\theta)$, depends on the alternative policies used, σ', so that this is a crucial step for identification. The inequality in $J(\theta)$ arises exactly from comparing σ and σ'. If we use policies very "far" from σ, the identified set should increase and if we use policies very "close" to σ, the identified set should shrink. Since we can produce as many alternative policies as desired, the (set) identification of the model can be improved by artificially generating non-optimal data and wisely choosing the alternative policies. For this reason I will focus on the point identified case. One particular issue to address is how σ' is constructed. One option is to slightly perturb the estimated policy function by adding some error. If the error is one sided (for example positive) this means that all perturbed actions will be larger. Then, the parameters are likely to be only set identified (one sided set). For example, if perturbed R&D start-up decisions are more frequent (i.e. only positive errors added to the optimal policy function), the alternative policies will generate high levels of R&D behavior. Since these are much higher than the ones actually observed in the data, the sunk costs of R&D which can be rationalized have to be above a certain level (bounded below), otherwise firms in the data must have performed more R&D. However, the sunk costs are not going to be bounded above because only positive errors have been added. For this reason the choice of alternative policies has an impact on identification so should be done carefully by adding, in this case, both positive and negative errors to the estimated policies.

In the choice of σ' if the alternative policies are not binding (i.e. they are chosen to be very loose meaning the distribution from which errors are draw has a very large standard deviation) again the parameters are less likely to be point identified. As the "looseness" of the alternative policies increases so should the identified set. However, if the alternative policies are very tight (very close to the estimated ones with a very small standard deviation of the errors for the alternative policies) it is going to be much harder to rationalize the observed actions and we will have more violations of equilibrium behavior (i.e. observed actions that cannot be rationalized by any given θ). If the error in the estimated policies is large, this could affect identification since we might be creating a more serious bias in θ by forcing a non-optimal (estimated) strategy to be optimal. Again a careful choice of the standard deviation for the errors affects identification. In the application below, these are set as standard normal errors for investment and normal errors with mean zero and standard deviation 0.5 for R&D and exit. These translate into the fact that 95% of the alternative policies for investment are in the interval ±200% from the optimal one.

5.3.1 Standard errors

Standard errors are estimated using sub-sampling or the bootstrap. An important remark is that only simulation error is produced in step 3. This error disappears as $n_J \to \infty$ for a given set of alternative policies. In practice, since bootstrapping requires very intense computations, the bootstrapped standard errors can overestimate actual standard deviations since they may still contain some simulation error.

5.4 Identification

5.4.1 Technical

The identification problem in dynamic models is well known (Rust, 1994; Magnac and Thesmar, 2002; Pesendorfer and Schmidt Dengler, 2008; and Bajari et al., 2008). Assuming agent's optimal dynamic behavior, imposes no testable predictions. Without further restrictions, a given reduced form (observed) model can be rationalized by more than one parametric form for the structural model.

The structural objects to be identified are the period returns, distribution for the shocks, state transition function and discount factor: $(\pi(a_{it}, \mathbf{s}_t, \varphi_{it}), F(\varphi_{it}), p(s_{it+1}|s_{it}, a_{it}), \beta)$. These primitives are in general non-parametrically unidentified (Rust, 1994), i.e., there is an infinite amount of primitives that can rationalize observed decisions so that different models are observationally equivalent. Magnac and Thesmar (2002) provide some conditions for the identification in single agent models that is extended to dynamic games by Pesendorfer and Schmidt Dengler (2008). The solution for the unidentification result is to use exclusion restrictions (i.e. state variables that do not enter period returns) and/or normalization of the period returns for some outside alternative. Even in these cases the discount factor and distribution of costs shocks are still non-parametrically unidentified unless further restrictions are introduced.

Alternative solutions are when returns are observed and the return function can be estimated non-parametrically or to use parametric restrictions. In our case part of the return function ($\widetilde{\pi}(.)$) is identified in the second step because profits are observed. The evolution for the states $p(s_{it+1}|s_{it}, a_{it})$ is also estimated in the second step (assuming agents have rational expectations this allows us to recover their beliefs which coincide with the actual evolution for the state variables) while the distribution of cost shocks $F(\varphi_{it})$ is assumed to be normal and the discount factor, β, is set exogenously. The only object left to estimate is the cost function which is non-parametrically identified. To understand why, notice that in the model the return function is $\pi(a_{it}, s_{it}, S_t, \varphi_{it}) = \widetilde{\pi}(s_{it}, S_t) - C(s_{it}, a_{it}) + \varphi_{it}(a_{it})$. $\widetilde{\pi}(s_{it}, S_t)$ is estimated using observed data while the cost function takes the value zero when there is inaction (no investment and no R&D start-up), $C(s_{it}, a_{it} = 0) = 0$. This satisfies the exclusion restriction and normalization assumptions in Pesendorfer and Schmidt Dengler (2008).[22]

Notice two drawbacks. First the error distribution ($F(\varphi)$) has to be specified parametrically and there cannot be serially correlated fixed effects. Second, these identification results are asymptotic. However, in applied work, identification can sometimes be harmed by the availability of data in finite samples. This can be due to the fact that there is no sufficient variation in the data. In this case the parameters might be weakly identified.

5.4.2 Empirical

I will discuss now the particular features of the data that identify the parameters of interest (cost function). The second step is probably the most important part of the estimation. It recovers the profits that firms expect to earn at each state (gross of adjustment costs), how they make their decisions and also how they expect states to evolve over time. All these objects will play a decisive role in identifying the parameters of interest. These objects can be estimated non-parametrically and identification of these functions depends on having enough variation in the data (rank condition) and not having unobserved heterogeneity (model misspecification). In our case part of the unobserved heterogeneity is productivity, which is recovered in the first step.

Once the second step is concluded, the estimated dynamic (cost) parameters rationalize observed behavior (i.e. the parameters for which estimated policy functions, returns, and state transition are optimal). The lack of an analytical solution for the optimality conditions, forces us to use computational methods. I now describe how the process works and clarify that identification does not depend on implausible features of the data. Take our main object of interest, sunk costs of R&D. We know the size, productivity and the (rational) expectations of the firms who decided to start doing R&D. We also know (recovered) the gross profits they expect to earn from making these decisions. The estimated sunk costs compare the profits earned by the firms at a given state that decided to do R&D with the profits of the firms that decided not to do R&D. Had these costs been higher, we would have observed less R&D and had these costs been lower, we would have observed more R&D. The identification of the parameters is therefore very intuitive given the data and observed decisions.

Notice that the use of rationality here might be seen as imposing too much structure as it is not guaranteed that firms form rational equilibrium beliefs. However, this would not be needed if we knew how beliefs are formed. Imagine that firms did not have rational beliefs,

[22] Further restrictions given by economic theory can be used to guarantee identification like continuity and monotonicity (Matzkin, 1994)

but were instead using some ad-hoc (non micro-founded) forecasting method to construct their expectations. Since the equilibrium transition is estimated from the observed data, one could argue that the beliefs they were using would match what we (as econometricians) also observe. This would be along the lines of a behavioral model.[23]

6 The data

The data is part of a database compiled by the Portuguese Central Bank ("Central de Balanços"). Observations are for the period between 1994-2003 for the five-digit NACE (rev 1.1) industry, 29563. This database collects, financial information (balance sheet and P&L) together with other variables like number of workers, occupation of workers (5 levels), total exports, R&D, founding year and current operational status (e.g. operating, bankrupt, etc). Industry aggregate variables for sales, number of firms, employment and value added come from the Portuguese National Statistics Office (INE, 2007) and industry price data from IAPMEI (2006). The detailed data appendix provides a thorough description of the dataset and variable construction.

6.1 Descriptive statistics

The dataset contains 1,290 observations for 231 firms over the period 1994-2003. There are 265 observations with positive R&D that correspond to 59 firms and 49 R&D start-ups (defined as the first year of positive R&D expenditures reported). On average, an R&D firm reports positive R&D for 2.5 consecutive years (Table B.V).

Due to the short nature of the panel, there are very few observations on entry and exit. A further complication arises due to the way data has been collected. Since answering the questionnaire is not compulsory, some firms might not be reported in the dataset but still be active in the industry. This complicates the identification of exiting and entering firms since they could have been operating in the market before first appearing in the dataset. This problem is addressed with two variables that help to identify entry and exit. For entry, firms report their founding year so this is matched with the year the firm first appeared in the sample and if this is within a 2 year window, the firm is considered a new entrant (this is reported in Table B.V under the column "entry"). Regarding exit, the central bank collects a variable that reports the "status" of the firm. The problem with this variable is that some firms that might have closed down are still reported as "active", so only a fraction of the total exits can be captured. Using this methodology identifies a total of 48 entries and 7 exits from the panel.

[23] An alternative recently explored in Aradillas-Lopez and Tamer (2009) is the use of rationalizability to derive bounds on the parameters by using weaker concepts than full rationality. Fershtman and Pakes (2009) also have some very interesting work on extending the rationality concept.

Tables B.I and B.IV report summary statistics for the main variables. The average firm in the sample sells goods worth 1.5 million Euros and employs 32 workers with an average labor productivity of 20,381 Euros. Over the period 1994-2003, real sales have grown at an average 8.9% and labor productivity at 5.7%.

After a decline until 1998, the total number of firms in the industry has grown to a maximum of 1,230 in 2003, employing 10,108 workers. The industry is populated by small and medium firms and there are no market leaders. R&D performers are larger and older and their labour productivity is on average 20% higher.

7 Results

As explained above, the estimation is performed in three steps. The first step recovers an estimate for productivity (TFP). In the second step the reduced form profits, the policy functions for investment, R&D and exit are estimated as well as the transition functions for productivity and the aggregate state (capital accumulation is deterministic). Finally in the third step the dynamic parameters are recovered so that they rationalize these estimated objects. Since any error or bias in the policies or transitions will be transmitted nonlinearly in the final step, the results are sensitive to the second step . For this reason several robustness checks were performed in the second step, using alternative polynomials, different static profit functions or productivity measures. Besides this, alternative specifications for the dynamic cost parameters are reported. Overall the evidence of relatively large sunk costs of R&D is quite robust.

7.1 First step

7.1.1 Productivity (production function)

Production function estimates are reported in Table I.[24]. The estimated labor and capital coefficients are 0.51 and 0.4, respectively, while the estimated demand elasticity implies a price-cost margin of 8%. These values are at a reasonable level and within the range of parameters found in the literature for other industries. In columns (iii) and (iv) results are also reported for a linear and a sigmoidal parametrization for productivity transition. The advantages of these two specifications is that they both preserve monotonicity. Overall the differences are negligible which gives us more confidence that functional forms are not restrictive.

Results using alternative methods are also reported. In particular using a simple fixed effects specification with time dummies (column (v)) does not perform well due to the fact that productivity is serially correlated. Adding the control function $E[\omega_{it}|\omega_{it-1}, R_{it-1}]$ to this

[24] The results are reported for value added (and not sales) production functions. The approach is identical under the assumption that materials are a constant share of total sales. Since by definition $Y_{it} = VA_{it} + M_{it}$, if $M_{it} = \alpha_m Y_{it}$ then $Y_{it} = VA_{it} + \alpha_m Y_{it}$ so that $Y_{it} = \frac{1}{(1-\alpha_m)} VA_{it}$.

	(i)		(ii)		(iii)		(iv)		(v)		(vi)	
Dependent Variable: log of value added va_{it}												
	First stage (OLS)		2nd stage (NLLS) *Cubic Approx.*		2nd stage (NLLS) *Linear Approx.*		2nd stage (NLLS) *Sigmoidal Approx.*		Fixed Effects		*Control function*	
	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*
α_l	0.47	*0.04*							0.69	*0.04*	0.52	*0.04*
α_k			0.37	*0.02*	0.41	*0.03*	0.38	*0.02*	0.37	*0.03*		
η			0.08	*0.04*	0.07	*0.05*	0.07	*0.04*				
γ_0^0			42.45	*33.63*	-0.10	*0.31*	4.32	*0.18*				
γ_1^0			-20.93	*16.77*	1.02	*0.06*	300.39	*191.51*				
γ_2^0			3.76	*2.84*			4.09	*0.09*				
γ_3^0			-0.21	*0.16*			270.20	*169.97*				
γ_0^1			7.76	*3.59*	0.33	*0.13*						
γ_1^1			-3.17	*1.70*	0.94	*0.02*						
γ_2^1			0.75	*0.28*								
γ_3^1			-0.04	*0.02*								
const.							3.57	*0.64*				
R^2	98%		92%		92%		92%		92%		93%	
										Year dummies		
Obs	1,271		1,044		1,044		1,044		1,271		1271	
Firms	227		223		223		223		227		227	

Labor coef,	0.51	0.51	0.74	0.69	
Capital coef.	0.40	0.44	0.40	0.37	-
PC Margin	8%	7%	7%	-	-

Notes: Column (i) reports the results for the first stage of the production function estimates. Columns (ii) to (iv) present the second stage estimates using a cubic, linear and sigmoidal approximation, respectively. Results in column (v) are for a simple fixed effects specification with year dummies and finally column (vi) is identical to column (i) added for fixed effects.

Table I: Production function estimates.

specification (column (vi)), reduces the magnitude of the labor coefficient which suggests that fixed effects and time dummies are not properly capturing the correlation between labor and productivity.

Firms are willing to pay a sunk cost for R&D, if they expect a gain (higher productivity) in the future. In Figure 1 the productivity distribution for R&D and no-R&D firms is reported. There is evidence that R&D firms have better productivity draws and TFP is on average 26% larger.

Finally, notice that results for the dynamic parameters in step three should be sensitive to these productivity estimates. To resolve this issue several alternatives are used. The important feature which affects final estimates is the higher productivity for R&D firms and this is quite robust across all specifications.

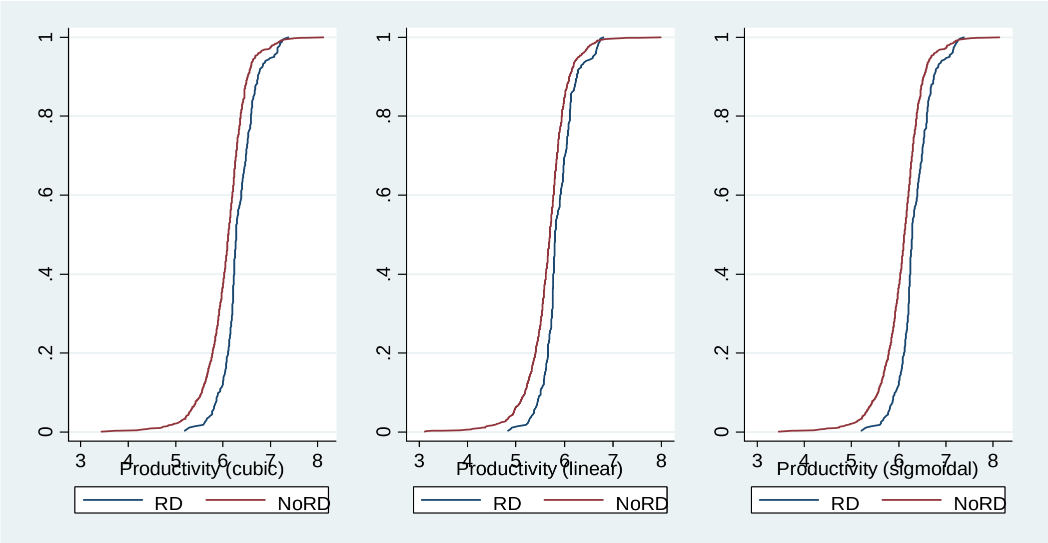

Figure 1: Productivity distribution (cubic, linear and sigmoidal approximation)

	(i)		(ii)		(iii)		(iv)		(v)		(vi)	
Dependent Variable:	Operational profits $\tilde{\pi}_{it}$											
	Cubic approx.		Cubic approx.		Cubic approx.		Cubic approx.		Linear approx.		Sigmoidal approx.	
	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*
α_1^π	0.42	*0.17*	0.43	*0.17*	1.15	*0.05*	1.14	*0.05*	0.40	*0.17*	0.42	*0.17*
α_2^π	0.88	*0.05*	0.88	*0.05*	0.71	*0.02*	0.70	*0.02*	0.90	*0.04*	0.88	*0.05*
α_3^π	0.04	*0.03*	0.04	*0.03*	0.04	*0.06*	0.04	*0.06*	0.04	*0.03*	0.04	*0.03*
α_0^π	-1.63	*1.12*	-1.71	*1.14*	-4.80	*0.83*	-4.72	*0.85*	-1.55	*1.14*	-1.63	*1.12*
α_4^π*	-10.7	*14.6*					-7.6	*14.5*	-10.5	*14.6*	-10.7	*14.6*
α_6^π*	111.8	*25.3*	110.9	*25.3*	126.7	*25.0*	127.3	*25.1*	112.1	*25.4*	111.8	*25.3*
α_5^π	-0.59	*0.37*	-0.58	*0.37*					-0.63	*0.41*	-0.59	*0.37*
R^2	85%		87%		87%		85%		84%		84%	

Notes: Results for the reduced form profit function using different specifications and alternative productivity estimates. *The coefficients α_4^π and α_6^π are scaled down by a factor of 1000.

Table II: Reduced form profit function estimates.

7.2 Second step

7.2.1 Static profits

Using reported profits (cash flow) the reduced form profit function (gross of adjustment costs) in equation (11) can be estimated. On average, firms report 366,146 Euros in profits (with a range from negative 740,000 to more than 14 million Euros). Fixed operating costs seem to be important and increasing for firms with larger capital stocks since both $\hat{\alpha}_4^\pi$ and $\hat{\alpha}_5^\pi$ are negative. R&D firms are estimated to earn on average 111,776 Euros more. All these results are robust across all specifications as reported in the remaining columns of Table II.

7.2.2 Transition function

Aggregate state The aggregate state will be a normally distributed Markov process if the assumptions of Corollary 3 are satisfied. This is useful since we are only required to estimate three parameters: the mean, variance and autocorrelation. These are:

$$\mu_S = 13.18 \quad \sigma_S = 0.28 \quad \rho_S = 0.79$$

However, since the aggregate state is average industry deflated sales, it is not guaranteed that it can be represented as the sum of independent and identically distributed conditional variables. Alternatively, the transition function, $q(S_{t+1}|S_t)$ can be estimated. Results using a polynomial approximation are reported in Table III.

Dependent Variable: Aggregate State $ln(S)$	(i) Coef.	(i) *s.e.*	(ii) Coef.	(ii) *s.e.*	(iii) Coef.	(iii) *s.e.*	(iv) Coef.	(iv) *s.e.*
$ln(S_{t-1})$	0.54	*0.14*	0.39	*0.30*	16.11	*14.54*	1592.84	*1049.42*
$ln(S_{t-2})$			-0.11	*0.21*				
$ln(S_{t-1})^2$					-0.60	*0.56*	-121.61	*80.54*
$ln(S_{t-1})^3$							3.09	*2.06*
Constant	6.15	*1.87*	9.52	*2.53*	-95.03	*94.51*	-6940.60	*4556.67*
Observations	11		10		11		11	
Adjusted R^2	62%		25%		66%		75%	
Mean $ln(S)$	13.18							
St. Dev $ln(S)$	0.28							
Autocorrelation $ln(S)$	0.79							

Notes: Column (i) specifies a linear first order Markov process and column (ii) a second order Markov process. Columns (iii) and (iv) present results for a second and third degree polynomial.

Table III: Aggregate state transition estimates.

Specification test A test which rejects the results from Proposition 1, would cast doubts on the aggregate state model. In particular, rejection of a Markovian aggregate state would raise concerns about using an aggregate state model to represent industry dynamics. The problem arises because, even by restricting players to use Markovian strategies (dependent on payoff relevant variables), the resulting equilibrium evolution for the aggregate state might not be Markovian.[25]

$$p(S_{t+1}|S_t, S_{t-1}, ...S_0) \neq p(S_{t+1}|S_t)$$

The violation of Assumption 3.3 could lead to a history dependent evolution for the aggregate state. Results in column (ii) of Table III do not reject the aggregate state model.[26]

[25] In general, an agggregate state which is a collection of several independent Markovian variables of order one, will not be Markovian of order one in itself.

[26] I have also tested the significance of distribution moments for the state variables (ω_{it}, k_{it}) conditional on

Productivity Using the productivity estimates from step one, we estimate the transition function in eq. (16), separately for R&D and non R&D firms. Again, other parametric specifications are reported. The use of a cubic polynomial fits the data well (Table IV). R&D firms have a smaller productivity dispersion.

	(i)		(ii)		(iii)		(iv)		(v)		(vi)	
	Non-RD		RD		Non-RD		RD		Non-RD		RD	
Dependent Variable: Productivity $[\omega_t]$												
	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*
$[\omega_{t-1}]$	-2.80	*0.98*	-8.12	*4.65*	0.92	*0.02*	0.93	*0.02*				
$[\omega_{t-1}]^2$	0.61	*0.17*	1.39	*0.74*								
$[\omega_{t-1}]^3$	-0.03	*0.01*	-0.07	*0.04*								
const.	7.97	*1.87*	19.94	*9.66*	0.46	*0.09*	0.46	*0.10*				
$[\gamma_0]$									7.04	*0.03*	7.43	*0.03*
$[\gamma_1]$									64.65	*1.80*	90.32	*2.75*
R^2	84%		93%		82%		92%		-		-	
Obs.	790		254		790		254		790		254	
Firms	197		59		197		59		197		59	
s.e. resid.	0.18		0.09		0.18		0.10		0.23		0.13	

Notes: Columns (i) and (ii) present results for the productivity transition using a 3rd degree polynomial, columns (iii) and (iv) a linear specification and columns (v) and (vi) for the sigmoidal approximation.

Table IV: Productivity transition estimates.

7.2.3 Investment, R&D and Exit policies

Lastly, estimates of the policy functions are needed for the third step where they will be used to simulate optimal behavior. The results are reported for lower order polynomial approximations. The choice against higher order polynomials is due to their weak performance in subsets of the state space with little observations (particularly at the tails). The fitted functions might not preserve basic properties like monotonicity. This can generate very inaccurate predictions for optimal actions, particularly at the tails of data distribution where there are very few observations. Incorrect predicted actions at the extremes might generate very high/low returns because they have a significant impact on average estimates. As mentioned by Aguirregabiria and Mira (2007), noise in the estimates is magnified in the third step due to the nonlinearity in the minimum distance estimator.

The R&D (equation (15)) and exit policy functions were estimated using a probit model whereas the investment policy function (equation (13)) was estimated by ordinary least squares. For the exit policies a simple linear probit on the state variables is used because of data limitations.

S_t. Their significance would again reject the aggregate state model, i.e. it would test the hypothesis that $p(S_{t+1}|g(\mathbf{s}_t), S_t) = p(S_{t+1}|S_t)$. Results (not reported) show that further moments of the state variables are not statistically significant.

	(i)		(ii)		(iii)		(iv)		(v)	
Dep. Var.:	RD probit				Investment				Exit probit	
					RD firms		Non-RD firms			
	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*
$ln(S_{t-1})$	-0.06	*0.22*			-0.18	*0.16*	-0.09	*0.33*	0.05	*0.47*
$ln(K_{t-1})$	2.32	*0.98*	2.09	*0.90*	0.63	*0.42*	-2.71	*1.66*	-0.002	*0.11*
$ln(K_{t-1})^2$	-0.08	*0.04*	-0.07	*0.03*	0.00	*0.02*	0.13	*0.06*		
$[\omega_{t-1}]$	-1.34	*2.23*	0.10	*0.22*	2.47	*1.56*	16.61	*6.17*	-0.50	*0.32*
$[\omega_{t-1}]^2$	0.12	*0.19*			-0.15	*0.13*	-1.27	*0.48*		
Constant	-13.32	*7.97*	-16.87	*6.09*	-4.85	*4.88*	-27.92	*17.50*	-0.21	*6.33*
R^2	-		-		40%		38%		-	
Observations	838		838		801		204		1044	
Firms	212		212		208		51		223	

Notes: Columns (i) and (ii) contain results for the RD start-up probit regression. Columns (iii) and (iv) contain results for the invesment OLS results for the non-RD and RD firms. Finally column (v) contains results for the exit probit regression. All results use productivity estimates with a cubic approximation.

Table V: Policy function estimates: RD, investment and exit.

The results are presented in Table V. The probability of doing R&D is increasing and concave in capital stock which means that larger firms are more likely to pay the sunk cost probably because they are also able to extract a larger benefit from doing R&D. On the other hand, it is only weakly increasing in productivity and there is no clear evidence that more productive firms are more likely to start R&D (selection). It seems that the selection effect occurs non monotonically with firms at the extremes of the productivity distribution being more likely to start R&D but this result is not statistically significant.

Regarding investment decisions, larger firms or those with higher productivity tend to invest more. This effect is stronger for non R&D firms. Finally, larger and more productive firms are less likely to exit but given the number of exits observed, these results are not statistically significant.

7.3 Third step

To estimate the dynamic parameters reported in Table VI (linear and quadratic investment cost, R&D sunk cost and exit value), the third step implements the minimum distance estimator outlined above. Standard errors were calculated using the bootstrap.

The values are estimated with the expected signs. Investment costs are increasing and convex. The exit value is estimated at around 1.8 million Euros. However, given the very small number of exits observed, this exit cost is very imprecise and not statistically significant. Finally for the parameter we are interested in, R&D sunk costs are estimated at about 3.4 million Euros which is almost 2 times the average firm level sales in the industry and more than one year worth of sales for an average R&D firm.

Alternative specifications where the quadratic investment cost term, μ_2, is dropped or a fixed operating cost, F, added are also reported. Overall, the precision of the linear cost for

	μ_1	μ_2	λ	e	F
Cubic approximation					
Coefs	-0.61	-5.82	-3,403,000	1,776,700	-
s.e.	*0.87*	*3.02*	*1,254,081*	*3,490,173*	-
Coefs	-1.61		-4,010,500	759,500	-
s.e.	*0.75*	-	*1,234,338*	*3,933,133*	-
Coefs	-1.62		-3,868,200	94,488,600	4,174,210
s.e.	*0.77*	-	*1,598,430*	*137,614,269*	*6,195,137*

Notes: Estimates for the dynamic parameters and bootstraped standard errors.

Table VI: Estimates for the dynamic parameters.

investment improves and it gives an estimated investment cost of 1.6 Euros for each euro of investment, so that indirect investment costs for investment are around 60%.

A fixed operating cost has been estimated in the profit function $(\alpha_4^\pi, \alpha_5^\pi)$ as reported in Table VI. Further introducing another fixed operating cost automatically increases the exit value, which is now unrealistically large. This illustrates the identification problems where the two parameters are not separately identified because they are almost a combination of each other. For example, if we assume no stochastic environment, a firm can decide to exit today and collect the exit value e or stay one more period and exit tomorrow and collect $\pi + \hat{F} + \beta\hat{e}$. An indifferent firm will have $\pi + \hat{F} + \beta\hat{e} = \hat{e}$. In the model with no fixed cost (F), for an indifferent firm we get $\pi + \beta\widetilde{e} = \widetilde{e}$. Replacing and solving $\hat{e} - \widetilde{e} = \frac{1}{1-\beta}\hat{F}$. This illustrates the difficulty in separately identifying $\hat{F}$ and $\hat{e}$ because in this case there is no normalization for the outside alternative.[27] For this reason, estimating the fixed operating cost in the first step using observed profits is clearly a preferred approach since it normalizes the profits for the outside alternative (i.e. not exiting).

As explained above, bias in the policy function estimates will translate non-linearly into the dynamic parameters' estimates. Several alternative specifications for the policy functions using different degrees for the polynomials were tried. The estimated dynamic parameters are relatively robust to these alternative polynomials. One issue not addressed here is the possibility of unobserved state variables. This is a problem which can bias the estimates but the literature with methods for properly addressing it is still at an early stage.

Using a simple "back of the envelope" calculation we can compare average profits of an R&D firm in the period before it started doing R&D against average profits of an R&D firm and this gives us a difference of $230,000$ Euros. Discounting this difference over a 40 year horizon (imagine on average firms expect to live for 40 years) with a discount factor $\beta = 0.96$ gives a current discounted value of 4.6 million Euros which is slightly above our estimates. This larger value comes from the fact that this rough measure does not account for selection into R&D by larger or more productive firms and it also does not account for capital investment

[27] The estimated values of $\hat{e} = 94,488,600$ and $\widetilde{e} = 1,776,700$ with $\beta = 0.96$ rationalize a "net" value (i.e. $(1-\beta)(\hat{e} - \widetilde{e}) = 3,708,406$) close to the fixed cost estimate, $\hat{F} = 4,174,210$. This illustrates the identification problem since there will be an infinite combination of pairs $(\hat{F}, \hat{e})$ which can rationalize observed decisions.

after the R&D decision. Therefore, this value is in line with the estimates and reinforces the credibility of our results.

Finally, these estimates are relatively robust to different approximations for the policy and transition functions as well as the reduced form profit function. They are also relatively robust to different discount factors, i.e. estimates for the sunk costs increase (decrease) with an increase (decrease) in the discount factor within a sensible range (e.g. $\beta \in [0.92; 0.98]$).

8 Counterfactual Experiments

This section reports the results from three simple policy changes. The objective is to assess their impact on industry R&D, productivity and investment. In the first, sunk costs of R&D are exogenously decreased by 10%, in the second market size is reduced to early 1990's levels (from 405 to 150 million Euros) and finally entry costs are increased by 10%. The simplest example of the first policy could either be a direct R&D start-up subsidy or some more general incentive like the creation of a public research agency dedicated to advising firms during R&D start-ups or the supply of training for workers with very specific skills required to do R&D. These are probably more effective because some of the start-up sunk costs might be duplication costs and a research agency would be able to explore the economies of scale. The second policy, for example, illustrates the effects of increasing trade barriers. The final policy could be the result of an argument whereby the development of sufficiently large firms should be supported in order for these big firms to start doing R&D.

To simulate the effects of these policies we are now required to solve the model. The new equilibrium industry evolution, $q(S_{t+1}|S_t)$, has to be calculated. Some parameters have to be set. These are the distribution for entry cost, total number of players, market size, discrete grid for the state variables and the productivity distribution for entrants. The mean and variance for the productivity distribution of entrants is matched with the actual value in the dataset (5.76 and 0.587 respectively). Total number of players is set at 1,000 and market size at 405 million Euros. The grid used to discretize the state variables is similar to the distribution of the state variables in the data.[28] Finally, the mean and variance of entry costs is calibrated, so that we get an equilibrium number of firms consistent with those observed in the data. Since estimated exit values were negative and very poorly estimated, the exit value distribution is calibrated jointly with the entry distribution to get sensible entry and exit rates of 5% per year.[29]

[28]For productivity (ω): {4; 4.5; 5; 5.5; 5.75; 5.9; 6.05; 6.25; 6.5; 6.75; 6.9; 7.05; 7.2; 7.4; 7.75}
For capital (k): {5; 8.5; 9; 10; 10.5; 10.9; 11.2; 11.5; 11.8; 12; 12.2; 12.4; 12.6; 12.7; 12.8; 12.9; 13; 13.2; 13.4; 13.6; 13.8; 14.1; 14.4; 14.7; 15.25; 16.5}
For the aggregate state (S): {12; 12.15; 12.3;... 14.7; 14.85; 15}

[29]Both distributions are assumed normal. For the entry distribution the mean is set to 660,000 and the standard deviation to 130,000. For the exit value distribution, the mean is set to -360,000 and the standard deviation to 50,000.

	Original		Policy Change		
	Data	Simulated	Sunk Cost	Market Size	Entry Cost
Average Sales (EUR)	*529,665*	*680,171*	*768,042*	*581,636*	*696,345*
Average Number of firms	*681*	*601*	*654*	*280*	*535*
% of RD firms	*21%*	*22%*	*74%*	*14%*	*22%*
Average Productivity	*6.13*	*6.29*	*6.36*	*6.21*	*6.24*
Average Capital Stock	*351,512*	*383,310*	*577,001*	*303,246*	*414,033*
Entry Rate	-	*5.36%*	*2.87%*	*9.30%*	*3.12%*
Exit Rate	-	*5.31%*	*2.63%*	*9.23%*	*3.06%*

Notes: Simulated results for the impact on market structure of a 10% reduction in RD sunk costs, decrease in market size from 405 to 150 millions of euros and 10% increase in entry cost.

Table VII: Counterfactual results.

After setting these I use the algorithm provided in the technical appendix to calculate the equilibrium for the model using the estimated structural parameters. Notice that these experiments are only possible using the aggregate state model which is also relatively fast.[30] Solving a full dynamic game would be computationally prohibitive.

Results are presented in Table VII. A 10% decrease in sunk costs leads to a strong increase in R&D performance, a 7% increase in average productivity and 50% increase in average capital. This means that by reducing sunk costs of R&D the average firms gets larger. Reducing market size to the equivalent of the early 90's leads to a reduction in R&D performance from 22% to 14%. There is also a decrease in average productivity (8%) and capital stock (20%) which illustrate the *trade-induced innovation* mechanism. Finally, the increase in entry costs has a negative effect on productivity (by reducing exit of less efficient firms) while virtually no effect on R&D performance. This shows that while increasing entry costs could potentially be seen as a positive measure in the presence of sunk costs, this would actually have negative effects on average productivity.

9 Conclusion

In this paper I have estimated the sunk costs of R&D for the Portuguese moulds industry using a model, which is computationally tractable and, can be implemented empirically with the most common type of firm level datasets. The model both avoids the "curse of dimensionality" and the existence of unobserved firms in the data. The empirical findings suggest a role for *trade-induced innovation*. In the presence of sunk costs of R&D, access to large external markets might create the necessary conditions for an industry to develop itself and become more competitive by investing in R&D. This seems to be what happened in the Portuguese moulds industry after the country joined the EEC in 1986.

[30] Solving the model takes about 150 minutes of computer time on a simple 2.0 Ghz Pentium Core2 Duo with 2GB Memory RAM.

The idea explored to simplify the complex models' dynamics, was to summarize the industry state by the (payoff relevant) aggregate state. As explained, this implicitly imposes more structure in terms of strategic interactions, i.e. the firms react symmetrically to all of its competitors independently of its state (size, productivity, etc). This simplification does not seem severe for the moulds manufacturing industry because each firm specializes in a particular product, does not observe what its competitors offer, produces almost per piece and prices are contract specific. We have reasons to accept that demand can be reasonably well approximated with a constant elasticity of substitution framework. This simplification goes a long way in allowing us to construct the counterfactuals and answer some significant policy questions.

Finally, the sunk costs of R&D for the Portuguese moulds industry are recovered by using a structural estimation method for microdata (BBL). These are estimated at around 3.4 million Euros, more than one year worth of sales for an R&D firm. The magnitude of the sunk costs suggest that policies cannot disregard the discreteness of the R&D decision. Particularly, policies targeted at reducing the sunk costs and increasing R&D start-ups will be more effective at increasing overall industry productivity.

Using the aggregate state framework, we are able to solve the model and perform counterfactual experiments. In particular, market size increase (similar to entering the EU) has a positive effect on R&D performance and productivity. Furthermore, a decrease in the sunk costs of R&D will have a similar effect while an increase in protectionism by increasing entry costs, will have a negative effect on productivity via selection (exit of less productive firms).

The existence of serially correlated unobservables and the extension to estimation techniques which can be more efficient but require equilibrium calculations (Rust, 1987) are important concerns, left for future research.

References

[1] Ackerberg, D., Benkard, L., Berry, S., and Pakes, A. (2007) "Econometric Tools for Analyzing Market Outcomes", Handbook of Econometrics, in: J.J. Heckman & E.E. Leamer (ed.), edition 1, volume 6, chapter 63

[2] Ackerberg, D., Caves, K and Frazer, G. (2006) "Structural Identification of Production Functions", mimeo

[3] Aghion, P., Bloom, N., Blundell, R., Grifith, R., and Howitt, P. (2005); "On Competition and Innovation: An Inverted U Relationship"; Quarterly Journal of Economics 120 (2)

[4] Aguirregabiria, V. and Ho, C. (2009) "A Dynamic Oligopoly Game of the US Airline Industry: Estimation and Policy Experiments", manuscript

[5] Aguirregabiria, V. and Mira, P. (2002); "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models"; Econometrica 70 (4)

[6] Aguirregabiria, V. and Pedro Mira (2007); “Sequential Estimation of Dynamic Discrete Games”, Econometrica, 75(1), 1-53.

[7] Aradillas-Lopez, A. and Tamer, E. (2008); "The Identification Power of Equilibrium in Simple Games" Journal of Business and Economic Statistics 26 (3), 261-283

[8] Aw, Roberts and Xu (2009): "R&D Investment, Exporting and Productivity Dynamics" NBER Working Paper 14670

[9] Bajari, P. Benkard, C. and Levin, J. (2007), "Estimating Dynamic Models of Imperfect Competition", Econometrica, 75 (5), 1331-1370

[10] Bajari, P., Chernozhukov, V., Hong, H., and D. Nekipelov (2008); "Nonparametric and Semiparametric Estimation of Dynamic Games of Incomplete Information", working paper

[11] Bajari, P., Hong, H., and Steve Ryan (2007); "Identification and Estimation of A Discrete Game of Complete Information", Working Paper, MIT

[12] Baumol, W. and Robert Willig (1981); "Fixed Cost, Sunk Cost, Entry Barriers and Sustainability of Monopoly ", Quarterly Journal of Economics 95, 405-431

[13] Beira, E., Crespo, C., Gomes, N. and Menezes, J. (2003); "Dos moldes à engenharia do produto, a trajectória de um cluster", in M. Heitor, J.M.B. Brito e M.F. Rollo (eds), "Momentos de Inovação e Engenharia em Portugal no seculo XX", Lisboa: Dom Quixote

[14] Beira, E. and Menezes,J. (2003) “Inovação e indústria de moldes em Portugal: a introdução do CAD/CAM/CAE nos anos 80”, in Brito et al, vol. III

[15] Blundell, R., Griffith, R. and J. Van Reenen (1999): "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms, Review of Economic Studies 66 (3)

[16] Bond, S. and Måns Söderbom, (2005); "Adjustment costs and the identification of Cobb Douglas production functions", IFS Working Papers, W05/04.

[17] CEFAMOL, (2008); "Current Situation of the Portuguese Mouldmaking Industry", http://www.cefamol.pt/

[18] Collard-Wexler, A. (2006); "Demand Fluctuations in the Ready-Mix Concrete Industry", mimeo

[19] De Loecker, J (2007); "Product Differentiation, Multi-product Firms and Estimating the Impact of Trade Liberalization on Productivity", NBER WP13155

[20] Dixit, A. (1988); "A General Model of R&D Competition and Policy", The RAND Journal of Economics, 19(3), 317-326

[21] Doraszelski, U. and Satterthwaite, M. (2007), "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity", mimeo

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

[23] Foster, L. Haltiwanger, J. and Syverson, C. (2008); "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" American Economic Review, 98(1)

[24] Hashmin, A. and Van Biesebroeck (2009) "Market Structure and Innovation: A Dynamic Analysis of the Global Automotive Industry", manuscript

[25] Hayashi, F. (1982); "Tobin's Marginal q and Average q: A Neoclassical Interpretation", Econometrica 50, 213-224

[26] Hopenhayn, H. (1992), "Entry, Exit and Firm Dynamics in Long Run Equilibrium", Econometrica, 60(5), 1127-1150

[27] Hotz, V. and Miller, R. (1993); "Conditional Choice Probabilities and the Estimation of Dynamic Models"; Review of Economic Studies 60 (3); pp 497-529

[28] Hotz, V. J., Miller, R. A., Sanders, S., and J. Smith (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice," Review of Economic Studies, 61:2, 265-289.

[29] IAPMEI, (2006); "A Industria Portuguesa de Moldes no Contexto Internacional - contributos para o conhecimento da actividade", www.iapmei.pt

[30] Jovanovic, B. (1982); "Selection and the Evolution of Industry", Econometrica, 50(3)

[31] Klette, T. and Griliches, Z., (2000), "Empirical Patterns of Firm Growth and R&D Investment: A Quality Ladder Model Interpretation", The Economic Journal, 110, 363-387

[32] Klette, T. and Kortum, S. (2004); "Innovating Firms and Aggregate Innovation", Journal of Political Economy 112 (5), 986-1018

[33] Krusell, P. and A. Smith (1998); "Income and Wealth Heterogeneity in the Macroeconomy", Journal of Political Economy 106(5), 867-896

[34] Levinsohn, J and Amil Petrin, (2003); "Estimating Production Functions Using Inputs to Control for Unobservables", Review of Economic Studies, 317-342.

[35] Magnac, T. and Thesmar, D. (2002); "Identifying Dynamic Discrete Decision Processes", Econometrica, 70, 801-816

[36] Maskin, E., and Tirole, J. (1988); "A Theory of Dyamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs", Econometrica 56(3), 549-569

[37] Maskin, Eric and Tirole, Jean; (2001); "Markov Perfect Equilibrium: I. Observable Actions"; Journal of Economic Theory; 100 (2); 191-219

[38] Matzkin, R.L. (1994) "Restrictions of Economic Theory in Nonparametric Methods," Handbook of Econometrics, Vol. 4, edited by C.F. Engel and D.L. McFadden, Elsevier.

[39] Melitz, M. (2000); "Estimating Firm-Level Productivity in Differentiated Product Industries; Harvard University

[40] Olley, S. and Ariel Pakes, (1996); "The Dynamics of Productivity in the Telecommunications Equipment Industry," Econometrica, vol. 64, no. 6, 1263-98.

[41] Pakes, A. and McGuire, P. (2001); "Stochastic Algorithms, Symetric Markov Perfect Equilibrium, and the "Curse" of Dimensionality"; Econometrica; 69 (5); 1261-1281

[42] Pakes, A., Ostrovsky, M. and Steve Berry (2007); "Simple Estimators for the Parameters of Discrete Dynamic Games (with entry/exit examples", RAND Journal of Economics 38(2)

[43] Pavcnik, N. (2002): "Trade liberalization, exit, and productivity improvement: Evidence from Chilean plants", Review of Economic Studies 69(1), 245-276

[44] Pesendorfer, M. and Philipp Schmidt-Dengler (2008): "Asymptotic Least Squares Estimators for Dynamic Games", Review of Economic Studies, 75 (3), 901-928

[45] Ryan, S. (2006): "The Costs of Environmental Regulation in a Concentrated Industry", Working Paper 0510, MIT

[46] Rust, J. (1987), "Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher", Econometrica 55 (5); 999-1033

[47] Rust, J. (1994), "Structural Estimation of Markov Decision Processes" chapter 51 in D. McFadden and R. Engle (eds.) Handbook of Econometrics 4 North Holland, 3082–3139.

[48] Santos, C. (2008) "Recovering the Sunk Costs of R&D: The Moulds Industry Case"

[49] Santos, C. and Van Reenen, J. (2008) "Identifying financial constraints in a dynamic structural model of R&D and investment: the US Iron and Steel industry", mimeo

[50] Schmalensee, R. (1992); "Sunk Costs and Market Structure: A Review Article", The Journal of Industrial Economics, 40 (2), 125-134

[51] Schmidt-Dengler, P. (2006); "The Timing of New Technology Adoption: The Case of MRI", mimeo

[52] Stiglitz, J. (1987); "Technological Change, Sunk Costs and Competition", Brooking Papers on Economic Activity, 3, 883-947

[53] Sutton, J. (1991): Sunk Costs and Market Structure, MIT Press

[54] Sutton, J. (1998): Technology and Market Structure, MIT Press

[55] USICT, United States International Trade Comission, (2002) ; "Tools, dies and industrial molds: competitive conditions in the United States and selected foreign markets", Investigation no. 332-435, USITC Publication 3556

[56] Varela-Irimia, X. (2008) "Entry Costs and Economies of Scope in Multiproduct Firms' Desicions", manuscript

[57] Vives, X. (2004); "Innovation and Competitive Pressure", CEPR DP 4369

[58] Weintraub, G. Benkard, L. Van Roy, B. (2008), "Markov Perfect Industry Dynamics with Many Firms", Econometrica, 76 (6), 1375-1411

[59] Xu, D. (2008) "A Structural Empirical Model of R&D, Firm Heterogeneity and Industry Evolution", manuscript

A Appendix

A.1 Proof of proposition 1

Proof. Using Assumptions 3.1 to 3.3, S_t is the payoff relevant variable and $g(\mathbf{s}_t|S_t,...S_0) = g(\mathbf{s}_t|S_t)$. Define (i) the conditional industry state evolution resulting from assumption 3.1 and the Markovian strategies as $p^{\sigma}(\mathbf{s}_{t+1}|\mathbf{s}_t, S_t) = \Pi_{i=1}^{N} \int_{\varphi_{it}} p(s_{it+1}|s_{it}, \sigma(s_{it}, S_t, \varphi_{it}))dF^{\varphi_{it}}$ (ii) $f^{\sigma}(\mathbf{s}_{t+1}|S_t, ..., S_0)$ as the industry state distribution conditional on the whole history for the aggregate state. The distribution for the aggregate state conditional on the history is

$$\begin{aligned}
f^{\sigma}(S_{t+1}|S_t, S_{t-1}, ..., S_0) &= \int_{\varepsilon_{t+1}} \int_{\mathbf{s}_{t+1}:S_{t+1}=S(\mathbf{s}_{t+1})+\varepsilon_{t+1}} f^{\sigma}(d\mathbf{s}_{t+1}|S_t, ..., S_0)dF_{\varepsilon_{t+1}} \\
&= \int_{\varepsilon_{t+1}} \int_{\mathbf{s}_{t+1}:S_{t+1}=S(\mathbf{s}_{t+1})+\varepsilon_{t+1}} \int_{\mathbf{s}_t} p^{\sigma}(d\mathbf{s}_{t+1}|\mathbf{s}_t, S_t, ..., S_0)g(d\mathbf{s}_t|S_t, ..., S_0)dF_{\varepsilon_{t+1}} \\
&= \int_{\varepsilon_{t+1}} \int_{\mathbf{s}_{t+1}:S_{t+1}=S(\mathbf{s}_{t+1})+\varepsilon_{t+1}} \int_{\mathbf{s}_t} p^{\sigma}(d\mathbf{s}_{t+1}|\mathbf{s}_t, S_t)g(d\mathbf{s}_t|S_t, ..., S_0)dF_{\varepsilon_{t+1}} \\
&= \int_{\varepsilon_{t+1}} \int_{\mathbf{s}_{t+1}:S_{t+1}=S(\mathbf{s}_{t+1})+\varepsilon_{t+1}} \int_{\mathbf{s}_t} p^{\sigma}(d\mathbf{s}_{t+1}|\mathbf{s}_t, S_t)g(d\mathbf{s}_t|S_t)dF_{\varepsilon_{t+1}} \\
&= \int_{\varepsilon_{t+1}} \int_{\mathbf{s}_{t+1}:S_{t+1}=S(\mathbf{s}_{t+1})+\varepsilon_{t+1}} f^{\sigma}(d\mathbf{s}_{t+1}|S_t)dF_{\varepsilon_{t+1}} \\
&= q^{\sigma}(S_{t+1}|S_t)
\end{aligned}$$

where the first step follows from using the law of total probability; the second step from the definition of $p(\mathbf{s}_{t+1}|\mathbf{s}_t, S_t)$ given above; the third step from 3.3; and the final step again from the law of total probability. ■

B Detailed Data Appendix

B.1 Data and sample construction

The data comes from three sources: Aggregate variables (sales, value added, employment) come from the Portuguese National Statistics Office (INE); Industry price deflators are collected from

IAPMEI (2006); Finally firm level data was extracted from the Bank of Portugal database on firms across the economy (Central de Balancos, 5 digit NACE code industry 29563 - moulds industry).

Some notes on the "Central de Balancos": The dataset has been collected by the Central Bank since 1986. However, due to changes in accounting rules, it is only comparable from 1990. The quality of the data between 1990 and 1994 is not considered reliable. From 2000, the sampling method (simple random sampling) was changed to stratified sampling and this caused a drop in the number of observed (mostly smaller) firms in 1999 and 2000.

Representativeness: The sample is representative of the whole industry, in particular for the early periods. It covers 90% of total sales and industry employment in 1994 and this coverage decreases to a minimum of 50% of sales (40% of employment) in 2003. This reduction is mainly due to changes in the sampling procedure as explained above. There is an obvious gap in productivity trends between the sample and the industry (also total sales and employment). Labor productivity in the sample increased by roughly 60% while this was only 40% in the industry. For this reason none of the aggregate variables are calculated using the sample but come directly from the collected industry wide variables.

Variable construction:

- Capital stock was calculated using the perpetual inventory method with a 8% depreciation rate

$$K_{it+1} = (1 - depreciation) * K_{it} + I_{it}$$

- Value added is equal to sales subtracted from materials and external services expenditures

$$VA_{it} = Y_{it} - M_{it} - ESE_{it}$$

- R&D dummy variable takes a value equal to one whenever positive R&D was reported in the past or present and zero otherwise.

Both aggregate and individual sales and value added were deflated with the industry price deflator.

In 11 observations the number of workers reported was zero and these were dropped.

There were 9 holes identified in the sample, i.e. firms that interrupt reporting for 1 or more consecutive years. In these cases either the earlier or later periods are dropped, minimizing the total number of observations lost.

Entry and exit are difficult to identify since it is not compulsory for firms to report to the central bank. However, the dataset has information on the founding year and current firm "status" (i.e. active, bankrupt, merged, etc). Using this information 48 actual entries and 7 exits were identified.

B.1.1 Aggregate State

Market definition: The market is defined as total worldwide demand for Portuguese moulds. This is mainly for simplicity reasons since there is no good data on worldwide production and our dataset contains total industry sales (national plus exports). This adds the restriction that firms in the Portuguese market take as exogenous the evolution of demand for their market. Imagine that Y^W is world demand for moulds and $Y = Y^W - Y^{NP}$ is total demand for Portuguese moulds. The assumption is that foreign competition is exogenous so that Y^{NP} evolves exogenously (Y^W, is assumed to evolve exogenously depending on the economic conditions, etc).

Variable definition: The aggregate state is defined as average deflated total industry sales, which approximately matches the variable in the demand system: $\frac{Y/N}{\widetilde{P}}$.

This can be divided into three variables. The first is total industry sales (Y) and can easily be assumed to evolve exogenously. As explained above, this is total demand for Portuguese moulds. The second variable is the industry price ($\widetilde{P}$) and is the solution to the static pricing game. In the technical appendix the pricing strategies are shown to be a mapping from states onto the pricing space. Therefore, this variable evolves endogenously. Finally, the total number of incumbents (N) is also endogenous and it depends on market size. Modeling these three variables separately would involve taking into account (and estimating) all cross correlations.

In Figure B.2 I plot the evolution of all variables. We can observe that the market was growing mainly between 1993 and 2000 and pauses until 2003 and the number of firms share a similar pattern. On the other side prices were increasing slightly over this period and decreased in the later years. This characterizes most of what has already been explained before. The industry grew substantially after 1994 due to the strong increase in demand for Portuguese moulds. Together with this increase in demand we also observe an increase in labor productivity and R&D. The cross correlations are as expected with prices being negatively correlated with number of firms and market size, and the number of firms being positively correlated with market size. The evolution of these three variables will be summarized by the evolution of the single index variable, $\frac{\widetilde{Y}}{\widetilde{P}}$(average deflated sales). This also helps to address potential non-stationarity problems (see below).

Addressing non-stationarity issues: The industry grew substantially in the period 1994-2003 so this raises concerns over non-stationarity. To analyze this we can look at the evolution of average deflated industry sales. The plot in Figure B.2 clearly shows that while it is true that average sales were growing between 1994-1998, it seems to have stabilized over the later period. The justification for this performance is the increase in the number of firms in the later period (after 1998) as is evident in Table B.I.

	Number of firms	Production (EUR mio)	Exports (EUR mio)	Exports %	Total Employment	Value Added (EUR mio)	Price (EUR/ton)
1994	644	171	132	77%	5,133	101	24.43
1995	570	193	151	78%	5,796	114	25.25
1996	452	244	191	78%	7,316	143	25.71
1997	477	293	220	75%	7,821	166	25.73
1998	461	322	232	72%	7,740	167	24.62
1999	549	362	250	69%	8,429	208	25.23
2000	604	412	277	67%	8,879	228	26.49
2001	612	421	328	78%	8,919	240	26.74
2002	722	378	310	82%	9,312	235	24.97
2003	738	403	303	75%	8,766	227	22.86
2004	1109	455	340	.	9,846	259	20.33
2005	1230	468	298	.	10,108	256	18.69

Source: National statistics office, INE 2007

Table B.I: Aggregate variables

	1970	1980	1985	1990	1995	2000	2003
1	USA	USA	USA	USA	USA	France	Germany
2	UK	UK	UK	France	France	USA	France
3	W. Germ.	Sweden	Russia	Germany	Germany	Germany	Spain
4	Canada	Mexico	Israel	UK	UK	Spain	USA
5	Venezuela	W. Germ.	Venezuela	Netherlands	Netherlands	UK	UK
6	Nd	France	France	Spain	Israel	Sweden	Sweden
7	Nd	Netherlands	Netherlands	Sweden	Belg./Lux.	Netherlands	Netherlands
8	Nd	Venezuela	Sweden	Israel	Sweden	Israel	Romania
9	Nd	Spain	Spain	Belg./Lux.	Brazil	Belg./Lux.	Switzer.

Source: CEFAMOL, 2008

Table B.II: Export ranking by destination country.

	France	Germany	Spain	USA	UK	Sweden	Netherlands	Belgium-Lux
1996	31,044	23,912	6,746	30,737	10,181	18,130	14,393	7,995
1997	30,416	31,462	8,740	33,714	21,333	19,179	11,771	7,561
1998	26,456	35,230	11,176	32,115	25,079	12,670	9,323	7,289
1999	45,767	36,314	23,172	37,876	17,058	9,760	9,103	9,210
2000	51,829	37,869	28,843	46,857	27,670	13,055	11,862	7,229
2001	71,222	53,863	35,659	36,687	25,133	10,979	12,940	8,444
2002	65,368	53,007	47,796	36,210	24,541	18,377	7,911	7,971
2003	61,633	66,837	39,909	44,102	16,177	15,364	6,527	6,527
2004	71,766	61,395	42,781	30,720	33,618	13,556	5,478	5,478
2005	68,221	47,233	40,399	20,074	16,615	11,586	9,275	9,275

Source: CEFAMOL, 2008

Table B.III: Exports to main destinations, thousands of euros.

	Mean	Std. Dev.	Min	Max
All firms: 1274 observations				
Sales (EUR)	1,574,073	2,869,201	3,292	34,700,000
Exports (EUR)	891,333	2,483,554	0	31,800,000
Capital Stock (EUR)	1,058,104	2,130,734	135	23,800,000
Employment	32	39	1	258
Labor Productivity (EUR)	20,381	9,044	359	74,632
Investment rate	0.20	0.25	0.00	5.32
Sales growth	8.9%	34.5%	-195.8%	469.0%
Value added Growth	9.4%	40.5%	-289.3%	477.0%
Labor Productivity growth	5.7%	37.1%	-289.3%	284.3%
Non RD firms: 1009 observations				
Sales (EUR)	1,198,854	2,321,233	3,292	26,800,000
Exports (EUR)	640,879	1,919,257	0	25,200,000
Capital Stock (EUR)	835,706	1,854,294	135	20,600,000
Employment	27	35	1	230
Labor Productivity (EUR)	19,609	9,178	359	74,632
Investment rate	20.9%	27.5%	0.0%	531.7%
Sales growth	9.9%	37.9%	-195.8%	469.0%
Value added Growth	10.4%	45.2%	-289.3%	477.0%
Labor Productivity growth	6.2%	41.1%	-289.3%	284.3%
RD firms: 265 observations				
Sales (EUR)	3,002,735	4,066,477	99,206	34,700,000
Exports (EUR)	1,844,947	3,811,178	0	31,800,000
Capital Stock (EUR)	1,904,897	2,802,605	53,161	23,800,000
Employment	52	45	3	258
Labor Productivity (EUR)	23,321	7,861	7,148	59,923
Investment rate	16.7%	14.2%	0.0%	77.5%
Sales growth	5.6%	20.1%	-101.8%	123.3%
Value added Growth	6.3%	19.6%	-113.3%	102.4%
Labor Productivity growth	3.9%	19.9%	-87.2%	116.9%
RD to sales ratio	0.9%	3.4%	0.0%	46.5%

Source: "Central de Balanços", Bank of Portugal

Table B.IV: Summary statistics, by RD status.

Year	*Number of firms*	*Number of non-RD firms*	*Number of RD firms*	*RD start-ups*	*Entry*	*Entry in the dataset*	*Exits*
1994	144	134	10	-	2	3	0
1995	157	137	20	10	12	14	2
1996	165	141	24	4	8	14	0
1997	170	145	25	2	11	20	2
1998	164	135	29	7	9	33	0
1999	136	108	28	3	2	46	1
2000	92	68	24	7	2	8	0
2001	88	56	32	9	1	5	0
2002	88	53	35	4	1	2	0
2003	86	48	38	3	0	0	2
Total	**1290**	**1025**	**265**	**49**	**48**	**145**	**7**

Source: "Central de Balanços", Bank of Portugal

Table B.V: Descriptive statistics.

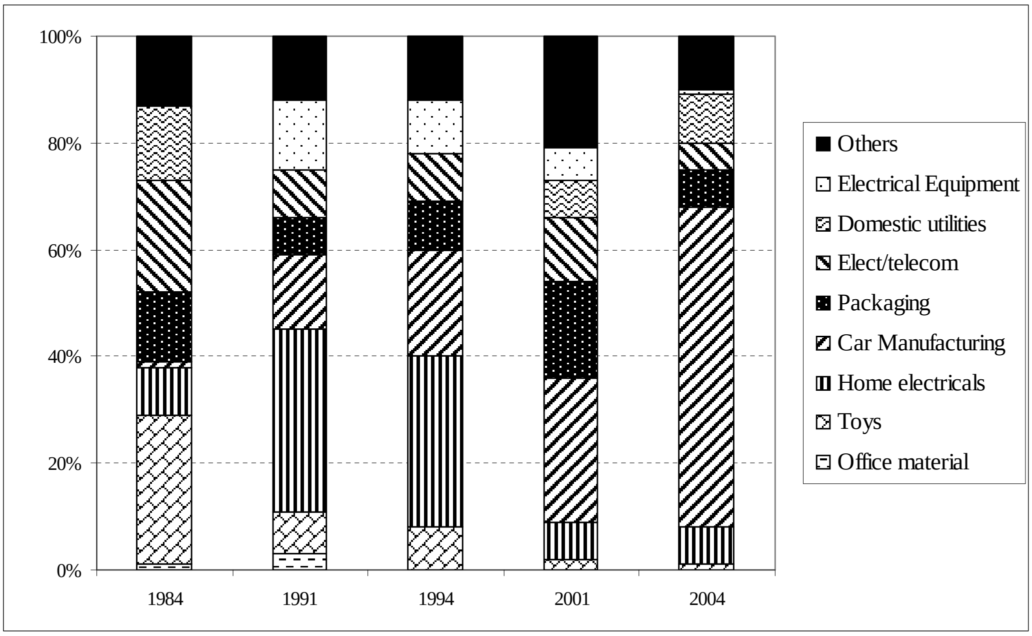

Figure B.1: Sales composition, by client type

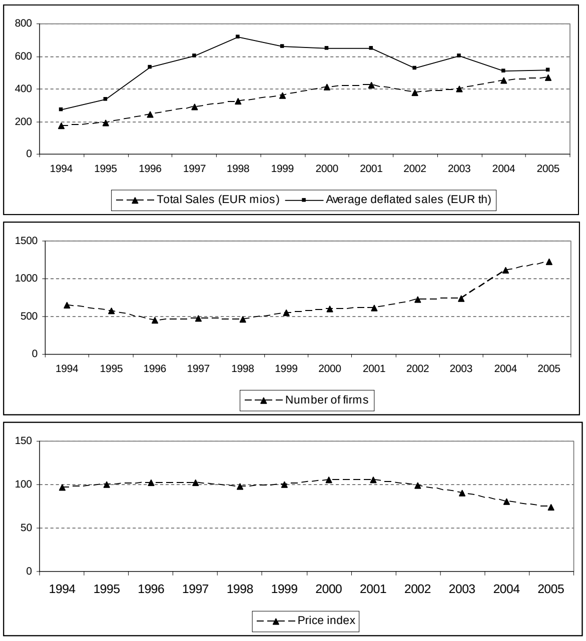

Figure B.2: Aggregate State

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

957	Nicholas Oulton Ana Rincon-Aznar	Rates of Return and Alternative Measures of Capital Input: 14 Countries and 10 Branches, 1971-2005
956	Tim Leunig Chris Minns Patrick Wallis	Networks in the Premodern Economy: the Market for London Apprenticeships, 1600-1749
955	Urban Sila	Can Family-Support Policies Help Explain Differences in Working Hours Across Countries?
954	John T. Addison Alex Bryson Paulino Teixeira André Pahnke Lutz Bellman	The Extent of Collective Bargaining and Workplace Representation: Transitions between States and their Determinants. A Comparative Analysis of Germany and Great Britain
953	Alex Bryson Harald Dale-Olsen Erling Barth	How Does Innovation Affect Worker Well-being?
952	Nathan Foley-Fisher Bernardo Guimaraes	US Real Interest Rates and Default Risk in Emerging Economies
951	Yann Algan Christian Dustmann Albrecht Glitz Alan Manning	The Economic Situation of First- and Second-Generation Immigrants in France, Germany and the United Kingdom
950	Jérôme Adda Francesca Cornaglia	The Effect of Bans and Taxes on Passive Smoking
949	Nicholas Oulton	How to Measure Living Standards and Productivity
948	Alex Bryson Bernd Frick Rob Simmons	The Returns to Scarce Talent: Footedness and Player Remuneration in European Soccer
947	Jonathan Wadsworth	Did the National Minimum Wage Affect UK Wages?
946	David Marsden	The Paradox of Performance Related Pay Systems: 'Why Do We Keep Adopting Them in the Face of Evidence that they Fail to Motivate?'
945	David Marsden Almudena Cañibano	Participation in Organisations: Economic Approaches

944	Andreas Georgiadis Alan Manning	One Nation Under a Groove? Identity and Multiculturalism in Britain
943	Andreas Georgiadis Alan Manning	Theory of Values
942	Kristian Behrens Giordano Mion Yasusada Murata Jens Südekum	Trade, Wages and Productivity
941	David Marsden Richard Belfield	Institutions and the Management of Human Resources: Incentive Pay Systems in France and Great Britain
940	Elhanan Helpman Oleg Itskhoki Stephen Redding	Inequality and Unemployment in a Global Economy
939	Norman Ireland Robin A. Naylor Jeremy Smith Shqiponja Telhaj	Educational Returns, Ability Composition and Cohort Effects: Theory and Evidence for Cohorts of Early-Career UK Graduates
938	Guy Mayraz Jürgen Schupp Gert Wagner	Life Satisfaction and Relative Income: Perceptions and Evidence
937	Nicholas Bloom Raffaella Sadun John Van Reenen	The Organization of Firms Across Countries
936	Jean-Baptiste Michau	Unemployment Insurance and Cultural Transmission: Theory and Application to European Unemployment
935	João M. C. Santos-Silva Silvana Tenreyro	Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically
934	Christian Morrisson Fabrice Murtin	The Century of Education
933	João M. C. Santos-Silva Silvana Tenreyro	Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator
932	João M. C. Santos-Silva Silvana Tenreyro	On the Existence of the Maximum Likelihood Estimates for Poisson Regressioin

The Centre for Economic Performance Publications Unit
Tel 020 7955 7284 Fax 020 7955 7595 Email info@cep.lse.ac.uk
Web site http://cep.lse.ac.uk