Sunk Costs of R&D, Trade and Productivity: the moulds industry case*

Carlos Daniel Santos[†]

Tilburg University

December 30, 2011

Abstract

The evidence suggests that productivity improves with international trade via (i) selection of the best firms and (ii) within firm productivity growth. In this article I focus on the second channel operating through innovation. Trade liberalization allows the best firms to exploit the economies of scale in the R&D process. I estimate an equilibrium model of industry dynamics with endogenous size and productivity decisions. I then show how the European economic integration induced innovation in the Portuguese Moulds industry.

Keywords: Industry Dynamics, Innovation, Markov Equilibrium, Moulds Industry, R&D, Structural Estimation, Sunk Costs, Trade

JEL Classification: C61, D92, F12, L11, L22, O31

1 Introduction

Trade liberalization can trigger innovation. Firms competing internationally can exploit economies of scale by having access to larger markets. In the presence of sunk costs, trade

*Acknowledgments: 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, Adam Sanjurjo and 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 and was previously named "Recovering the Sunk Costs of R&D: the moulds industry case". Research supported by the Portuguese Science and Technology Foundation grant SFRH/BD/12092/2003.

[†]Tilburg University, Netherlands; c.d.santos@uvt.nl

liberalization can thus create the conditions for innovation. In this article I illustrate how foreign trade created the conditions for the best firms to grow and triggered innovation for an export-oriented industry in a small economy. The increase in innovation arises from the interaction of heterogeneity in firm size and the costs of innovation. In this setting, while trade increases the sales of all firms, only the best will cross the threshold and pay the sunk costs of innovation. One particularly important aspect of this industry is that almost all firms export. I use this fact to focus the analysis on the channel of trade induced innovation and isolate it from the selection mechanism that has been the focus of most of the literature.

The trade literature has emphasized the role of firm heterogeneity in explaining why the most productive firms select into exporting, in line with the evidence (e.g. Eaton and Kortum (2002), Melitz (2003), Bernard et al. (2003)). Empirical studies have also documented that firms exhibit productivity growth after entering the export market (e.g. Pavnick (2002), Bernard et al. (2006), De Loecker (forthcoming)). To match this fact, recent work allows for endogenous productivity choices (e.g. Constatini and Melitz (2008), Aw, Roberts and Xu (2009), Atkenson and Burstein (2010)) or learning by exporting (De Loecker (forthcoming)). However, most existing models predict market size neutrality, which is at odds with the empirical literature (e.g. Bresnahan and Reiss (1991), Eaton, Kortum and Kramarz (2008)). Allowing for more plausible competition features can relax this prediction and create size effects (e.g. Melitz and Ottaviano (2008), Zhelobodko, et al. (2011)). Size effects can also be created by introducing adjustment costs for physical capital (Cooper and Haltiwanger (2006)).

In this article I incorporate these empirical facts, namely, endogenous productivity decisions and size effects. I specify a dynamic equilibrium model with endogenous innovation and physical capital to analyze the effect of trade on equilibrium innovation and R&D. To do so I abstract from the fixed and sunk costs of exporting thereby abstracting from the separate decision of selling to the internal and external markets. In practice this assumption shuts down the selection mechanism, making the interaction of heterogeneity and export costs irrelevant, consistent with the particular industry studied here. Instead, I focus on the interaction between heterogeneity and the costs of innovation. Trade liberalization can allow the best firms to exploit the economies of scale and innovate. The framework matches the empirical facts for the Portuguese Moulds industry, where I estimate the model. First, the industry has some characteristics that make it fit to address

this particular question. The internal market is negligible and 80% of the firms export 90% of the production. Thus, I can abstract from import competition and focus on the size effect of trade liberalization. Second, the selection mechanism plays a very marginal role for this industry allowing me to safely ignore export costs. Incorporating export costs is straightforward and would cause the selection mechanism to reinforce the effects analyzed here.

The model is similar in spirit to the industry equilibrium models of Hopenhayn (1992) and Ericson and Pakes (1995). In building the model I bring together three strands from the literature on trade, investment (Cooper and Haltiwanger (2006)) and innovation (Klette and Kortum (2004)). I introduce investment in order to explain the large investment rates observed in this industry. Modelling investment in physical capital is important because firm size is the main explanatory (state) variable for innovation, consistent with the existence of sunk costs of innovation. On the technical side, I introduce incomplete information and assume that individual players do not observe the state or actions of their competitors. Players' strategies and beliefs are Markovian and depend on payoff relevant states. These assumptions allow me to break the "curse of dimensionality" and solve the model with a large number of players. Notice that since the industry is very fragmented and industrial Moulds are intermediate products, the assumption of incomplete information about the state of competition is plausible.

Estimation is done in three steps using the method developed by Bajari, Benkard and Levin (2008), henceforth BBL. The first step recovers total factor productivity by estimating a production function. In the second step I estimate the policy and transition functions. Finally, in the third step I search for the structural (cost) parameters that rationalize the estimated policy functions as optimal. I estimate the model using data for the Portuguese moulds industry. The dataset cover the period 1994 to 2003, after the country joined the EU (1986) and the European Common Market (1993). I then illustrate how the *trade-induced innovation* effect operates by varying trade costs and simulating the new stationary equilibrium. The results show that an increase in trade costs is expected to cause a reduction in R&D performance, average productivity, physical capital and number of firms - in line with the evidence. To my knowledge this article is the first to propose and estimate a dynamic equilibrium model with many players using microdata.

During the analysis I abstract from export costs and general equilibrium. As explained

above, abstracting from export costs is a reasonable approximation for this particular industry and allows me to isolate the trade induced innovation mechanism. Extending the model to incorporate export costs is straightforward and would simply require extra data to guarantee identification of such costs (i.e. variation in export decisions). Extending the framework to general equilibrium is conceptually possible, but it would require more than one single industry. There is thus a difficult trade-off between using rich features in partial equilibrium and estimating more parsimonious general equilibrium models. The conclusions can be interpreted in light of the modeling choices. First, when the trade selection mechanism is relevant, we would expect the interaction of innovation and trade to reinforce the effects studied here as illustrated by Aw, Roberts and Xu (2009). On the other hand, as Atkenson and Burstein (2010) show, trade is expected to have general equilibrium economy wide effects on wages that could partially offset the effects in the long run.

2 The moulds industry

The Portuguese moulds industry is an interesting case of success. The ability to develop itself even in the absence of a national market and the fact that most of the production has always been exported illustrates the importance of foreign markets. Between 1994 and 2003 the growth in exports was mainly to other European countries in detriment of the US, traditionally the larger market (Table I). After 2001 the industry stabilized and by 2005 nearly 90% of total production was exported (out of which 72% to the automotive industry).

Rank	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	Venez.	W. Germ.	Venez.	Holland	Holland	UK	UK
6	Nd	France	France	Spain	Israel	Sweden	Sweden
7	Nd	Holland	Holland	Sweden	Belg./Lux.	Holland	Holland
8	Nd	Venez.	Sweden	Israel	Sweden	Israel	Romania
9	Nd	Spain	Spain	Belg./Lux.	Brazil	Belg./Lux.	Switzerland

Source: CEFAMOL (2008)

Table I: Ranking of export destinations.

Portuguese mouldmakers benefit from a good international reputation. A report by the US international trade commission (USITC (2002)) emphasized the fast delivery, technology, quality and competitive price as the main strengths.

> *"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 (Tools, Dies and Moulds) 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)

Entering the *European Economic Community* in 1986 and the establishment of the *Common European Market* in 1993 were two important factors that explain the observed growth. With such events, firms were able to reach important European producers and later increase R&D and innovation. Beira et al. (2003) and IAPMEI (2006) document how the close collaboration with car-makers constituted a strong push for the development of new processes and products, normally in the form of research projects. The typical research projects have well defined targets and objectives and are implemented over a period of 2 to 3 years. Projects need to be completed for the gains to materialize. The data is consistent with the project-based view of R&D for this industry. R&D spans last on average 2.5 years and most firms report uninterrupted R&D over the R&D span.[1] The project-based nature of R&D motivates the sunk cost specification. An alternative per unit or fixed cost structure, would be better suited for industries where R&D is performed continuously (for example the chemical or pharmaceutical industries).

Brief industry history

> The industry's history dates back to the 1930's and 1940's when the development of plastics generated a demand for moulds. The Portuguese moulds industry

[1] Only 16 out of 59 firms interrupt R&D. Eight interrupt it for one year, seven for two years and one for five years.

grew in the late 1950's as a major supplier of moulds to the glass (where it inherited some of its expertise) but mainly to the toy industry. In the late 1980's the production started shifting from toy manufacturing towards the growing industries of automobiles and packaging. During the 1990's the largest markets shifted from the US to France, Germany and Spain (Table I, IAPMEI (2006)).

The industry underwent several changes with the introduction of new technologies (*e.g.* CAD, CAM, Complex process, In-mould Assembling) and an increasing importance of innovation and R&D. For example, the technology used in computer operated machines radically changed from the 1980's to the 1990's. State of the art machinery allows flexibility at a low cost. It also allows a close collaboration with the client in the design phase. Design teams can work closely with engineers and produce 3D virtual versions of the mould before it is produced in the final phase. Controlling the technology is a major requirement from car-makers.

3 The framework

I will start by outlining the model to be estimated and describe the elements of a more general model together with a characterization of the equilibrium in the Technical Appendix. The framework captures the relevant industry features described above, namely, the size effects and the increase in Research and Development expenditures. Incumbent firms can increase size by investing in physical capital or increase productivity by engaging in R&D activities. R&D and innovation takes the form of projects, implemented over a given period (for simplification I assume this period is one year). The projects cannot be undone so I model the cost of improving productivity to be sunk. Once the R&D project is finished and implemented the gains become permanent. Firms can also exit the market and collect a scrap value. New firms can enter the market by paying an entry fee. To close the model I define the equilibrium for the dynamic game played by all the firms. Since the industry is very fragmented (largest market share is below 10% in any given year), I will abstract from strategic interactions and summarize industry competition with an aggregate state. Firms have rational expectations about future competition. In equilibrium, the expectations (beliefs) must match actual behavior. The equilibrium

is stochastic due to both aggregate uncertainty and a finite number of players.[2]

The features that distinguish this from previous work are the introduction of size (capital) and endogenous productivity. Having access to larger markets triggers the size effects induced by endogenous capital choices. The increase in size makes firms more likely to do R&D and subsequently improve their productivity.

Notice that the unimportant internal market allows me to both ignore import competition and abstract from modeling exports as a separate decision. Since innovation is profitable for large firms, the increase in demand generated by a reduction in trade costs explains the large observed investment rates.

3.1 Demand

The internal market is virtually inexistent. Modelling the demand from the external market and abstracting from considering the internal market, let quantity demanded (Q_i) for a moulds producer in Portugal with price P_i be

$$Q_i = Q_0 \frac{Y^W}{(P^W)} \left(\frac{P^P}{P^W}\right)^{-\sigma_w} \left(\frac{P_i}{P^P}\right)^{-\sigma} e^{\varepsilon_i^d} \tag{1}$$

where Q_0 is a constant, P^W is the world price index, Y^W is the world market size, ε_i^d is a demand shock and σ_w and σ are the across and within country price elasticities. The world variables are taken as exogenous by the firms in Portugal (small country assumption). Notice that the final price P_i is equal to the producer price times the trade cost of exporting to the world, τ. Let P^P be the final price index for Portuguese moulds and Y^P the size of the market for Portuguese Moulds. We can write the total market for portuguese moulds as $Y^P = Q_0 Y^W \left(\frac{P^P}{P^W}\right)^{1-\sigma_w}$ and the demand for the products of firm i is

$$Q_i = \frac{Y^P}{P^P} \left(\frac{P_i}{P^P}\right)^{-\sigma} e^{\varepsilon_i^d} \tag{2}$$

This formulation is quite general as it allows the market for portuguese moulds to have the same elasticity as the world market (the standard case), smaller or larger elasticity. If $\sigma_w = \sigma$ we are in the simple case and if $\sigma_w = 1$ we have cobb douglas utility. When

[2] The models of Hopenhayn (1992) or Melitz (2003) remove any aggregate uncertainty to obtain a deterministic equilibrium.

$\sigma_w = 1$ the shares of total revenues are determined exogenously and a price decrease is compensated by an exact quantity increase leaving total revenues unchanged. When $\sigma_w > 1$, for a reduction in price there is an increase in total revenues.

There are two reasons to write the demand as a function of (Y^P, P^P) instead of (Y^W, P^W). The first is data availability for the national aggregates but not worldwide. The second is the definition of the relevant geographic market. It is unclear what is the correct geographic market affecting the national producers (i.e. world total, total for a selected group of countries, etc). It is probably better to construct the relevant indices using different weights for different countries with more important destinations receiving a larger weight.

3.2 Production function

Production function takes the form of a standard Cobb-Douglas with capital (K), labour (L), materials (M) and Hicks-neutral technology (ω)

$$Q_i(K_i, L_i, M_i, \omega_i) = A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m} \tag{3}$$

where $\alpha_k, \alpha_l, \alpha_m$ are the capital, labour and materials elasticities and A_0 is a constant.

3.3 Static Profits

Capital is a state variable subject to adjustment costs, while labour and materials can be freely adjusted. Firms do not observe the demand shock ε^d before choosing labor and materials. The firm's objective function (integrated over ε^d) for given wages (w) and price of materials (P_m) is

$$\max_{L_i, M_i} \tau^{-1} P(Q_i, Y, P) Q_i(K_i, L_i, M_i, \omega_i) - wL_i - P_m M_i - FC(K_i, R_i)$$

where FC are simply the fixed costs of operation that can depend on current capital stock and R&D. Focusing on the variable component and replacing (2) and (3) in the objective function

$$\max_{L_i, M_i} [A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m}]^{(\sigma-1)/\sigma} \tau^{-1} (Y^P)^{1/\sigma} \left(P^P\right)^{(\sigma-1)/\sigma} - wL_i - P_m M_i$$

which after solving (see the technical appendix) results in

$$\Pi^G = \frac{1}{\gamma}\left[\left(\frac{\alpha_l}{w}\right)^{\alpha_l}\left(\frac{\alpha_m}{P_m}\right)^{\alpha_m}\left(\frac{\sigma-1}{\sigma}\right)^{\frac{\sigma}{\sigma-1}}\right]^{\gamma}\left[A_0 e^{\omega_i} K_i^{\alpha_k}\right]^{\gamma}\tau^{-\frac{\sigma}{\sigma-1}\gamma}\left(\frac{Y^P}{(P^P)^{1-\sigma}}\right)^{\frac{\gamma}{\sigma-1}} \tag{4}$$

where $\gamma = \frac{(\sigma-1)}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}$.

Let $\alpha = (\alpha_l, \alpha_m, \alpha_k, \sigma)$ be the set of unknown parameters. Notice this is equivalent to the short run gross profit function. The net profit function will include the costs of capital investment and R&D (see below).

3.4 Aggregate state

We can define the aggregate state as $S_t = \frac{Y^P}{(P^P)^{1-\sigma}}$. The technical appendix contains a discussion about the aggregate state from the model and the one observed in the sample and its importance for estimation and counterfactual simulations. For the simulations we need to construct the aggregate state from the individual variables. We get the expression

$$\frac{Y^P}{(P^P)^{1-\sigma}} = g1.\tau^{\frac{(\sigma-\sigma_w)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}\left[\sum\left((A_0 e^{\omega_i} K_i^{\alpha_k})^{\gamma}\right)\right]^{\frac{1}{\gamma}\left(\frac{(\sigma_w-\sigma)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}\right)} \tag{5}$$

where $g1 = \left(Q_0 Y^W\left(\frac{1}{P^W}\right)^{1-\sigma_w}\right)^{\frac{(\sigma-(\alpha_l+\alpha_m)(\sigma-1))}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}\left[\left(\frac{\alpha_l}{w}\frac{\sigma-1}{\sigma}\right)^{\alpha_l}\left(\frac{\alpha_m}{P_m}\frac{\sigma-1}{\sigma}\right)^{\alpha_m}\right]^{\frac{-(\sigma-\sigma_w)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$.

Notice that we can construct the aggregate state up to a constant $g1$, since the components of $g1$ are unknown. To construct the final aggregate state I first calibrate the model at the estimated parameters to recover the scaling constant.

3.5 The dynamics

I now define the state variables, their transition and the unobserved heterogeneity.

3.5.1 States and actions

Each firm is characterized by four state variables: Capital, productivity, R&D status (denoted by $R \in \{0,1\}$) and the active status (where $\chi = 1$ if the firm is an incumbent and $\chi = 0$ if the firm is either a potential entrant or exited the industry). The state of firm i in period t is written in compact notation as

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

Unobserved heterogeneity is required to fit the model to the data. I specify four privately observed payoff shocks: to investment costs ε_{it}^{I} , to the sunk cost ε_{it}^{RD}, to the scrap value ε_{it}^{scrap}, to demand ε_{it}^{d}, and to profits ε_{it}^{π}. The distribution of these shocks is nonparametrically unidentified (see section 4.4). The shocks also satisfy conditional independence (Rust (1994)). Thus, I will assume the vector of payoff shocks $\boldsymbol{\varepsilon} = (\varepsilon^{I}, \varepsilon^{RD}, \varepsilon^{entry}, \varepsilon^{scrap}, \varepsilon^{d}, \varepsilon^{\pi})$ are independently drawn with distribution, $F(\varepsilon)$.

Firms' available choices are the investment level ($I \in \mathbb{R}_{+}$) which adds to the current capital stock and R&D ($R \in \{0,1\}$) which leads to productivity improvements. Firms can also decide to enter or exit ($\chi \in \{0,1\}$). The choices are summarized in the vector of actions a_{it}

$$a_{it} = (I_{i,t+1}, R_{i,t+1}, \chi_{i,t+1})$$

For convenience I will sometimes use the compact notations for the vector of states and actions (s_{it}, a_{it}).

3.5.2 State transition

Individual states evolve over time. Most naturally, the evolution is stochastic for productivity and deterministic for capital.

Productivity R&D firms have better prospects for their productivity than non-R&D firms (in a probabilistic sense) and productivity follows a controlled first order Markov process.

$$\begin{aligned} \omega_{i,t+1} &= E(\omega_{i,t+1}|\omega_{it}, R_{it}) + \eta_{i,t+1} \qquad (6) \\ &= P^{\cdot,\omega}(\omega_{it}, R_{it}) + \eta_{i,t+1} \end{aligned}$$

where η_{it} is independently and identically distributed across firms and time and $P^{\cdot,\omega}(.)$ is a parametric function used to approximate the conditional expectation function. The transition for individual productivity is estimated separately for R&D and non-R&D firms and several functional forms for $P^{\cdot,\omega}(.)$ will be reported, including polynomial functions of order n.

Capital stock The capital stock depreciates at rate δ and investment adds to the stock

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

R&D A non-R&D firm ($R_{it} = 0$) can decide to pay the sunk cost and do R&D ($R_{is} = 1$ for all $s > t$). R&D is an absorbing state and the effect is a permanent technological upgrade. A more flexible alternative, where firms can switch in and out of R&D or perform more than one project, would require a longer time series to be estimated accurately given that such switches are not frequent in the data. The permanent shift might be sufficient to understand the steady state effects of trade liberalization on R&D and productivity.

Entry/exit A potential entrant can enter and an incumbent remain in the market, $\chi_{i,t+1} = 1$. Alternatively, an incumbent can exit and a potential entrant can stay out, $\chi_{i,t+1} = 0$. Exiting or staying out are also absorbing states in the sense that firms cannot reenter the market at a later period.

3.6 Cost functions

The net profit is equal to the short run gross profits (equation 4) minus the adjustment costs (investment, R&D, exit).

$$\Pi(s_{it}, S_t, a_{it}, \boldsymbol{\varepsilon}_{it}) = \Pi^G(s_{it}, S_t, \boldsymbol{\varepsilon}_{it})e^{\varepsilon^d_{it}} - FC_i - C(s_{it}, S_t, a_{it}) \tag{7}$$

where $C(.)$ is the cost function that includes four adjustment costs which I now specify.

Investment cost Lumpy investment behavior at the plant level is well documented (e.g. Cooper and Haltiwanger (2006)). It can be rationalized by the existence of strong non-convex adjustment costs and irreversibilities. I specify a quadratic adjustment cost function with total irreversibility.

$$C^K(I_{it}, K_{it-1}) = \left[\mu_1 I_{it} + \mu_2 \frac{I_{it}^2}{K_{it-1}}\right] + \varepsilon^I_{i,t-1} I_{it} \quad \text{if} \quad I_{it} \geq 0 \tag{8}$$

where μ_2 indexes the degree of convexity. This simple parametrization still captures

$$
\begin{aligned}
&\Pi(s_{it}, S_t, a_{i,t+1}, \boldsymbol{\varepsilon}_{it}; \alpha, \theta) = \\
= \; &\Pi^G(s_{it}, S_t, \alpha) e^{\varepsilon^d_{it}} - FC_i + \varepsilon^{\pi}_{it} - \mu_1 I_{i,t+1} - \mu_2 \frac{I^2_{i,t+1}}{K_{it}} - \varepsilon^I_{it} I_{i,t+1} \\
&-(\lambda + \sigma^{\lambda} \varepsilon^{RD}_{it})(R_{i,t+1} - R_{it}) R_{i,t+1} + (1 - \chi_{i,t+1})(e + \sigma^e \varepsilon^{scrap}_{it})
\end{aligned} \tag{9}
$$

Where $\Pi^G(.)$ is the gross profit function (gross of adjustment costs) and the payoff shocks enter additively (Rust (1994)).

3.7 Value function

We can finally write down the dynamic problem faced by an individual firm. Players choose actions to maximize the discounted sum of profits

$$
\begin{aligned}
V(s_{it}, S_t) &= \max_{\{a_{is}\}_{s=t}^{\infty}} E_t \sum_{s=t}^{\infty} \beta^{s-t} \Pi(s_{is}, S_s, a_{is}, \boldsymbol{\varepsilon}_{is}) \\
&= \max_{a_{it}} \Pi(s_{it}, S_t, a_{it}, \boldsymbol{\varepsilon}_{it}) + \beta E_t^{a_{it}} V(s_{i,t+1}, S_{t+1})
\end{aligned} \tag{10}
$$

where the expected continuation value is integrated over the possible values of the state variables conditional on the current state and actions

$$
E_t^{a_{it}} V(s_{i,t+1}, S_{t+1}) = \int_{s_{i,t+1}} \int_{S_{t+1}} V_{i,t+1} p(a_{it}, s_{it}, s_{i,t+1}) q(s_{it}, S_t, S_{t+1}) ds_{i,t+1} dS_{t+1}
$$

the transition $p(a_{it}, s_{it}, s_{i,t+1})$ is the probability of reaching state $s_{i,t+1}$ from state s_{it} when action a_{it} is chosen, and the function $q(s_{it}, S_t, S_{t+1})$ is the probability of reaching state S_{t+1} from state (s_{it}, S_t). The individual transition $(p(.))$ is a primitive of the model assumed known to all players, whereas the industry state transition $(q(.))$ represents the beliefs about the evolution of the industry state.

Equilibrium In equilibrium, beliefs $q(s_{it}, S_t, S_{t+1})$ are consistent with optimal behavior (see the Technical Appendix) . The (stochastic) *equilibrium* transition for the aggregate state is $q^*(s_{it}, S_t, S_{t+1})$.

4 The estimation procedure

The Appendix provides a description of the sample and variable construction. Estimation is done in three steps. In the first step, I estimate the production function, the productivity transition ($p(\omega_{i,t+1}|\omega_{it}, R_{it}, \chi_{it})$ and the profit function ($\Pi(\omega_{it}, K_{it}, R_{it}, S_t, \alpha)$). In the second step I estimate the transition for the aggregate state $q(S_{t+1}|S_t)$, and the equilibrium policy functions for investment ($I(\omega_{it}, K_{it}, R_{it}, S_t)$), R&D ($R(\omega_{it}, K_{it}, R_{it}, S_t)$) and exit ($\chi(\omega_{it}, K_{it}, R_{it}, S_t)$).[4] In the third step the dynamic parameters ($\mu_1, \mu_2, \lambda, \sigma^{\lambda}, e, \sigma^e$) are recovered using the equilibrium conditions. Since the model has no analytical solution, I use numerical methods.

4.1 Step 1: Static parameters

The production function, demand elasticity and productivity transition are estimated with the method proposed by Gandhi et al (2011) which is outlined in the Appendix. Total factor productivity is then recovered as the residual from the production function.[5]

4.2 Step 2: Policies and transitions

4.2.1 Policies

Investment The optimal solution to the dynamic problem faced by an individual firm in equation (10) is the investment function, estimated separately for R&D and non-R&D firms using a flexible polynomial

$$i_{i,t+1} = P^{n,i}(\omega_{it}, K_{it}, S_t, R_{it}) + \varepsilon^I_{it} \tag{11}$$

A lower choice of n to approximate the polynomial function $P^{n,\cdot}(.)$ is preferred. Higher order polynomials can create large distortions and generate a poor approximation (e.g. Runge's phenomenon). This is more likely to occur in intervals of the data with less observations that are normally at the tails of the distribution. Such large distortions in the policy functions' estimates at the tails can create a large bias in the (averaged) estimates. As a further check on the goodness of fit of the approximations, I evaluate how well the estimated policies match the data by comparing the simulated predictions with the observed behavior.

[4] Assuming players own effect on the aggregate state is negligible we can write $\boldsymbol{q}(\boldsymbol{s}_{it}; \boldsymbol{S}_t; \boldsymbol{S}_{t+1}) = \boldsymbol{q}(\boldsymbol{S}_t; \boldsymbol{S}_{t+1})$. This is the case when players are infinitesimal.

[5] See Ackerberg et al. (2007) for a survey on methods to estimate production functions.

R&D The probability of doing R&D is

$$\begin{aligned}\Pr(R_{i,t+1} &= 1|R_{it}=0, s_{it}, S_t) = \\ &\Phi\left(-\lambda + \beta\begin{bmatrix} E\{V(s_{i,t+1}, S_{t+1})|R_{i,t+1}=1\} \\ -E\{V(s_{i,t+1}, S_{t+1})|R_{i,t+1}=0\}\end{bmatrix}\right)\end{aligned}$$

which is approximated by

$$\Pr(R_{i,t+1}=1|R_{it}=0) = \Phi\left(P^{n,rd}(\omega_{it}, K_{it}, S_t, R_{it}=0)\right) \tag{12}$$

The same argument in favor of lower order polynomials is in place here.

Exit Finally, the exit function is treated similarly

$$\Pr(\chi_{i,t+1}=0|\chi_{it}=1) = \Phi\left(P^{n,\chi}(\omega_{it}, K_{it}, S_t, R_{it})\right) \tag{13}$$

Equation (11) is estimated by OLS and equations (12) and (13) by maximum likelihood.

4.2.2 Aggregate state transition

The equilibrium transition for the aggregate state is estimated directly from the data. The aggregate state is the sum of conditionally independent variables. Using the central limit theorem its transition can be approximated parametrically by

$$\ln(S_{t+1}) = \mu_{SS} + \rho_S \ln(S_t) + \nu_{t+1} \tag{14}$$

where ν_{t+1} is a normal zero mean random variable with variance $\sigma_{SS}^2 = \sigma_S^2(1-\rho_S^2)$.[6] The variance of the aggregate state represents the aggregate uncertainty affecting investment. The intercept is $\mu_{SS} = (1-\rho_S)\mu_S$, and $(\mu_S, \sigma_S, \rho_S)$ are the unconditional mean, variance and autocorrelation for the $\ln(S)$ process.

[6] Notice that when the mean is significantly larger than the standard deviation, the log-normal approximates the normal distribution.

4.3 Step 3: Minimum distance estimator

Using the estimated profits, policies and transition functions, the cost parameters are finally recovered as follows:

1. Choose n_s different starting values $\{s_{it}^{(j)}, S_t^{(j)}\}_{j=1}^{n_S}$ for $t = 0$. For example the starting values can be the observations in the sample;

2. For each value $(s_{it}^{(j)}, S_t^{(j)})$, draw a vector of payoff shocks, $(\varepsilon_{it}^{I}, \varepsilon_{it}^{RD}, \varepsilon_{it}^{scrap})$.[7]

3. Calculate actions $(a_{it}^{(j)})$ at each state $(s_{it}^{(j)}, S_t^{(j)})$ using the estimated policy functions and the payoff shocks drawn in 2.;

4. Draw shocks to productivity $(\eta_{i,t+1})$ and to the aggregate state (ν_{t+1}) and update states $(s_{i,t+1}^{(j)}, S_{t+1}^{(j)})$ using the shocks drawn, the estimated transition functions and the actions calculated in 3.;

5. Repeat steps 2. to 4. for $\bar{T}$ periods, and construct a sequence of actions and states: $\{a_{it}(s_{i0}^{(j)}, S_0^{(j)}), s_{it}(s_{i0}^{(j)}, S_0^{(j)}), S_t(S_{i0}^{(j)})\}_{t=1}^{\bar{T}}$. Use this sequence to calculate the discounted stream of profits for a given parameter vector θ and for each $j = 1, .., n_s$:

$$V^{(j,m)}(s_{i0}^{(j)}, S_0^{(j)}) = \sum_{t=0}^{\bar{T}} \beta^t \Pi(a_{it}, s_{it}, S_t, \boldsymbol{\varepsilon}_{it}; \hat{\alpha}, \hat{P}^n, \theta)$$

6. Repeat steps 2. to 5. for $m = 1, ..., n_J$ times and calculate an average estimate for the expected value at each (j) state:

$$\widehat{EV}(s_{i0}^{(j)}, S_0^{(j)}; \tilde{\alpha}, \theta) = \frac{1}{n_J} \sum_{m=1}^{n_J} V^{(j,m)}(s_{i0}^{(j)}, S_0^{(j)})$$

The equilibrium conditions imply that at equilibrium beliefs, $q^*(.)$, strategy $a(.)$ is

an equilibrium if for all $a' \neq a$ the following condition holds:

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

[7]Notice we can abstract from the demand and profit shocks ("^{d}; "$^\pi$) since they are neutral when integrated out.

Given the linearity of the value function in the dynamic parameters we can write

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

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

7. Calculate the difference between the optimal and non-optimal value function for each policy/state pair $(X_k, k = 1, ...n_I)$, where $X_k = (a'_{it}, s_{i0}, S_0)$ and there are $n_I = n_a * n_s$ such pairs:

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

The equilibrium condition implies that estimated policies are optimal and the expected value of using strategy a should not be smaller than using alternative a'. A violation of equilibrium conditions occurs when $\hat{g}(X_k, \theta, \hat{\alpha}, \hat{P}^n) < 0$. The empirical minimum difference estimator minimizes[8] the squared violations of such equilibrium conditions

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

I set the time of each path at $\bar{T} = 90$, the discount factor at $\beta = 0.93$, the number of starting configurations is the total number of observations ($n_s = 1,271$), the number of simulations for each configuration $n_J = 200$ and the number of alternative policies $n_a = 150$, giving a total number of "differences", $n_I = 190,650$.

Choice of alternative policies The vector of cost parameters, θ, must rationalize the estimated strategy profile, $\hat{a}$. In general, θ can be point or set identified, depending on the model and available data. The minimum distance estimator above can also lead to a loss of point identification due to the generation of alternative policies. Bajari et al. (2007) also propose an estimation method when the parameters are set identified.

[8] When the objective functions is not smooth (*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, the Nelder-Mead algorithm) to minimize the empirical minimum distance ($\hat{\boldsymbol{J}}$) helps to circumvent these problems. Non-smooth functions occur with finite $\boldsymbol{n}_I$, because of the min operator in the empirical objective function, $\hat{\boldsymbol{J}}$. This can create non-differentiability even when $\boldsymbol{g}()$ is differentiable.

The inequality in the objective function, $J(\theta)$, arises from comparing $\hat{a}$ with alternative policies, a'. The way alternative policies are chosen will influence estimation and identification. If we choose alternative policies very *far* from $\hat{a}$, the identified set increases and we only achieve parametric identification while if we choose alternative policies very *close* to $\hat{a}$, the identified set shrinks. Since we can produce as many alternative policies as required, we can improve the estimation by choosing the alternative policies and artificially generating non-optimal observations. However this also leads to the problem of too many uninformative moments and loss of efficiency.

This raises the question of how to choose the alternative policies? One option is to add a slight perturbation to the estimated policy function. If the perturbation is positive, all alternative actions will be larger and the parameters are only identified in a one sided set. For example, if alternative R&D start-up decisions are more frequent (i.e. positive perturbations added to the optimal policy function), the alternative policies will generate high levels of R&D behavior. Firms will do more R&D than actually observed in the sample which can only be rationalized if the sunk costs are high (bounded below). However, the sunk costs are not bounded above unless we also add negative perturbations. I add zero-mean normally distributed errors to the investment, R&D and exit policies with standard deviations of 0.15, 0.075 and 0.075. This translates in a 95% probability for the alternative policies for investment to be in the interval $\pm 30\%$ from the estimated one. I also evaluate the sensitivity of the estimates to other choices of a'.

4.4 Identification

Identifying restrictions 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 identification problem in dynamic models is well known (Rust (1994), Magnac and Thesmar (2002), Pesendorfer and Schmidt Dengler (2008) and Bajari et al. (2008)).

The unknown structural objects are the period returns, distribution for the shocks, state transition function and discount factor: $\Pi(a_{it}, \mathbf{s}_t, \boldsymbol{\varepsilon}_{it}), F(\boldsymbol{\varepsilon}_{it}), p(s_{i,t+1}|s_{it}, a_{it}), \beta$. These objects cannot be separately point identified so we either introduce assumptions or estimate the identified sets. Pesendorfer and Schmidt Dengler (2008) extend the work of Magnac and Thesmar (2002) on single agent models to dynamic games and provide one

solution to the identification problem. One solution is to normalize the period returns for some outside alternative and to use exclusion restrictions (i.e. state variables that can be excluded from the period returns but enter the policy functions). Even so, without further restrictions the discount factor and distribution of cost shocks are nonparametrically unidentified. There are also other alternatives. One is the use of parametric restrictions on the return function . Another alternative is to estimate part of the return function directly when some measure of returns is observed.

Part of the return function ($\Pi^G(.;\alpha)$) is estimated in the first step. Assuming agents have rational expectations, their beliefs are recovered by estimating the evolution for the states in the second step. The discount factor and the distribution of shocks $(\beta, F(\varepsilon))$ are set exogenously. The only object left to estimate is the cost function. This function takes the value of zero when there is inaction (no investment and no R&D start-up), $C(s_{it}, a_{it} = 0) = 0$ and is unaffected by productivity and the aggregate state. Thus, the cost function satisfies both the normalization and exclusion restrictions.

Sample features The cost function is identified from easy to intepret sample features. The cost parameters are estimated to rationalize the observed choices given the observed returns and state transitions. For example, the estimated sunk costs compare the profits earned by the firms that decided to do R&D at a given state with the profits of the firms that decided not to do R&D. Had these costs been higher, we would observe less R&D and had these costs been lower, we would observe more R&D.

Beliefs Firms operate in a dynamic and uncertain environment and have to form beliefs about the future. In general, these beliefs are not known or even estimable by the econometrician. For this reason, the hypothesis of rationality is kept and the beliefs are matched to the observed outcomes. Rationality can be relaxed in some particular cases. For example, if the forecasting technology used by firms is known to the econometrician, we can try to estimate the beliefs from the data. Rationality is just one particular forecasting technology that is internally consistent and that can be estimated from the data. Thus, any solution necessarily requires an explicit treatment of beliefs.[9]

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

5 Results

The sample is part of the *Central de Balanços* compiled by Portuguese Central Bank from 1994 to 2003 and a detailed description is provided in the Appendix. Estimation takes three steps. An evaluation of the sensitivity to errors and bias in the profits, policies and transitions is also conducted. As a robustness check I also try alternative polynomials in the second step and different static profit functions. Finally, alternative specifications for the cost function are also reported.

5.1 First step: static parameters

The estimates for the production function and productivity using the GMM estimator outlined in the Appendix are reported in Table II.

Parameter	Estimate	*s.e.*	t
$_{k}$	0.32	*0.12*	2.74
$_{l}$	0.43	*0.07*	6.15
	6.76	*0.68*	9.98
$_{r,0}$	0.61	*0.83*	0.73
$_{r,1}$	0.53	*1.09*	0.49
$_{r,2}$	0.13	*0.45*	0.28
$_{r,3}$	-0.01	*0.06*	-0.20
$_{nr,0}$	1.05	*0.18*	5.97
$_{nr,1}$	0.11	*0.28*	0.39
$_{nr,2}$	0.26	*0.12*	2.12
$_{nr,3}$	-0.03	*0.02*	-1.71
$^{a}_{0}$	-27.55	*13.49*	-2.04
$^{a}_{fc,0}$	-24.75	*0.12*	-207.29
$^{a}_{fc,rd}$	90.05	*28.64*	3.14
$_{fc,K}$	34.77	*0.59*	58.66
Observations	1,044		
Firms	223		

Notes: [a] Coefficients and standard errors divided by 1,000.
Results for the production function (top) and fixed costs component (bottom)

Table II: Production function estimates.

Production function The estimated labor and capital coefficients are 0.43 and 0.31, while the estimated demand elasticity implies a price-cost margin of 17%. The

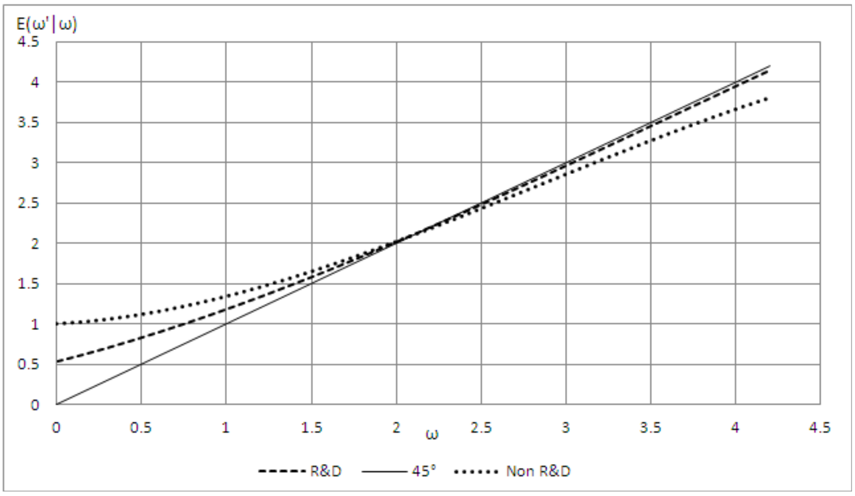


Figure 1: Productivity transition function

values are at a reasonable level and within the range of parameters found in the literature for other industries.

Productivity transition Productivity transition is approximated with a cubic polynomial which is allowed to differ between R&D and non R&D firms. The results reported in Figure 1 illustrate a very persistent productivity for both types of firms, being more persistent for the R&D firms. Also, productivity is more persistent at high productivity than at low productivity levels.

Productivity Total factor productivity is calculated as the residual from a production function. Firms are willing to pay the sunk cost of innovation if they expect a gain in the future, in this case, larger productivity. R&D firms are on average 10% more productive than non R&D firms.

Profit function Finally, the intercept and the fixed cost component of the profit function $(\alpha_0, \alpha_{fc,0}, \alpha_{fc,rd}, \alpha_{fc,K})$ cannot be recovered from the first order conditions of the production function and are estimated separatelly using observed cash flows

$$CashFlow_{it} = e^{\alpha_0}\left[e^{\omega_{it}}K_{it}^{\hat{\alpha}_k}\right]^{\hat{\gamma}}\left(\frac{Y_t^P}{P_t^P}\right)^{\frac{\hat{\gamma}}{\hat{\sigma}-1}} e^{\hat{\varepsilon}_{it}^d} - \alpha_{fc,0} - \alpha_{fc,rd}R_{it} - \alpha_{fc,K}K_{it} + \varepsilon_{it}^{\pi}$$

The estimates in the second panel of Table II report that firms have an average fixed cost of 30 thousand euros with R&D firms earning 90 thousand euros more. Overall the

results illustrate the importance of both physical capital and productivity as determinants of profitability. The fit is quite good and almost 80% of the total variation in profits is explained by the state variables. The persistence of capital and productivity will characterize the persistence of profits.

5.2 Second step: Policy functions

5.2.1 Investment, R&D and Exit

Estimates for the policy functions are reported in Table III. The R&D function in equation (12) is estimated with a probit model, and the investment function in equation (11) is estimated by ordinary least squares. Due to the limited number of observations, a linear probit is used for the exit policies. The preference for low order polynomials is related with the poor fit of the high order polynomials in regions with few observations (particularly at the tails) and the resulting sensitivity of averages to incorrect optimal predictions at the extremes. Also, in some circumstances we would like the fitted functions to preserve basic properties like monotonicity.

	(i)		(ii)		(iii)		(iv)	
Dep. Var.:	RD		Investment				Exit	
			RD firms		Non RD firms			
	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*	Coef.	*s.e.*
Constant	-20.22	*7.18*	-19.10	*12.45*	-7.92	*3.85*	-4.78	*9.83*
$\boldsymbol{ln}(\boldsymbol{S}_{t-1})$	0.01	*0.22*	1.30	*0.53*	0.40	*0.22*	0.31	*0.73*
$\boldsymbol{ln}(\boldsymbol{K}_{t-1})$	2.50	*0.96*	0.42	*1.40*	0.71	*0.36*	-0.09	*0.09*
$\boldsymbol{ln}(\boldsymbol{K}_{t-1})^2$	-0.09	*0.04*	0.02	*0.05*	0.01	*0.01*		
$[!_{t-1}]$	0.32	*0.87*	3.34	*1.30*	2.28	*0.54*	-0.34	*0.33*
$[!_{t-1}]^2$	0.03	*0.17*	-0.55	*0.25*	-0.30	*0.11*		
$\boldsymbol{R}^2$	25%		56%		42%		5%	
Observations	838		801		204		1044	
Firms	212		208		51		223	

Notes: Column (i) contains results for the RD start-up probit regression. Columns (ii) and (iii) contain results for the invesment OLS results for the non-RD and RD firms. Finally column (iv) contains results for the exit probit.

Table III: Policy function estimates: R&D, investment and exit.

Investment For non-R&D firms, investment is increasing in both size and productivity. For R&D firms, the size effect is milder while the productivity effect is stronger.

This suggests size-driven investment for the most productive non-R&D firms, up until the optimal level. Once the optimal level is reached, R&D can be triggered and the size effect becomes less pronounced while productivity becomes the main driver of investment. The aggregate state has a positive effect.

R&D probit There are strong and significant size effects and insignificant productivity effects. Small firms have a probability of doing R&D close to zero, which increases strongly with capital and levels out for large capital stocks. The strong size effect supports evidence for the potential gains from trade. However, conditional on size there is no evidence that more productive firms are more likely to do R&D. The non significant productivity effect suggests the selection of more productive firms into R&D is not meaningful.

Exit Larger and more productive firms are less likely to exit. However, the results are not statistically significant due to the small number of observations for exit. Alternatively we could fit an exogenous exit probability with aggregate moments. Doing so, would not account for selection due to exit and overestimate the value of firms with low productivity, and vice-versa. Either of the approaches has small effects on the value function.

Implications Summarizing, investment of non-R&D firms is strongly influenced by size and productivity. The R&D decision is affected by size but not by productivity. Finally, large and more productive firms are less likely to exit.

The results suggest the following model. Relatively more productive firms start by investing in physical capital. Once they reach a certain size, R&D becomes a profitable investment. After R&D is done, productivity increases and since firms are now larger, investment *rates* decrease. The mechanism is also consistent with the descriptive statistics (Table A.I). Growth rates are larger for non-R&D firms (for revenues, value added and labor productivity) and investment rates are lower for R&D firms. Furthermore, R&D firms are larger and more productive. Even though labor productivity *growth* is larger for non-R&D firms, labor productivity *levels* are larger for the R&D firms. Taken together, these features characterize firm level dynamics and motivate the use of a dynamic model with productivity and capital investment.

5.2.2 Aggregate state transition

The aggregate state transition of equation (14) is fully specified by three parameters: the mean, variance and autocorrelation of the aggregate state. The estimated sample values are $\hat{\mu}_S = 13.18$, $\hat{\sigma}_S = 0.28$ and $\hat{\rho}_S = 0.79$.

5.3 Third step

The minimum distance estimator outlined above is now implemented to recover the dynamic parameters reported in the first panel (Model 1) of Table IV: the linear and quadratic investment cost, R&D sunk cost and exit value. Confidence intervals were constructed using the bootstrap.[10]

The values are estimated with the expected signs. Average sunk costs are estimated at about 3.2 million Euros, almost two years worth of revenues of an average firm and more than one year for an R&D firm. The exit value is estimated to be negative at -42 million Euros with a standard deviation of 17.5 million Euros. Although the unconditional mean exit value is negative, given its distribution, the actual conditional mean exit value is 1 million euros. Investment costs are increasing and convex but the quadratic component is not statistically significant. Alternative specifications where the quadratic investment cost term, μ_2, is dropped or a fixed operating cost, F, added are also reported in the second and third panels (Models 2 and 3). The precision of the linear cost improves when we drop the quadratic term. The estimate of 2.04 Euros for each Euro of investment suggests indirect investment costs are about 100%.

The cost function of Model 3 is no longer normalized 0 when there is inaction. This violation of the identification condition illustrates the difficulty of separately identifying $\hat{F}$ and $\hat{e}$.[11] Estimating the fixed operating cost in the first step using the available data on observed profits allows us to normalize the profits for the action of not exiting.

[10] Step 3 contains only simulation error. This error is reduced as the number of simulation draws increases ($\boldsymbol{n}_J \to \infty$).

[11] The estimated values of $\hat{\boldsymbol{e}} = 19.17;\ \hat{}^{e} = 5.66$ of model 3 and $\tilde{\boldsymbol{e}} = 13.85;\ \tilde{}^{e} = 6.02$ of model 2 with $= 0.93$ rationalize a "net" expected value $(1 - \ )(\boldsymbol{E}(\hat{\boldsymbol{e}}|\boldsymbol{exit}) - \boldsymbol{E}(\tilde{\boldsymbol{e}}|\boldsymbol{exit})) = 0.436$ where the probability of exit is 0.635%. This net expected value is close to the fixed cost estimate, $\hat{F} = 0.429$. Due to weak identification, different combinations of pairs $(\hat{F}; \hat{\boldsymbol{e}}; \hat{}^{e})$ can rationalize the observed decisions. For example, in a non-stochastic setting, a firm can decide to exit today and collect the average exit value $\bar{\hat{\boldsymbol{e}}}$ or stay one more period and collect $+\hat{F} + \ \bar{\hat{\boldsymbol{e}}}$. For an indifferent firm $+\hat{F} + \ \bar{\hat{\boldsymbol{e}}} = \bar{\hat{\boldsymbol{e}}}$ where $\bar{\hat{\boldsymbol{e}}}$ is the mean of $\hat{\boldsymbol{e}}$ conditional on exit. In the absence of fixed costs (F), the indifference condition is $+ \ \bar{\tilde{\boldsymbol{e}}} = \bar{\tilde{\boldsymbol{e}}}$ where $\bar{\tilde{\boldsymbol{e}}}$ is the mean of $\tilde{\boldsymbol{e}}$ conditional on exit. Replacing and solving the two conditions we get $\bar{\hat{\boldsymbol{e}}} - \bar{\tilde{\boldsymbol{e}}} = \frac{1}{1-\beta}\hat{F}$.

Parameters	$_1$	$_2$	^{1}	$^{\lambda,1}$	$\boldsymbol{e}^1$	^{e,1}	$\boldsymbol{F}^1$
Model 1							
Coefs	1.31	3.63	3.21	1.32	42.53	17.52	-
5th percentile	*-0.83*	*-10.92*	*1.35*	*0.40*	*1.30*	*0.00*	-
95th percentile	*5.05*	*10.62*	*17.86*	*8.34*	*77.12*	*31.97*	-
Model 2							
Coefs	2.04	-	2.14	0.51	13.85	6.02	-
5th percentile	*-0.25*	-	*1.19*	*0.17*	*1.95*	*0.00*	-
95th percentile	*4.14*	-	*14.57*	*6.72*	*57.62*	*32.66*	-
Model 3							
Coefs	2.05	-	1.74	0.25	19.17	5.66	0.43
5th percentile	*-0.20*	-	*0.78*	*0.04*	*-1117.91*	*0.00*	*-152.51*
95th percentile	*4.14*	-	*14.66*	*6.67*	*685.71*	*14.76*	*46.22*

Notes: Estimates for the dynamic parameters and bootstraped CI's
[1] Millions of euros

Table IV: Cost parameters estimates.

How reasonable are the estimated parameters? We can compare average profits earned by an R&D firm in the period before it started doing R&D with the average profits earned afterwards. The difference is $230,000$ Euros. This value discounted over a 50 year horizon is almost 3.2 million Euros, slightly below the estimated sunk costs from the structural model.

Notice that the estimated sunk costs cannot be compared with reported R&D expenditures. If they could, there would be no need to estimate them. There are many reasons why sunk costs of R&D will not show up in financial statements. One such reason is because the sunk costs can be related to internal resources (capital and labor) that have to be allocated to R&D instead of production and will thus not show up in company accounts. It is because such costs are unobseved that we need to adopt a structural model and use revealed preferences to recover them.

5.4 Robustness

Given the sensitivity of the parameter estimates to the first and second step, I now estimate the model under alternative specifications for the discount factor, profit function and policies.

Discount factor (Table V) Lower discount factors reduce the continuation value and, as expected, decrease the estimated investment and sunk costs. This relation is monotonic. For example, with a discount factor of 0.85 the sunk costs are estimated

at 1.28 million Euros and the investment costs at 0.93 for the linear and 1.64 for the quadratic component. This illustrates why the discount factor cannot be identified.

Discount factor	$_1$	$_2$	^{1}	$_{\lambda,1}$	$\mathbf{e}^1$	$_{e,1}$
0.850	0.93	1.64	1.28	0.48	13.38	5.59
0.875	1.01	2.12	1.68	0.66	18.11	7.53
0.900	1.11	2.76	2.30	0.95	25.50	10.56
0.925	1.23	3.67	3.28	1.44	38.03	15.69
0.950	1.42	4.64	4.95	2.30	61.10	25.11
0.975	1.92	4.25	7.66	3.69	106.17	43.45

Notes: [1] Millions of euros

Table V: Cost parameters estimates for different discount factors.

Profit function (Table VI) The results in Table VI show that, depending on the exact parametrization for the profit funcion, the sunk costs are estimated between 3.2 and 9.4 million Euros (while the average conditional sunk costs vary much less). Also, the investment costs are reduced when $\alpha_{fc,K}$ is removed from the fixed cost function. However, the results from Table II show that all the components of the profit function are quite significant suggesting a preference for the baseline case.

Model	$_1$	$_2$	^{1}	$_{\lambda,1}$	$\mathbf{e}^1$	$_{e,1}$
Baseline	1.31	3.63	3.21	1.32	42.53	17.52
$_{fc,0} =$ $_{fc,rd} =$ $_{fc,K} = 0$	-1.69	12.66	5.66	3.55	6.80	2.75
$_{fc,K} = 0$	-1.36	10.44	9.40	3.40	6.59	2.86

Notes: [1] Millions of euros

Table VI: Cost parameters estimates for different specifications of the profit function.

Policy functions (Table VII) The policy functions estimated in the second step play a central role in the estimation. Due to the nonlinearities, even small bias in the estimated policies can get magnified into a large bias in the structural parameters. To evaluate how sensitive the results are, I estimate the model with simpler (linear) policies. Using a linear investment function, sunk cost are estimated at about 4.5 million Euros and investment costs at 0.87 for the linear and 6.22 for the quadratic component. Alternatively, using a linear probit for R&D, sunk cost are estimated at 3.66 million Euros and adjustment costs at 1.59 for the linear and 2.4 for the quadratic component. If we allow both the investment and R&D policies to be linear, the sunk cost is estimated at 4.6 millions.

	$_1$	$_2$	^{1}	$^{\lambda,1}$	$\mathbf{e}^1$	^{e,1}
Linear R&D probit	1.59	2.40	3.66	1.49	38.07	15.68
Linear investment	0.87	6.22	4.52	2.07	50.30	20.69
Linear R&D probit and investment	1.34	3.93	4.60	1.92	46.51	19.11

Notes: [1] Millions of euros

Table VII: Cost parameters estimates for different policy functions.

After confirming that the model delivers relatively robust estimates, a final exercise is to evaluate how well the estimated model matches the data.

5.5 Validation

How well does the model fit the evidence? As opposed to other estimation methods, the minimum distance estimator does not match the aggregate moments.[12] Thus, I can use the comparison between the actual data and the simulated aggregate moments as a validity check. Before progress is made I must set some parameters: the distribution for the entry costs and the distribution for the productivity of entrants.

The parameters are set as follows. Mean and variance for the entry costs are set at 4 million Euros for the average entry cost and 1.75 millions for the standard deviation while the mean and variance for the productivity of entrants is 1.95 and 0.25. These number provide a good fit for the average number of firms and the entry and exit rates. The estimated standard deviation for sunk costs also have to be adjusted to 528,000 euros since otherwise the model would predict too much R&D.

The aggregate state is constructed using equation 5. The scalling constants for the aggregate state are $g1 = \exp(-11.75)$ and $g2 = \exp(-5.9)$.[13] The demand elasticity across countries is set at $\sigma_w = 4$ and smaller than the estimated within country elasticity of $\hat{\sigma} = 6.76$ implying that firms in Portugal are closer substitutes than firms outside Portugal.

Using the full set of parameters, we can now solve the model and simulate the industry for 100 periods. To do so, the model has to be solved for the new equilibrium industry evolution, $q^*(S_t, S_{t+1})$. The computational burden of solving a full dynamic game would

[12] BBL (2007) also propose a method of moments estimator to reduce the bias in the estimates by averaging out first stage estimation error.

[13] The statistics are reported for the empirical aggregate state $\left(\frac{Y^P}{\bar{P}}\right)$ and not for the theoretical aggregate state $\left(\frac{Y^P}{(P^P)^{1-\sigma}}\right)$ to make them comparable with the observed values. The model has been solved using the theoretical aggregate state (see the appendix for more details).

Moments	All	Data Pre- '97	Post- '97	Simulated
Average number of firms	680	536	753	740
Average log(Y/P)	13.18	12.94	13.35	13.34
Std. dev. log(Y/P)	0.28	0.37	0.10	0.21
Autocorrelation log(Y/P)	0.80	0.87	0.30	0.80
Percentage of R&D firms	0.21	0.13	0.29	0.34
Average productivity	2.21	2.15	2.24	2.22
Std. dev. of productivity	0.47	0.45	0.48	0.48
Average log-capital	12.77	12.43	13.00	13.00
Std. dev. of log-capital	1.57	1.49	1.58	1.88
Average investment rate	0.35	0.41	0.31	0.33
Std. dev. of inv. rate	0.61	0.65	0.59	1.29
Entry rate (%)	3.70	5.12	2.30	2.17
Exit rate (%)	-	-	-	1.86

Notes: Sample moments pre and post 1997 and moments for the simulated sample using the estimated parameters

Table VIII: Sample and simulated moments.

be prohibitive while the aggregate state model can be computed in a reasonable time.[14] After dropping the first 50 periods I calculate the moments for the stationary market structure, reported in Table VIII: the mean and standard deviation for productivity, capital, and investment rate, and the aggregate state. Also reported are the number of firms and the entry rates but these values are not free since they have been "calibrated" by setting the entry distribution.

Table VIII shows that the model can explain well the first and second moments for capital, productivity and investment rates. The average investment rates are very high both in the data and the model because for small firms it can be as high as 600%. The average 31% investement rate compares with the 16% median investment. Finally, the model predicts an entry rate of around 2%. Using these values as a benchmark, we can now evaluate the effects of policy changes.

6 Counterfactual Experiments

This section reports the results for a change in trade costs resulting from trade liberalization and assess the impact on industry R&D, productivity and investment. I also report two other experiments: changes in (i) sunk costs of R&D, and (ii) entry costs. Because

[14] A MATLAB algorithm to solve the model is available from the author. Solving the model takes about 40 minutes on a 2.27 Ghz Phenom double Quad-core computer with 12GB RAM.

the model is stylized, the particular numbers generated by the counterfactual simulations should be seen as merely suggestive as the effects of any policy change might depend on other factors not captured by the model.

Increase in trade costs The goal is to assess the models' predictions in case we return to the pre EU integration era, by evaluating the effects of an increase in trade costs. Since 80% of the firms export 90% of the production I have abstracted from the decision of selling at home or abroad. This shuts down the selection mechanism resulting from the interaction of heterogeneity with export costs that has been well discussed in the literature. On the other hand I focus on the effects of heterogeneity interacted with innovation costs. If export costs were significant so that only a fraction of the firms exported, the selection mechanism would reinforce the mechanism analyzed here.

The results in Table IX suggest that a 25% increase in trade costs will cause a reduction in average R&D, capital and productivity. The figures predicted by the model are similar to the values observed in the sample for the early periods. Also, entry and exit rates would increase. More importantly, however, while the number of firms decreases with the increase in trade costs the average size of a firm is now smaller. This is the important size effect documented in other studies (*e.g.* Eaton, Kortum and Kramarz (2008)) that is not captured by the models of Melitz (2003) or Eaton and Kortum (2003).

Trade costs	-25%	-10%	0%	10%	25%
Average number of firms	1130	835	745	680	625
Average log(Y/P)	14.71	13.84	13.32	12.87	12.23
Std. dev. log(Y/P)	0.12	0.17	0.22	0.24	0.29
Autocorrelation log(Y/P)	0.67	0.73	0.78	0.78	0.79
Percentage of R&D firms	0.52	0.41	0.34	0.27	0.18
Average productivity	2.26	2.23	2.22	2.20	2.18
Std. dev. of productivity	0.50	0.49	0.47	0.46	0.45
Average log-capital	14.42	13.44	13.01	12.67	12.30
Std. dev. of log-capital	1.44	1.87	1.88	1.86	1.72
Average investment rate	0.49	0.38	0.33	0.30	0.26
Std. dev. of inv. rate	1.66	1.40	1.29	1.20	1.08
Entry rate (%)	0.83	1.76	2.15	2.48	2.83
Exit rate (%)	0.45	1.39	1.84	2.25	2.67

Table IX: Predicted moments from changes in trade costs.

The mechanism causing the size effect illustrates the need of physical capital and adjustment costs in the model. Models without adjustment costs predict no effect of trade on firm size. As the simulated results show, this no longer holds in the general

case. Aggregate shocks together with adjustment costs for capital create an option value for investment and generate two effects from changes in trade costs. First, reductions in trade costs push investment up. The equilibrium average capital stock becomes larger which, in the presence of sunk costs, leads to an increase in R&D. Second, as trade costs are reduced more firms see it profitable to enter in the market. In equilibrium, the reduction in trade costs leads to an increase in the number and the size of firms in the market.

Reduction in the sunk costs of R&D The sunk costs of R&D play an important role shaping productivity and trade. A reduction of these sunk costs can also promote innovation and trade. Some examples of policies to reduce the sunk costs would be direct subsidies to R&D start-up or the creation of public R&D labs to explore economies of scale and avoid duplication costs. The results in Table X show that a reduction in sunk costs is expected to lead to an increase in R&D, productivity, capital and investment. More importantly average sales increase which suggests that reducing the sunk costs of R&D can promote firm exports. These effects are similar to a reduction in trade costs and can also rationalize the observed industry change.

Sunk costs	-50%	-25%	0%	25%	50%
Average number of firms	904	781	740	722	711
Average log(Y/P)	13.37	13.43	13.34	13.24	13.16
Std. dev. log(Y/P)	0.18	0.19	0.21	0.22	0.23
Autocorrelation log(Y/P)	0.84	0.82	0.80	0.78	0.76
Percentage of R&D firms	0.74	0.42	0.34	0.27	0.21
Average productivity	2.23	2.23	2.22	2.20	2.19
Std. dev. of productivity	0.48	0.48	0.48	0.47	0.46
Average log-capital	13.25	13.15	13.00	12.91	12.84
Std. dev. of log-capital	1.84	1.88	1.88	1.87	1.86
Average investment rate	0.35	0.34	0.33	0.32	0.31
Std. dev. of inv. rate	1.32	1.32	1.29	1.25	1.23
Entry rate (%)	1.56	1.98	2.17	2.27	2.34
Exit rate (%)	1.20	1.65	1.86	1.98	2.06

Table X: Predicted moments from changes in sunk costs of R&D.

Reducing the sunk costs affects only the value of non-R&D firms and leaves the value of R&D firms unchanged (excluding equilibrium effects). Changes in sunk costs affect the value function through three different mechanisms. First, smaller sunk costs cause a direct reduction in costs. Second, the probability of doing R&D increases with a decrease in sunk costs having an indirect effect and increasing expected benefits. Finally, the

equilibrium effects have an opposite effect on the value function. As sunk costs decrease, the value of entering is increased and more firms enter the market leading to an increase in the number of firms. The decrease in sunk costs leads to an increase in R&D and productivity. As a consequence, firms invest more and become larger. The equilibrium effect is not sufficient to counteract the first two effects and the final result of a decrease in the sunk costs is more firms that are both larger and more productive.

Increase in entry costs Lastly, there is often political support for the creation of large firms under the argument that this might spur innovation. Large firms also account disproportionately for total exports. One way to create large firms is by increasing entry costs. Large entry costs will decrease entry rates and protect incumbent firms from competition, allowing them to grow. Larger firms will then be willing to pay the sunk cost and start innovating.

As reported in Table XI, an increase in entry costs boosts innovation by generating an increase in average size, investment and R&D. However, it also leads to a reduction in the number of firms and entry/exit rates and bad firms are now less likely to exit and be replaced by better firms. Overall, this selection effect might lead to a decrease in average industry productivity, specially if the boost in innovation is not sufficiently strong (or all firms are already innovating). As expected, the increase in entry costs protects firms and leads to an increase in average size.

entry costs	-25%	-10%	0%	10%	25%
Average number of firms	1,129	911	743	589	413
Average log(Y/P)	12.92	13.12	13.33	13.61	14.07
Std. dev. log(Y/P)	0.19	0.20	0.21	0.21	0.22
Autocorrelation log(Y/P)	0.93	0.85	0.80	0.75	0.78
Percentage of R&D firms	0.27	0.30	0.34	0.39	0.49
Average productivity	2.20	2.21	2.22	2.23	2.25
Std. dev. of productivity	0.47	0.47	0.48	0.48	0.49
Average log-capital	12.72	12.85	13.01	13.29	13.77
Std. dev. of log-capital	1.84	1.86	1.88	1.88	1.81
Average investment rate	0.30	0.31	0.33	0.36	0.41
Std. dev. of inv. rate	1.21	1.24	1.29	1.35	1.48
Entry rate (%)	2.33	2.29	2.16	1.87	1.27
Exit rate (%)	2.18	2.04	1.84	1.55	1.08

Table XI: Predicted moments from changes in entry costs.

7 Conclusion

The empirical evidence suggests that size is an important driver of endogenous technological choices. I have proposed and estimated an equilibrium model with endogenous choices of size and technology for the Portuguese Moulds industry. I find evidence of relatively large sunk costs of R&D that play an important role when there is trade liberalization. Getting access to foreign markets allows the firms to exploit the economies of scale in R&D and invest in physical capital. Sunk costs then explain why R&D is done by larger firms. The two results explain the observed performance of some industries after trade liberalization episodes and the important role for size.

Some challenging questions are left for future research. The model proposed here is partial equilibrium. A trade liberalization event will certainly trigger other general equilibrium effects on wages and returns on capital. How these effects affect the innovation decisions is an interesting but difficult question to address empirically. Second, the model has abstracted from import competition and strategic interactions. While this is an advantage of the particular industry studied here and allowed me to focus on the trade-induced innovation mechanism, it is not the general case. Evaluating the effects of trade in other industries/settings necessarily means taking the selection and the strategic effects into account but requires further features to disentangle and identify the different mechanisms.

A Appendix

A.1 Sample construction and descriptive statistics

The sample is constructed from three sources: Aggregate variables (revenues, value added, employment) from the Portuguese National Statistics Office (INE); industry price deflators from IAPMEI (2006); and the firm level accounting data from the Bank of Portugal's database (*Central de Balancos*, five digit NACE code industry 29563 - Moulds Industry).

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

1999 and 2000. The data sampling scheme can be accounted for in the structural model. The dataset contains detailed financial information and other characterization variables: number of workers, total exports, R&D, founding year and current status.

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

There are 1,290 observations for 231 firms, out of which 265 observations with positive R&D corresponding to 59 firms and 49 R&D start-ups (Table A.II). Average firm has revenues of 1.5 million Euros and employs 32 workers with a labor productivity of 20,381 Euros. The industry is populated by many small and medium firms and no market leaders. R&D firms are three times larger, 20% more productive and export more (Table A.I).

Variable construction:

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

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

Value added is equal to revenues 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.

Cash flow is constructed as the sum of accounting profits plus depreciation and amortization.

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

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

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

	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 R&D 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%
R&D 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%
R&D to sales ratio	0.9%	3.4%	0.0%	46.5%

Source: Central de Balanços, Bank of Portugal

Table A.I: Summary statistics.

There are few observations on entry and exit. Due to the way the data was collected some firms might not be reported in the sample but still be active in the industry complicating the identification of entry and exit. Firms could have been operating in the market before first appearing in the sample and they might still be operating after leaving the

sample. The problem is addressed with two variables that help to identify entry and exit. For entry, firms report their founding year and it was considered to be a true entry if the firm first appeared in the sample within a 2 year window from the reported founding year (values reported in Table A.II under the column "entry"). Regarding exits, the central bank collects a variable for the "status" of the firm. The quality of the exit variable is poor and some firms might have closed down and still be reported as "active", so only a fraction of actual exits is captured. A total of 48 entries and 7 exits from the panel can be identified.

Year	*Number of firms*	*Number of non R&D firms*	*Number of R&D firms*	*R&D 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**

Table A.II: R&D, Entry and Exit: number of observations per year.

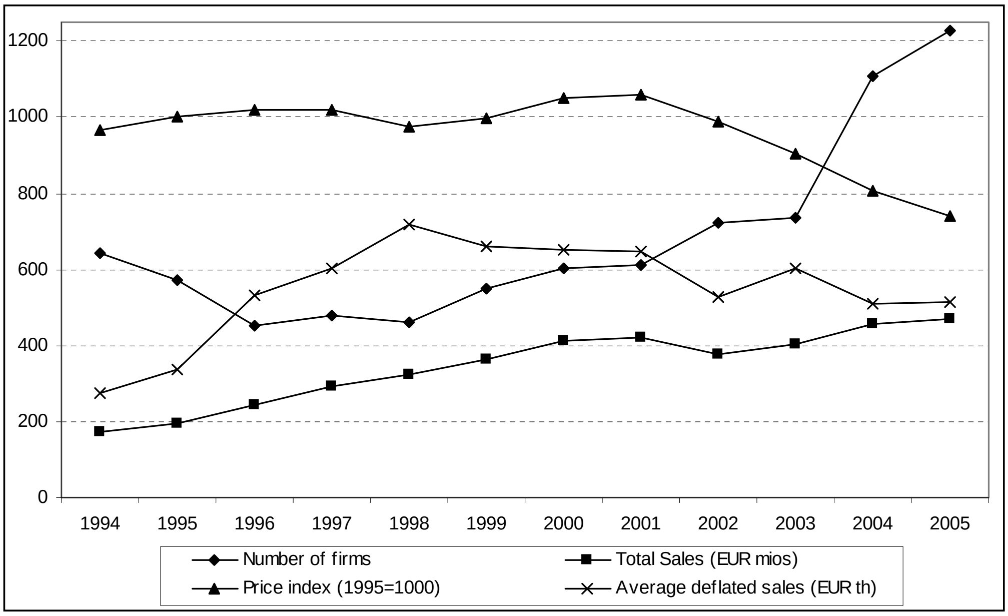
A.1.1 Aggregate State

Aggregate state definition: The aggregate state is defined as average deflated total industry revenues: $\frac{Y/N}{\widetilde{P}}$. It can be divided into three components. In Figure A.1.1 I plot the evolution of the different components. The market was growing between 1993 and 2000 and stopped until 2003. The number of firms share a similar pattern. On the other side prices were increasing slightly and decreased in the later years. The industry grew substantially after 1994 due to the strong increase in demand for Portuguese moulds. 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 the three variables is summarized by the evolution of the single index variable, $\frac{\widetilde{Y}}{\widetilde{P}}$(average deflated revenues).

	Number of firms	Production (EUR mio)	Exports (EUR mio)	Exports % of sales	Employm.	Val. 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 A.III: Aggregate variables



Aggregate variables: Number of firms, price index, average and total sales

Stationarity Over the period, production grew at an average 8.9% per year and labor productivity at 5.7% (Tables A.III and A.I) raising some concerns over non-stationarity. Production grows substantially over 1994 to 1998 period and stabilizes thereafter (Figure A.1.1). This pattern is consistent with a structural shock occurring somewhere before 1994 with the observed years of 1994 to 1998 being a period of transitional dynamics. The structural shock was the European economic integration that culminated with the establishment of the *Common Market* in 1993. Thus, it is likely that the data covers a

period of transitional dynamics and convergence to the new equilibrium.

References

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

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

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

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

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

[6] Bernard, A., J. Eaton, B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade", *American Economic Review*, 93(4), 1268-1290.

[7] Bernard, A., J. Jensen and P. Schott (2006): "Trade costs, Firms and Productivity", *Journal of Monetary Economics*, 53(5), 917-937.

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

[9] Constantini, J., and M. Melitz (2008): "The Dynamics of Firm-Level Adjustment to Trade Liberalization," in, *The Organization of Firms in a Global Economy*, eds. E. Helpman, D. Marin, and T. Verdier (eds.), Harvard University Press

[10] Cooper, R., and J. Haltiwanger (2006); "On the Nature of the Capital Adjustment Costs", *Review of Economic Studies*, 73(3), 611-633.

[11] De Loecker, J. (forthcoming). "Product Differentiation, Multi-product Firms and Estimating the Impact of Trade Liberalization on Productivity", *Econometrica.*

[12] Doraszelski, U., and M. Satterthwaite (2010): "Computable Markov-Perfect Industry Dynamics", *Rand Journal of Economics*, 41(2), 215-243.

[13] Eaton, J., and S. Kortum (2002): "Technology, Geography and Trade", *Econometrica*, 70(5), 1741-1779.

[14] Eaton, J., S. Kortum, and F. Kramarz (2008): "An Anatomy of International Trade: Evidence from French Firms", Working Paper WP14610, NBER.

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

[16] Fershtman, C., and A. Pakes (2009): "Finite State Dynamics with Asymmetric Information: A Framework for Applied Work", Working Paper, Harvard University.

[17] Gandhi, A., S. Navarro and D. Rivers (2011): "Value Added Bias and the Identification of Production Functions: How heteogeneous is Productivity", University of Western Ontario, CIBC Centre for Human Capital and Productivity Working Papers 20119

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

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

[20] Jovanovic, B. (1982): "Selection and the Evolution of Industry", *Econometrica*, 50(3).

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

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

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

[24] Lopez, R. (2005): "Trade and growth: Reconciling the macroeconomic and microeconomic evidence", *Journal of Economic Surveys* 19, 623-648.

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

[26] Melitz, M. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71(6) 1695-1725.

[27] Melitz, M., and G. Ottaviano (2008): "Market Size, Trade and Productivity", *Review of Economic Studies*, 75(1), 295-316

[28] Olley, S., and A. Pakes (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry”, *Econometrica*, 64(6), 1263-98.

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

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

[31] Rust, J. (1994): "Structural Estimation of Markov Decision Processes" in *Handbook of Econometrics*, vol. 4, ch. 51, ed. by C. Engel and D. McFadden, North Holland, 3082–3139.

[32] Ryan, S. (2008): "The Costs of Environmental Regulation in a Concentrated Industry", Working Paper 0510, MIT.

[33] Schmidt-Dengler, P. (2006): "The Timing of New Technology Adoption: The Case of MRI", Working Paper, London School of Economics.

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

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

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

[37] 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.

B Technical Appendix (not for publication)

B.1 Profit Function

The first order conditions are

$$\begin{aligned} \alpha_m \tfrac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m} \tfrac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma} \left(P^P\right)^{(\sigma-1)/\sigma} &= P_m M_i \\ \alpha_l \tfrac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m} \tfrac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma} \left(P^P\right)^{(\sigma-1)/\sigma} &= w L_i \end{aligned}$$

Replacing this in the objective function we get

$$\Pi^G = \left[1 - \frac{\sigma-1}{\sigma}(\alpha_m + \alpha_l)\right]\left[A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m} \frac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma}$$

From the first order conditions we can solve for M and L

$$\left(\frac{\alpha_m}{P_m}\frac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} \frac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma}\right)^{\alpha_m(\sigma-1)/[\sigma-\alpha_m(\sigma-1)]} = M_i^{\alpha_m(\sigma-1)/\sigma}$$

$$\left(\frac{\alpha_l}{w}\frac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} M_i^{\alpha_m} \frac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma}\right)^{\alpha_l(\sigma-1)/[\sigma-\alpha_l(\sigma-1)]} = L_i^{\alpha_l(\sigma-1)/\sigma}$$

or

$$\begin{aligned} \left(\tfrac{\alpha_m}{P_m}\tfrac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} \tfrac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma}\right)^{\frac{\sigma}{[\sigma-\alpha_l(\sigma-1)-\alpha_m(\sigma-1)]}} &= M_i \\ \left(\tfrac{\alpha_l}{w}\tfrac{\sigma-1}{\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k} \tfrac{1}{\tau}\right]^{(\sigma-1)/\sigma} (Y^P)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma}\right)^{\frac{\sigma}{[\sigma-\alpha_l(\sigma-1)-\alpha_m(\sigma-1)]}} &= L_i \end{aligned}$$

Replacing back in the objective function we get the final profit function

$$\Pi^G = \frac{1}{\gamma}\left[\left(\frac{\alpha_l}{w}\right)^{\alpha_l}\left(\frac{\alpha_m}{P_m}\right)^{\alpha_m}\left(\frac{\sigma-1}{\sigma}\right)^{\frac{\sigma}{\sigma-1}}\right]^{\gamma}\left[A_0 e^{\omega_i} K_i^{\alpha_k}\frac{1}{\tau}\right]^{\gamma}\left(\frac{Y^P}{(P^P)^{1-\sigma}}\right)^{\frac{\gamma}{\sigma-1}}$$

where $\gamma = \frac{(\sigma-1)}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}$. Alternatively

$$\frac{\Pi^G}{P^P} = \frac{1}{\gamma}\left[\left(\alpha_l\frac{P^P}{w}\right)^{\alpha_l}\left(\alpha_m\frac{P^P}{P_m}\right)^{\alpha_m}\left(\frac{\sigma-1}{\sigma}\right)^{\frac{\sigma}{\sigma-1}}\right]^{\gamma}\left[A_0 e^{\omega_i} K_i^{\alpha_k}\frac{1}{\tau}\right]^{\gamma}\left(\frac{Y^P}{P^P}\right)^{\frac{\gamma}{\sigma-1}}$$

Finally using the demand function the gross price for firm i is

$$P_i = \left[c_{\alpha_l,\alpha_m,\sigma} A_0 e^{\omega_i} K_i^{\alpha_k} \tau^{-(\alpha_m+\alpha_l)}\right]^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}} \left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{\frac{[1-(\alpha_l+\alpha_m)]}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}} \tag{A.1}$$

where $c_{\alpha_l,\alpha_m,\sigma} = \left(\frac{\alpha_l}{w}\frac{\sigma-1}{\sigma}\right)^{\alpha_l}\left(\frac{\alpha_m}{P_m}\frac{\sigma-1}{\sigma}\right)^{\alpha_m}$. First notice that

$$[A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m} \tau^{-1}]^{-1/\sigma} = \left(c_{\alpha_l,\alpha_m,\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k}\tau^{-1}\right]\left(\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{1/\sigma}\right)^{(\alpha_l+\alpha_m)}\right)^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

So that

$$P_i = Q_i^{-1/\sigma}\left(Y^P\right)^{1/\sigma}\left(P^P\right)^{(\sigma-1)/\sigma} = \left(A_0 e^{\omega_i} K_i^{\alpha_k} L_i^{\alpha_l} M_i^{\alpha_m}\tau^{-1}\right)^{-1/\sigma}\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{1/\sigma}$$

$$= \left(c_{\alpha_l,\alpha_m,\sigma}\left[A_0 e^{\omega_i} K_i^{\alpha_k}\tau^{-1}\right]\left(\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{1/\sigma}\right)^{(\alpha_l+\alpha_m)}\right)^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{1/\sigma}$$

$$= \left[c_{\alpha_l,\alpha_m,\sigma} A_0 e^{\omega_i} K_i^{\alpha_k}\tau^{-1}\right]^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{\frac{[1-(\alpha_l+\alpha_m)]}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

C Aggregate State

Using the optimal price for firm i derived in (A.1)

$$P_i = \left(\left(c_{\alpha_l,\alpha_m,\sigma} A_0 e^{\omega_i} K_i^{\alpha_k}\left(e^{\varepsilon_i^d}\right)^{[(\alpha_l+\alpha_m)-1]}\tau^{-1}\right)^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}\left(\frac{Y^P}{(P^P)^{\sigma-1}}\right)^{\frac{[1-(\alpha_l+\alpha_m)]}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}\right) \tag{A.2}$$

The price index is

$$P^P = \left[\sum_{i\in P} P_i^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

Let

$$p1 = \left[c_{\alpha_l,\alpha_m,\sigma}\right]^{\frac{-1}{[\sigma-\alpha_l(\sigma-1)-\alpha_m(\sigma-1)]}}\tau^{\frac{1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

$$p2 = \left(P^P\right)^{\frac{(1-(\alpha_l+\alpha_m)](\sigma-1)}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

$$p3 = \left(\left[A_0 e^{\omega_i} K_i^{\alpha_k}\left(e^{\varepsilon_i^d}\right)^{[(\alpha_l+\alpha_m)-1]}\right]\right)^{\frac{-1}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

$$p4 = \left(Y^P\right)^{\frac{(1-(\alpha_l+\alpha_m)]}{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}}$$

So that we can write

$$P_i = p1p2p3p4$$

Solving for the price index

$$P^P = \left[\sum (p1p3p4)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}} \tag{A.3}$$

From the demand specified in (2)

$$Y^P = Q_0 Y^W \left(\frac{P^P}{P^W}\right)^{1-\sigma_w} \tag{A.4}$$

Replacing the price index from above and solving

$$\begin{aligned} Y^P &= \left(Q_0 Y^W \left(\frac{1}{P^W}\right)^{1-\sigma_w}\right)^{\frac{1}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left(\left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}}\right)^{\frac{1-\sigma_w}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \\ &= (p5')^{\frac{1}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left(\left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}}\right)^{\frac{1-\sigma_w}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \end{aligned}$$

where $p5 = Q_0 Y^W \left(\frac{1}{P^W}\right)^{1-\sigma_w}$

Finally replacing the expression for Y^P in the price index results

$$\begin{aligned} P^P &= \left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}} \left(Y^P\right)^{(1-(\alpha_l+\alpha_m)]} \\ &= \left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}\left(\frac{1}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}\right)} (p5)^{\frac{(1-(\alpha_l+\alpha_m)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \end{aligned}$$

So we have the two equations for the aggregate revenues and price index

$$Y^P = (p5)^{\frac{1}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left(\left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}}\right)^{\frac{1-\sigma_w}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$$

and

$$P^P = \left[\sum (p1p3)^{1-\sigma}\right]^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{1-\sigma}\left(\frac{1}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}\right)} (p5)^{\frac{(1-(\alpha_l+\alpha_m)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$$

Notice that we can drop the term $\left(e^{\varepsilon_i^d}\right)^{[(\alpha_l+\alpha_m)-1]}$ from $p3$. Because of independence this term becomes mean neutral and the variance converges to zero as the number of players increases.

Finally, the aggregate state is

$$\begin{aligned}
\frac{Y^P}{(P^P)^{1-\sigma}} &= (p5)^{\frac{(\sigma-(\alpha_l+\alpha_m)(\sigma-1))}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left[\sum (p1p3)^{1-\sigma}\right]^{\frac{1}{1-\sigma}\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}(\sigma-\sigma_w)\right)} \\
&= g1.\tau^{\frac{(\sigma-\sigma_w)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left[\sum ((A_0 e^{\omega_i} K_i^{\alpha_k})^\gamma)\right]^{\frac{1}{\gamma}\left(\frac{(\sigma_w-\sigma)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}\right)}
\end{aligned}$$

where $g1 = \left(Q_0 Y^W \left(\frac{1}{P^W}\right)^{1-\sigma_w}\right)^{\frac{(\sigma-(\alpha_l+\alpha_m)(\sigma-1))}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left[c_{\alpha_l,\alpha_m,\sigma}\right]^{\frac{-(\sigma-\sigma_w)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$.

Note that if $\sigma_w = \sigma$ the aggregate state is completely determined by the world state, i.e. $\frac{Y^P}{(P^P)^{1-\sigma}} = Q_0 Y^W \left(\frac{1}{P^W}\right)^{1-\sigma_w}$. Clearly only two coefficients matter: the elasticity of the aggregate state with respect to τ and the elasticity of the aggregate state with respect to $[\sum (e^{\omega_i} K_i^{\alpha_k})^\gamma]$.

Finally we can define the observed price index

$$\bar{P} = \sum_{i \in P} P_i = \left[\sum (p1p3)\right]^{\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}\right)} (p5)^{\frac{(1-(\alpha_l+\alpha_m)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$$

and the corresponding observed aggregate state

$$\begin{aligned}
\frac{Y^P}{\bar{P}} &= (p5)^{\frac{(\alpha_l+\alpha_m)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left(p1^{-\sigma_w}\left[\sum (p3)^{1-\sigma}\right]^{\frac{1-\sigma_w}{1-\sigma}} \left[\sum (p3)\right]^{-1}\right)^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \\
&= g2.\tau^{\frac{-\sigma_w}{[\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left(\frac{\left[\sum (([A_0 e^{\omega_i} K_i^{\alpha_k}])^\gamma)\right]^{\frac{1-\sigma_w}{1-\sigma}}}{\left[\sum (A_0 e^{\omega_i} K_i^{\alpha_k})^{\frac{\gamma}{1-\sigma}}\right]}\right)^{\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}
\end{aligned}$$

where $g2 = \left(Q_0 Y^W \left(\frac{1}{P^W}\right)^{1-\sigma_w}\right)^{\frac{(\alpha_l+\alpha_m)}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}} \left[c_{\alpha_l,\alpha_m,\sigma}\right]^{\frac{-\sigma_w}{(\sigma_w-(\alpha_l+\alpha_m)(\sigma_w-1)]}}$.

C.1 Relation between the derived aggregate state and the observed aggregate state

Notice that the aggregate state observed in the data is $\frac{Y^P}{\sum P_i}$ and not $\frac{Y^P}{(P^P)^{1-\sigma}}$ or $\frac{Y^P}{P^P}$. To compare the observed state with $\frac{Y^P}{(P^P)^{1-\sigma}}$ notice that

$$\frac{\frac{Y^P}{\sum P_i}}{\frac{Y^P}{\left(P^P\right)^{1-\sigma}}} = \frac{\left(P^P\right)^{1-\sigma}}{\sum P_i} = \left[\frac{\sum(p3)^{1-\sigma}}{\sum(p3)}\right]^{\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}\right)} p1^{\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}\right)} (p5)^{-\frac{\sigma(1-(\alpha_l+\alpha_m))}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}}$$

so

$$\frac{Y^P}{\sum P_i} = \frac{Y^P}{(P^P)^{1-\sigma}} \left[\frac{\sum(p3)^{1-\sigma}}{\sum(p3)}\right]^{\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}\right)} p1^{\left(\frac{[\sigma-(\alpha_l+\alpha_m)(\sigma-1)]}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}\right)} (p5)^{-\frac{\sigma(1-(\alpha_l+\alpha_m))}{(\sigma w-(\alpha_l+\alpha_m)(\sigma w-1)]}}$$

This discrepancy only affects the first stage of the estimation. In order for the estimates in the first stage (namely the production coefficients and the demand elasticity) to be consistent it is required that $\left(\log\left[\sum (p3)^{1-\sigma}\right] - \log\left[\sum (p3)\right]\right)$ be uncorrelated with the observed variables. This is true for some known distributions. For example the log-normal case, when $p3\tilde{}\log N(a_1, a_2)$ then $\log\left[\sum (p3)\right] \approx (\log N + a_1 + \frac{a_s^2}{2})$ and $\log\left[\sum (p3^{1-\sigma})\right] \approx \log N + (1-\sigma)\left(a_1 + (1-\sigma)\frac{a_2^2}{2}\right)$. So that $\left(\log\left[\sum (p3)^{1-\sigma}\right] - \log\left[\sum (p3)\right]\right) \approx -\sigma\left(a_1 + [2-\sigma]\frac{a_2^2}{2}\right)$ is a constant.

C.2 Algorithm convergence

The algorithm for solving the model involves an outer loop to find the fixed point equilibrium transition for the aggregate state. For particular parameter values the algorithm diverges if a naive strategy is used in the update step. This is related with the slope of the function that should be in absolute value smaller then 1 in order to guarantee convergence. This slope is related with the elasticity of the aggregate state with respect to $\sum\left(\left[A_0 e^{\omega_i} K_i^{\alpha_k}\right]\right)^{\gamma}$ and is thus easy to analyze for particular parameter values. The adoption of alternative methods (like Newton's (or secant) method, interpolation, etc) help to solve the problem. The method used was the secant method, i.e. the state is updated according to the formula $s_n = s_{n-1} - \left[s_{n-1} - g(s_{n-1})\right] \frac{s_{n-1} - s_{n-2}}{s_{n-1} - g(s_{n-1}) - [s_{n-2} - g(s_{n-2})]}$

D Production Function Estimation

Using the specifications for the production function, demand and the assumption on the productivity transition function we can write the following three equations

$$q = \omega + \alpha_l l + \alpha_k k + \alpha_m m \tag{A.5}$$

$$p = -\frac{1}{\sigma}q + \frac{1}{\sigma}(y^p - p^p) + p^p + \varepsilon^d \tag{A.6}$$

$$\omega_{t+1} = \left(\gamma_{r,0} + \gamma_{r,1}\omega_t + \gamma_{r,2}\omega_t^2 + \gamma_{r,3}\omega_t^3\right) R_t + \left(\gamma_{nr,0} + \gamma_{nr,1}\omega_t + \gamma_{nr,2}\omega_t^2 + \gamma_{nr,3}\omega_t^3\right)(1-R_t) + \eta_t \tag{A.7}$$

Putting (A.5) and (A.6) together

$$y - p^p = p + q - p^p = \frac{\sigma-1}{\sigma}\left(\omega + \alpha_l l + \alpha_k k + \alpha_m m\right) + \frac{1}{\sigma}(y^p - p^p) + \varepsilon^d$$

where the demand shock is not observed by the firm when setting the level of production. From profit maximization we can write

$$\alpha_m = \frac{\sigma}{\sigma-1}\mu_{m,i,t}e^{\varepsilon^d_{it}}$$

where $\mu_{m,i,t} = \frac{P_M M_{i,t}}{P_{i,t}Q_{i,t}}$ which we can estimate by

$$\log(\mu_{m,i,t}) = c - \varepsilon^d_{i,t}$$

with $c = \log(\frac{\sigma-1}{\sigma}) + \log(\alpha_m)$. We can finally replace this above and solve for productivity

$$\omega(\alpha) = \frac{\sigma}{\sigma-1}\left[(y-p^p) - e^{\hat{c}}m - \frac{\sigma-1}{\sigma}\left(\alpha_l l + \alpha_k k\right) - \frac{1}{\sigma}(y^p - p^p) - \varepsilon^d\right]$$

and from (A.7)

$$\begin{aligned}\eta_{i,t}(\alpha,\gamma_p) \quad = \quad & \omega_{i,t+1}(\alpha) - \left(\gamma_{r,0} + \gamma_{r,1}\omega_{i,t}(\alpha) + \gamma_{r,2}\omega_{i,t}(\alpha)^2 + \gamma_{r,3}\omega_{i,t}(\alpha)^3\right) R_{i,t} \\ & - \left(\gamma_{nr,0} + \gamma_{nr,1}\omega_{i,t}(\alpha) + \gamma_{nr,2}\omega_{i,t}(\alpha)^2 + \gamma_{nr,3}\omega_{i,t}(\alpha)^3\right)(1 - R_{i,t})\end{aligned}$$

where $\alpha = (\alpha_l, \alpha_k, \sigma)$ and $\gamma_p = (\{\{\gamma_{i,j}\}_{i=r,nr}\}_{j=0}^3)$. We can use the moment condition $E(\eta(\alpha)z) = 0$ to form

$$E(\eta(\alpha,\gamma_p)z)'WE(\eta(\alpha,\gamma_p)z)$$

where z is a vector of instruments that are uncorrelated with η and W is the weighting matrix. In my application the vector of instrument is the current capital stock, labor and

their interaction, together with a cubic polinomial on $\omega(\hat{\alpha})$ and RD where $\hat{\alpha}$ is a first step consistent estimate of α.

I use the k-step GMM estimator

$$\min_{\alpha,\gamma_p}\left(\frac{1}{N}\sum_{i,t}\eta_{i,t}(\alpha,\gamma_p)z_{i,t}\right)' W\left(\frac{1}{N}\sum_{i,t}\eta_{i,t}(\alpha,\gamma_p)z_{i,t}\right)$$

where $W=\left(\frac{1}{N}\sum_{i,t}z_{i,t}z_{i,t}'\right)^{-1}$ in the first step and $W=\left(\frac{1}{N}\sum_{i,t}\eta_{i,t}^2(\hat{\alpha},\hat{\gamma}_p)z_{i,t}z_{i,t}'\right)^{-1}$ in the further steps. Given the efficient weighting matrix, the covariance matrix for the estimator is

$$V(\hat{\alpha})=\left[\left(\sum\frac{\partial\eta_{i,t}(\alpha,\gamma_p)}{\partial(\alpha,\gamma_p)}z_{i,t}\right)\left(\sum_{i,t}\eta_{i,t}^2(\hat{\alpha},\hat{\gamma}_p)z_{i,t}z_{i,t}'\right)^{-1}\left(\sum\frac{\partial\eta_{i,t}(\alpha,\gamma_p)}{\partial(\alpha,\gamma_p)}z_{i,t}\right)'\right]^{-1}$$

once we estimate the parameters of the production function above we can construct

$$\hat{\alpha}_m=\frac{\hat{\sigma}}{\hat{\sigma}-1}e^{\hat{c}}$$

E The aggregate state dynamic model

This section describes the elements of the general aggregate state dynamic model and characterizes the equilibrium conditions when states are discrete. I draw on Pesendorfer and Schmidt-Dengler (2008) and extend some results for the aggregate state model.

E.1 States and actions

Time is discrete, $t=1,2,...,\infty$. There are N players where a player is denoted by $i\in\{1,...,N\}$.

States Agents are endowed with a state variable $s_{it}\in\mathbf{S}_i=\{1,...L\}$ and a vector of payoff shocks $\boldsymbol{\varepsilon}_{it}\in\mathbb{R}^K$. Both the state and the shocks are privately observed by the players. The econometrician observes the states, s_{it}, but not the payoff shocks, $\boldsymbol{\varepsilon}_{it}$.

The industry state is $\mathbf{s}_t=(s_{1t},...,s_{Nt})\in\mathbf{S}=S_i^N$. The cardinality of the state space is $m_s=L^N$. The state of competitors $s_{-it}=(s_{1t},..s_{i-1,t},s_{i+1,t},..,s_{Nt})\in\mathbf{S}_{-i}=S_i^{N-1}$ with cardinality $m_{s-1}=L^{N-1}$. The vector of payoff shocks is drawn independently from the strict monotone and continuous distribution $F(.|\mathbf{s}_t)$.

Actions Each player chooses an action $a_{it} \in A_i = \{0, .., K\}$. Decisions are simultaneous after players observe their state and payoff shock. A profile $\mathbf{a}_t = (a_{1t}, ..., a_{Nt}) \in \mathbf{A} = A_i^N$ and $\mathbf{a}_{-it} \in \mathbf{A}_{-i} = A_i^{N-1}$. The cardinality of the action space $\mathbf{A}$ is $m_a = (1+K)^N$ and $\mathbf{A}_{-i}$ is $m_{a-1} = (1+K)^{N-1}$.

State transition The individual state transition is described by a density function $p : A_i \times S_i \times S_i \to [0, 1]$. A typical element of $p(a, s, s')$ equals the probability that state s' is reached from state s when action a is chosen and $\sum_{s' \in S_i} p(a, s, s') = 1$. The industry transition can be represented by a density function $g : \mathbf{A} \times \mathbf{S} \times \mathbf{S} \to [0, 1]$. A typical element of $g(\mathbf{a}, \mathbf{s}, \mathbf{s}') = \prod_{j=1}^{N} p(a_j, s_j, s'_j)$ equals the probability that state $\mathbf{s}'$ is reached from state $\mathbf{s}$ when action profile $\mathbf{a}$ is chosen and $\sum_{\mathbf{s}' \in \mathbf{S}} g(\mathbf{a}, \mathbf{s}, \mathbf{s}') = 1$.

Per period payoff Firms receive per period returns which depend on the state of the industry, current actions and shocks $(\pi(\mathbf{a}_t, \mathbf{s}_t, \boldsymbol{\varepsilon}_{it}))$.

Assumption E.1 *(a) There exists a function $(S : \mathbf{S} \to J)$ that maps the vector of firm's individual states $(\mathbf{s}_t)$ into an aggregate index $(S(s_{1t}, .., s_{Nt}) \in J)$. The cardinality of J is $m_J \leq m_s$.*

(b) Per period payoffs can be written as

$$\pi(\mathbf{a}_t, \mathbf{s}_t, \boldsymbol{\varepsilon}_{it}) = \pi(\mathbf{a}_t, s_{it}, S_t) + \sum_{k=1}^{K} \varepsilon_{it}^k \cdot 1(a_{it} = k)$$

Under this assumption, S_t is the payoff relevant variable commonly observed by all agents. Notice that the payoff relevant shocks $(\boldsymbol{\varepsilon}_{it})$ have no impact on the stage game pricing. One type of demand which meets this assumption is Monopolistic Competition.

Assumption E.2 *(a) (Incomplete Information) Individual states and actions are private information and;*

(b) (Markov beliefs) players use the current state to form their beliefs: $\sigma(\mathbf{s}|S, s_i)$.

Under Assumption E.2 the only common information is the aggregate state. Moreover, it restricts agents to form beliefs using only the current state. This is the main assumption that distinguishes the aggregate state model from the literature with no privately observed state variables. The restriction on beliefs is important to circumvent the learning problem that would emerge otherwise.

Strategies I consider pure symmetric Markovian strategies, $a_{it}(s_{it}, S_t, \boldsymbol{\varepsilon}_{it})$. These strategies depend only on payoff relevant variables. Using symmetry we can drop the i subscript and imposing stationarity we can drop the t subscript:

$$a_{it}(s_{it}, S_t, \boldsymbol{\varepsilon}_{it}) = a(s_{it}, S_t, \boldsymbol{\varepsilon}_{it})$$

Let $p(a_i|s_i, S; \sigma)$ denote the probability of player i choosing action a_i when he observes state (s_i, S) and has beliefs σ.

Value function We can write the ex-ante value function defined as the discounted sum of future payoffs before player specific shocks are observed and actions taken, as

$$V(s_i, S; \sigma) = \sum_{\mathbf{s}\in\mathbf{S}} \sum_{\mathbf{a}\in\mathbf{A}} p(\mathbf{a}|\mathbf{s}, S; \sigma)\sigma(\mathbf{s}|s_i, S) \left[\pi_i(\mathbf{a}, \mathbf{s}) + \beta \sum_{s_i', S'} \widetilde{g}(\mathbf{a}, \mathbf{s}, s_i', S')V(s_i', S'; \sigma)\right] \tag{B.1}$$

$$+\sum_{k=1}^{K} E_\varepsilon[\varepsilon_i^k|a_i = k].p(a_i = k|s_i, S; \sigma)$$

where $\widetilde{g}(\mathbf{a}, \mathbf{s}, s_i', S') = \left(\sum_{\mathbf{s}'\in\mathbf{S}} g(\mathbf{a}, \mathbf{s}, \mathbf{s}') \cdot \mathbf{1}(S' = S(\mathbf{s}'), s_i' \in \mathbf{s}')\right)$, $p(\mathbf{a}|\mathbf{s}, S; \sigma) = \prod_{j=1}^{N} p(a_j|s_j, S; \sigma)\cdot$ $\mathbf{1}(S = S(\mathbf{s}), s_j \in \mathbf{s})$ and E_ε is the expectation over the payoff shocks. Equation (B.1) is satisfied at each $(s_i, S \in \mathbf{S}_i \times J)$. This can be written in matrix form.

$$\begin{aligned} \mathbf{V}(\boldsymbol{\sigma}) &= (\boldsymbol{\sigma} \cdot \mathbf{P}(\sigma))\mathbf{\Pi} + \mathbf{D}(p(\sigma)) + \beta(\boldsymbol{\sigma} \cdot \mathbf{P}(\sigma))\widetilde{G}\mathbf{V}(\boldsymbol{\sigma}) \\ &= [\mathbf{I}_{LJ} - \beta(\boldsymbol{\sigma} \cdot \mathbf{P}(\sigma))\widetilde{G}]^{-1}[(\boldsymbol{\sigma} \cdot \mathbf{P}(\sigma))\mathbf{\Pi} + \mathbf{D}(p(\sigma))] \end{aligned} \tag{B.2}$$

where $\mathbf{V}(\boldsymbol{\sigma}) = [V(s_i, S; \sigma)]_{s_i, S\in\mathbf{S}_i\times J}$ is the $(L \cdot m_J) \times 1$ dimensional vector of expected discounted sum of future payoffs. $\widetilde{G}$ is the $(m_a \cdot m_s) \times (L \cdot m_j)$ transition matrix with element $\widetilde{g}(\mathbf{a}, \mathbf{s}, s_i', S')$. Π is the $m_a \cdot m_s \times 1$ dimensional vector of profits, $\mathbf{P}(\sigma)$ is the $(L \cdot m_J) \times (m_a \cdot m_s)$ dimensional matrix with element in row (s_i, S) and column $(\mathbf{a}, \mathbf{s})$ equal to $p(\mathbf{a}|\mathbf{s}; \sigma) \cdot \mathbf{1}(s_i \in \mathbf{s}, S = S(\mathbf{s}))$. $\boldsymbol{\sigma}$ is the $(L \cdot m_J) \times (m_a \cdot m_s)$ dimensional matrix with element in row (s_i, S) and column $(\mathbf{a}, \mathbf{s})$ equal to $\sigma(\mathbf{s}|s_i, S) \cdot \mathbf{1}(S = S(\mathbf{s}), (s_i) \in \mathbf{s})$ for any $\mathbf{a}$. $(\boldsymbol{\sigma} \cdot \mathbf{P}(\sigma))$ is the element by element product of the two matrices. I_J is the $L \cdot m_J$

dimensional identity matrix. Finally, $D(p(\sigma))$ is the $(L \cdot m_J) \times 1$ dimensional vector of expected payoff shocks.

E.2 Equilibrium

Let $u^S(a_i;\sigma)$ denote the continuation value net of payoff shocks under action a_i and beliefs σ.

$$u^S(a_i;\sigma,\theta) = \pi(a_i, s_i, S)$$

$$+\beta \sum_{\mathbf{s}_{-i}\in\mathbf{S}_{-i}} \sum_{\mathbf{a}_{-i}\in\mathbf{A}_{-i}} p(\mathbf{a}_{-i}|\mathbf{s}, S;\sigma)\sigma(\mathbf{s}|s_i, S) \sum_{\mathbf{s}'\in\mathbf{S}} g(\mathbf{a},\mathbf{s},\mathbf{s}') \cdot \mathbf{1}(S' = S(\mathbf{s}'))V(s_i', S';\sigma) \quad \text{(B.3)}$$

where $p(\mathbf{a}_{-i}|s_i, \mathbf{s}_{-i}, S;\sigma) = \prod_{j=1,j\neq i}^{N} p(a_j|s_j, S;\sigma) \cdot \mathbf{1}(S = S(s_i, \mathbf{s}_{-i}), s_j \in \mathbf{s}_{-i})$. The optimality condition requires that action a_i is chosen when

$$u(a_i;\sigma,\theta) + \varepsilon_i^{a_i} \geq u(a_i';\sigma,\theta) + \varepsilon_i^{a_i'} \quad \text{for all} \quad a_i' \in \mathbf{A}_i \quad \text{(B.4)}$$

Writing equation (B.4) as the ex-ante optimal choice probability

$$\begin{aligned} p(a_i|s_i, S;\sigma) &= \Psi(a_i, s_i, S;\sigma) \\ &= \int \left[\prod_{k\in A_i, k\neq a_i} \mathbf{1}(u(a_i;\sigma) - u(k;\sigma) \geq \varepsilon_i^k - \varepsilon_i^{a_i}) \right] dF \end{aligned} \quad \text{(B.5)}$$

We can write this system of $(K \cdot \mathbf{S}_i \cdot J)$ equations in vector form

$$\mathbf{p}^\sigma = \mathbf{\Psi}(\mathbf{p}^\sigma)$$

A solution exists by Brouwer's fixed point theorem since this equation is a continuous self-map on $[0,1]^{K\cdot\mathbf{S}_i\cdot J}$ as shown in Schmidt-Dengler and Pesendorfer (2008). The solution is the optimal behavior for given beliefs, $\mathbf{p}^\sigma$.

Stationary distribution Let $\lambda(\mathbf{s};\sigma)$ be the probability of being in state $\mathbf{s}$ and let the transition be

$$\lambda(\mathbf{s}';\sigma)=\sum_{\mathbf{s}}\sum_{\mathbf{a}}p(\mathbf{a}|\mathbf{s};\sigma)g(\mathbf{a},\mathbf{s},\mathbf{s}')\lambda(\mathbf{s};\sigma) \tag{B.6}$$

Define $\mathbf{P}^{\sigma}$ as the $m_s\times(m_s\cdot m_a)$ dimensional matrix consisting of $p(\mathbf{a}|\mathbf{s};\sigma)$ in row $\mathbf{s}$, columns $(\mathbf{a},\mathbf{s})$ and zero in the remaining columns. G is the $(m_a\cdot m_s)\times m_s$ dimensional matrix. Convergence to a unique stationary distribution with $\lambda^*(\mathbf{s};\sigma)>0$ for all $\mathbf{s}\in\mathbf{S}$ occurs if $\mathbf{P}^{\sigma}G$ is a regular matrix. We can write equation (B.6) in vector form as

$$\lambda^{\sigma\prime}=\mathbf{P}^{\sigma}G\lambda^{\sigma}$$

and the stationary distribution is the eigenvector $(\lambda^{\sigma}_{\mathbf{s}})$, solving

$$(\lambda^{\sigma}_{\mathbf{s}})^*.(I_{m_s}-\mathbf{P}^{\sigma}G)=0 \tag{B.7}$$

The stationary conditional distribution is

$$\begin{aligned}\lambda^*(\mathbf{s}|S;\sigma) &= \frac{\Pr(\mathbf{s},S|\sigma)}{\Pr(S|\sigma)}=\frac{\Pr(S|\mathbf{s};\sigma)\Pr(\mathbf{s}|\sigma)}{\Pr(S|\sigma)}\\ &= \frac{\mathbf{1}(S=S(\mathbf{s}))\cdot\lambda^*(\mathbf{s};\sigma)}{\sum_{\mathbf{s}\in\mathbf{S}}[\mathbf{1}(S=S(\mathbf{s}))\cdot\lambda^*(\mathbf{s};\sigma)]}\end{aligned} \tag{B.8}$$

$$\lambda^*(\mathbf{s}|s_i,S;\sigma)=\frac{\Pr(\mathbf{s},s_i,S|\sigma)}{\Pr(s_i,S|\sigma)}=\frac{1(S=S(\mathbf{s})))1(s_i\in\mathbf{s})\lambda^*(\mathbf{s};\sigma)}{\sum_{\mathbf{s}\in\mathbf{S}}[\mathbf{1}(S=S(\mathbf{s}),s_i\in\mathbf{s})\cdot\lambda^*(\mathbf{s};\sigma)]}$$

and $\lambda^*(\mathbf{s};\sigma)$ is probability of element $\mathbf{s}\in\mathbf{S}$ occurring under the stationary distribution $(\lambda^{\sigma}_{\mathbf{s}})^*$. We can write this in matrix form

$$\boldsymbol{\lambda}^*_{\mathbf{s}|s_i,S}=\boldsymbol{\Lambda}(\boldsymbol{\sigma})$$

where $\boldsymbol{\lambda}^*_{\mathbf{s}|s_i,S}$ denotes a $(m_s\cdot S_i\cdot J)\times 1$ dimensional vector. In equilibrium, beliefs must be consistent with optimal behavior $\boldsymbol{\lambda}^*_{\mathbf{s}|s_i,S}=\boldsymbol{\sigma}$

$$\boldsymbol{\lambda}^*_{\mathbf{s}|s_i,S}=\boldsymbol{\Lambda}(\boldsymbol{\lambda}^*_{\mathbf{s}|s_i,S}) \tag{B.9}$$

Equation (B.9) is both a necessary and sufficient condition for a Markov perfect equilibrium.

Theorem 1 *An equilibrium exists.*

Equation (B.9) gives a continuous self-map on $[0,1]^{m_s \cdot S_i \cdot J}$. Brouwer's fixed point theorem implies there exists (at least) a fixed point $\boldsymbol{\lambda}^*$ to the function $\boldsymbol{\Lambda}$. This fixed point corresponds to an equilibrium.