Identifying Firm-level R&D Cooperation and Innovation Decisions*

S. Amoroso[1]

[1]CentER, Tilburg University

November 27, 2012

Abstract

This paper investigate the distribution of the fixed and sunk costs associated with the firm–level decisions of innovating, spending, and cooperating in R&D, adopting a dynamic structural framework. The basic idea of the paper is to model the firms' decisions to cooperate in R&D and to innovate with a dynamic discrete choice model. None of the existing studies on heterogeneity of cooperation strategies or innovation processes deals with the non-trivial dynamics deriving from uncertainty and sunk costs of investments. Identifying the firms' primitives on productivity and investment decisions is key to have an encompassing understanding on what are the determinants on which the firm bases its choice to innovate and/or cooperate. Additionally, the suggested structural framework of firm heterogeneity in cost functions offers a straightforward extension to policy impact evaluation.

1 Introduction

An important source of productivity differentials across firms is related to R&D and innovation activities. Many authors have studied the connection between spending for R&D and productivity growth (Griliches, 1980; Jones and Williams, 1998; Hall and Mairesse, 1995; Crépon et al., 1998). As a result, a large number of empirical studies estimates the effect of R&D investment on such growth, finding that R&D spending has a significantly positive effect on productivity growth, with a rate of return that is about the same size as (or to some extent larger than) the rate of return on conventional investments. Crépon et al. (1998), examining the structural links between productivity, innovation input, and innovation output at the firm level, find that the firm innovation output rises with its research efforts, and the firm productivity correlates positively with a higher innovative output.

Nonetheless, before Ericson and Pakes (1995), most of the empirical literature in industry dynamics assumes that firms are endowed with an exogenous level of productivity. The "lucky" firms with high productivity survive and prosper, the others fail, and eventually exit the market.

The modern IO literature relaxes this exogeneity assumption by letting the productivity to be dependent on the investment decisions, so as to enhance the firms' survival chance (Ericson and Pakes, 1995; Doraszelski and Jaumandreu, 2008; Aw et al., 2011). Typically, in this context, the investment taken into consideration is past R&D expenditure (Doraszelski and Jaumandreu, 2008), or both R&D expenditure and export market

*Corresponding author: `s.amoroso@uvt.nl`.

participation (Aw et al., 2011). However, the firm that wants to survive must not only be innovative, but also ready to outsource knowledge and develop research networks. In fact, firms increasingly rely on the external acquisition of new technological knowledge, as the institutional locations of such resources can be quite disparate. Although not the primary source of produced knowledge, R&D outsourcing[1] (or external R&D) has considerably increased in importance and accounts for a substantial share of the total innovation expenditure in a large number of firms.[2]

Therefore, in this paper we construct a model where firms invest in R&D activities with or without a research partner to improve their productivity levels. In particular, we develop and estimate a structural dynamic monopoly model to quantify the linkages between R&D spending, innovation and cooperation investment choices, and endogenous productivity. To our knowledge, our paper constitutes the first attempt to explicitly model the different collaborative R&D investment decisions adopting a dynamic structural framework.

All the other empirical studies aimed at determining whether research collaborations affect firm-level productivity rely on reduced-form regression approach. For example, Belderbos et al. (2004b), using data from two waves of the Dutch Community Innovation Survey (1996, 1998), analyze the impact of R&D cooperation on firm performance, regressing two measures of firm-level productivity growth on four different cooperation strategies. They also control for the effect of both own R&D efforts as well as the impact of incoming knowledge flows that are not due to cooperation. Carboni (2012) explores the variables that determine a firms R&D collaborative expenditure, in a regression analysis framework, correcting for heteroscedasticity and non-normality when dealing with a large number of zero response data.

Differently from these studies, the model we propose derives the firms' optimal R&D investment decisions where these depend on the past R&D activities and on the past level of productivity. Additionally, within the suggested framework we are able to model and retrieve the current fixed or sunk costs relative to the different (collaborative) R&D activities.

The literature on R&D cooperation shows that the risks and costs of innovation and the need to exploit complementary resources are the main motives for cooperative behavior, and therefore, that cooperative behavior may be positively related to a number of obstacles such as high risks and cost of innovation (Amoroso, 2011; Belderbos et al., 2004a,b). R&D cooperations, in fact, allow firms to share costs or to reduce risks of innovation. In this regard, we hypothesize that cooperating in research could reduce both the fixed and the sunk costs of introducing an innovation in the market.

We merge data on sales and factor inputs of Dutch manufacturing firms extracted from the Production Survey (PS), and three waves of the Community Innovation Survey (CIS) for the Netherlands, covering the period from 2002 to 2008. The leading sectors (chemicals, agri-food, transport, high-tech) in the Dutch manufacturing industry heavily depend on research and innovation, and these are, in turn, driven by a wide range of factors, such as firm performance, market conditions, policy interventions, and government requirements to reduce environmental damages. In this paper, we assume the firm bases its decision of engaging in R&D or in innovation with or without a research partner on past choices, firm-level total factor productivity, and a demand shifter, proxying for the

[1] R&D outsourcing refers to the contractually agreed, non-gratuitous and temporary performance of R&D tasks for a client primarily by private contract research and technology organizations, but also by some private non-profit and related hybrid organizations (Howells, 1999; Grimpe and Kaiser, 2010)

[2] Source: Eurostat. "Innovation in Europe. Results for the EU, Iceland and Norway."

industry characteristics.

In Section 2 we present the model that we use to retrieve information on both fixed and sunk costs, and consequently on the optimal R&D, innovation, and cooperation decisions. In Section 3, we discuss the empirical strategy used to retrieve estimates of the static parameters of the model. Namely, we illustrate how we obtain a measure of firm-level productivity, demand elasticity, and an aggregate demand shifter. Moreover, we present estimates of the fixed costs associated with each investment choice, in the static case, i.e., when the firm does not take into account the future payoffs in its profits maximization. Section 4 describes the steps of the algorithm developed by Imai et al. (2009) that is used to obtain the dynamic parameters estimates. Section 4 and 5 describe the data and the results, respectively. In Section 6 we present the results for a policy simulation and the last section concludes.

2 Structural Framework

The empirical model builds on the class of models of dynamic entry games in IO, where the dependent variable is the firm's decision to enter or not in a market. In the same spirit, this paper defines the entry decision as the adoption of a set of discrete decisions: investing in research and development (R&D), cooperating, innovating, and innovation cooperation. These decisions are assumed to be costly to reverse and, therefore, associated with sunk costs. As firms are assumed to be forward looking, they take into account the implications of their decisions (and the associated costs) on their future payoffs. Time is discrete and indexed by t. The single-agent dynamic optimization problem is solved for the N firms operating in the market, which we index by $i \in I = \{1, 2, \ldots, N\}$. Following the standard setting of Ericson and Pakes (1995), and adapting it to a monopolistic competitive setting, firms compete on two different dimensions: a static and a dynamic dimension. Within the dynamic dimension, a firm makes the investment choice indexed by $k \in \{na, rd, c, d, cd\}$, where the vector of choices is defined as $a_{it} = (na_{it}, rd_{it}, c_{it}, d_{it}, cd_{it})'$, with $a_{it} \in A_i \equiv \{0,1\}^5$. The firm-specific choice na_{it} takes value one if the firm does not engage in any activity other than operating in the market; rd_{it} takes value one if the firm decides to spend in R&D; choices c_{it} and d_{it} match firms' decisions to start a research collaboration and to invest in a technological upgrade, respectively; action cd_{it} tags the decision to both innovate and cooperate (with either another firm or a research institute or a supplier/customer).

2.1 Static decisions

In every period, firms are competing in prices in a static Bertrand model. Let P_{it}, the price, be the static decision variable of firm i at time t. The demand curve faced by the monopolistically competitive firm is assumed to follow a Dixit–Stiglitz form:

$$Q_{it}^D = Q_t^j (P_{it}/P_t^j)^\eta e^{u_{it}^d} \tag{1}$$

where Q_{it}^D is the demanded quantity for a firm i, Q_t^j and P_t^j are the sector j aggregate production and price index, respectively, $\eta < -1$ is the constant elasticity of demand, and u_{it}^d is a demand shock.

The production function is assumed to take the form of a Cobb-Douglas, with the gross output Q_{it} of firm i at time t function of three specific inputs and productivity:

$$Q_{it} = A_{it} K_{it}^{\theta_{iKt}} L_{it}^{\theta_{iLt}} M_{it}^{\theta_{iMt}}, \tag{2}$$

where K_{it} denotes capital, L_{it} labor, and M_{it} intermediate goods, consisting of materials and energy, for firm i at period t. $\theta_{iKt}, \theta_{iLt}, \theta_{iMt}$ are the elasticities of output with respect to capital, labor, and intermediate goods, respectively. A_{it} represents the Hicksian neutral efficiency level of firm i at time t. The logarithm of A_{it} is defined as $A_{it} \equiv \exp(\theta_0 + \omega_{it})$ and is defined as the sum of the mean productivity level across firms and over time, θ_0, and the productivity shock which is observable by the firm, but not to the econometrician (for example, managerial ability, quality of research), ω_{it}..

Following the literature on imperfect competition in both product and labor markets (Bughin, 1993, 1996; Crépon et al., 2002; Dobbelaere, 2004; Abraham et al., 2009; Dobbelaere and Mairesse, 2011; Amoroso et al., 2012), we relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power. The workers bargain with the firm over both the levels of employment, L_{it}, and of the wage, W_{it}. Additionally, we define the firm level profits as

$$\Pi_{it} \equiv P_{it}Q_{it} - W_{it}L_{it} - FC(K_{it}, M_{it}, a_{it}), \tag{3}$$

where $FC(\cdot)$ are the (avoidable) fixed costs (costs that do not vary with the quantity of output produced, but are not irrevocably committed; (Wang and Yang, 2001)), depending on capital, material, and innovative investment. Moreover, we define the union's utility function as

$$U_{it}(W_{it}, L_{it}) \equiv L_{it}(W_{it} - \bar{W}_{it}),$$

where $\bar{W}_{it}$ is the reservation wage. Finally, the efficient bargaining model can be written as a weighted average of the logarithms of workers' aggregate gain from union membership and the firm's profits:

$$\max_{L_{it}, W_{it}} \left[\phi_{it} \log(U_{it}(W_{it}, L_{it})) + (1 - \phi_{it}) \log \Pi_{it}\right],$$

where $\phi_{it} \in [0, 1]$ is the degree of union bargaining power. In the static setting, the firm maximizes only with respect to the variable costs, namely, the cost of labor. Amoroso et al. (2012) show that, maximizing with respect to labor, and taking into account the demand curve faced by the monopolistically competitive firm, results in the following expression for the elasticity of the labor input factor:

$$\theta_{iLt} \equiv \left(\frac{\eta}{1+\eta}\right) \frac{W_{it}L_{it}}{P_{it}Q_{it}}(1 - \mu^W_{it}). \tag{4}$$

Amoroso et al. (2012) define the bargained wage rate $\mu^W_{it} \equiv \frac{W_{it} - \bar{W}_{it}}{W_{it}}$ as the *wage markup*[3] From (4), after solving for L_{it} (see technical appendix), we derive the following expression for labor:

$$L_{it} = \left[(\exp(\theta_0 + \omega_{it})K_{it}^{\theta_K} M_{it}^{\theta_M})^{\frac{\eta+1}{\eta}} \frac{1}{1-\mu^W} \frac{\eta+1}{\eta} \frac{\theta_{iLt}}{W_{it}} \frac{P^j_t}{(Q^j_t)^{1/\eta}} (\exp(u^d_{it}))^{-1/\eta}\right]^{\eta/(\eta - \theta_{iLt}(\eta-1))} \tag{5}$$

Substituting (5) into (3), taking into account (2), and assuming, for simplicity, that the elasticity of labor is constant across firms and time, we obtain the final short-run profit

[3]In their paper, Amoroso et al. (2012) also show how, maximizing with respect to wages leads to an expression of the wage markup as a function of the bargaining parameter, ϕ_{it}, and the ratio between profits and cost of labor, $\mu^W_{it} = \frac{\phi_{it}}{1-\phi_{it}} \frac{\Pi_{it}}{W_{it}L_{it}}$.

function:

$$\Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) = \left(\frac{1-\gamma}{\gamma^{1-\delta}}\right) W_{it}^{1-\delta} \left[\left(\exp(\theta_0 + \omega_{it}) K_{it}^{\theta_K} M_{it}^{\theta_M} \right)^{\frac{\eta+1}{\eta}} \left(\psi_t (\exp(u_{it}^d))^{-1/\eta} \right) \right]^{\delta} \tag{6}$$

where $\psi_t \equiv \frac{P_t^j}{(Q_t^j)^{1/\eta}}$, $\gamma \equiv \theta_L \frac{\eta+1}{\eta} \frac{1}{1-\mu^W}$, and $\delta \equiv \eta/(\eta - \theta_{iLt}(\eta - 1))$.

2.2 Dynamic decisions

The decisions of doing R&D, cooperating, or innovating cannot be revoked, so we assume the costs associated with these actions to be sunk. We define the vector of fixed costs paid in case of investment in research, cooperation, innovation, or both cooperation and innovating as $\theta_i^{FC} = (0, \theta_i^{FC}(rd), \theta_i^{FC}(c), \theta_i^{FC}(d), \theta_i^{FC}(cd))'$. We also define the vector of sunk costs associated with every investment choice k, $\theta_i^{SC} = (0, \theta_i^{SC}(rd), \theta_i^{SC}(c), \theta_i^{SC}(d), \theta_i^{SC}(cd))'$. In particular, we assume that, besides the fixed and sunk costs of R&D and innovation, there are sunk costs of finding an efficient research partner, or fixed costs of maintaining the research alliance, such as managing the contractual costs (transaction costs).

Given their level of productivity, capital, materials, and present and past knowledge investment decisions, a_{it} and a_{it-1}, the firm faces the following profit function:

$$\begin{aligned}
&\Pi(a_{it}, a_{it-1}, \omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) = \\
&\quad \Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) - FC(K_{it}, M_{it}, a_{it}) - SC(a_{it}, a_{it-1}) \\
&\quad \equiv \Pi^{SR}(\omega_{it}, W_{it}, K_{it}, M_{it}, \psi_t) - \tilde{FC}(K_{it}, M_{it}) - \theta_i'^{FC} a_{it} - \theta_i'^{SC}(1 - a_{it-1}) a_{it},
\end{aligned} \tag{7}$$

where the function of the fixed costs of operation is defined as $FC(K_{it}, M_{it}, a_{it}) \equiv \tilde{FC}(K_{it}, M_{it}) - \theta_i'^{FC} a_{it-1}$

To simplify the framework, while retaining the salient features of the model, we make a set of assumptions. First, we omit the firm-level entry/exit decisions. Moreover, to reduce the dimensionality of the state vector on which firms are assumed to base their decisions, we consider a simpler framework, featuring imperfect competition only on the output market, and where capital and materials are assumed to be flexible inputs, not subject to adjustment costs. Assuming that the productivity, ω_{it}, and the aggregate state, ψ_t, are sufficient statistics for predicting the expected future profits, the short-run profit function under these restrictions is derived in the Appendix and can be written as

$$\Pi(a_{it}, a_{it-1}, \omega_{it}, \psi_t) = \varphi \psi_t \exp(\omega_{it})^{-(1+\eta)} - \theta_i'^{FC} a_{it} - \theta_i'^{SC}(1 - a_{it-1}) a_{it}, \tag{8}$$

where $\varphi \equiv -\frac{1}{1+\eta}\left(\frac{\eta}{1+\eta}\right)^{\eta}$.

2.2.1 State variables transition functions

We assume that the next period state of the aggregate variable ψ_t depends only on the current state. In particular, we specify the evolution of the aggregate state variable as

$$\psi_t = f(\psi_{t-1}) = \mu_0 + \rho \psi_{t-1} + \epsilon_{\psi}, \tag{9}$$

where ϵ_{ψ} is a normally distributed error term. Following Santos (2009), the variance of ϵ_{ψ}, $\sigma_{\epsilon}^2 = \sigma_{\psi}^2(1 - \rho^2)$, represents the aggregate uncertainty of the industry affecting the firm's investment choice.

Concerning the productivity, we follow Doraszelski and Jaumandreu (2008), and Aw et al. (2011), and model the evolution of the firm's productivity as a Markov process, allowing for the productivity to be affected by firms' past choices of innovation, cooperation, and R&D.[4] We define the evolution process of productivity level ω_{it} of firm i at time t as:

$$\omega_{it} \equiv \omega(\omega_{it-1}, a_{it-1}) + \xi_{it} \tag{10}$$

where ξ_{it} is the normally distributed stochastic shock to productivity, and $\omega(\cdot)$ is approximated by a third degree polynomial.

In particular, we propose the evolution process of productivity level ω_{it} of firm i at time t as a nonlinearly persistent process, depending on a broader set of R&D activities, namely (cooperative) research and innovation. The productivity transition becomes:

$$\begin{aligned} \omega_{it} &= \omega(\omega_{it-1}, c_{it-1}, d_{it-1}, cd_{it-1}, rd_{it-1}) + \xi_{it} \\ &= \beta_0 + \beta_1\omega_{it-1} + \beta_2\omega_{it-1}^2 + \beta_3\omega_{it-1}^3 + \beta_4 c_{it-1} + \beta_5 d_{it-1} \\ &\quad +\beta_6 c_{it-1}d_{it-1} + \beta_7 rd_{it-1} + \xi_{it}. \end{aligned} \tag{11}$$

The firm profit function as in (7) related to the set of choices a do not only differ in their fixed costs intercepts, but also in their arguments. In fact, the productivity process assumed in (11) depends on both the past level of productivity, and on the type of technological upgrade. Therefore, the variable ω_{it} associated with one choice might be different from that of an alternative investment choice.

Figure 3 reports the schematic representation of what the profile of all the optimal strategies for firm i and the relative payoffs, given the levels of productivity, could look like. The firms with a productivity level above a certain threshold decide to either invest in R&D ($\omega_{it} > \omega^{rd}$), or to cooperate with a research partner ($\omega_{it} > \omega^{c}$), as it might provide higher profits than doing R&D by themselves. Cooperating yields higher profits since firms reduce the costs and associated risks of research by sharing them. Enterprises observing a level of productivity high enough to bear the sunk costs of introducing an innovation, invest in a product or process improvement that offers a greater performance or a reduced cost of production ($\omega_{it} > \omega^{d}$). Firms with productivity $\omega_{it} > \omega^{cd}$ engage in both activities and are thus assumed to be the most productive.

2.2.2 Value and policy functions

To retrieve information about the sunk costs of R&D, innovating, and cooperating, and to identify the evolution of the productivity states depending on firms' research investment policies, we consider a dynamic programming problem in which a firm i makes a series of discrete choices over its infinite lifetime.

Let a_{it} be the control variable and let S be the set of space state points and let the firms' characteristics s_{it} be an element of S. To simplify the framework, without losing the generality of the model, we assume that the state of firm i at time t is defined only by the level of productivity, ω_{it}, the industry competition proxied by the aggregate state ψ_t, and the past investment actions, a_{it-1}; therefore the state vector is summarized as

[4] Doraszelski and Jaumandreu (2008) relax the exogeneity assumption usually made about productivity in the production function literature (see Olley and Pakes (1996), Levinsohn and Petrin (2003), Ackerberg et al. (2006)), by letting the R&D spending and related activities to determine the differences in and the evolution of productivity across firms and over time. Aw et al. (2011) take a step further and assume that productivity evolves as a Markov process which depends on both investments in R&D and export market participation.

$s_{it} = (\omega_{it}, \psi_t, a_{it-1})$. To fit the model to the data, we need to add unobserved heterogeneity. In particular, we introduce the vector of payoff shocks $\epsilon_{it} = \{\epsilon_{it}(k)\}_{k\in\{na,rd,c,d,cd\}}$ observed only by the firm. The unobserved characteristics ϵ_{it} are independently and identically distributed over time with continuous support and multivariate distribution function $F_\epsilon(\epsilon_{it})$. In particular, I assume that ϵ_{it}'s are *i.i.d.* extreme value distributed and enter the profit function in an additively separable way. These assumptions are not strictly necessary, but useful as they lead to a closed form likelihood function and a closed form expression for the expected maximum of the choice-specific value functions.

The observed state variable ω_{it} evolves as a Markov process depending stochastically on the choices of the firm because of the assumption in equation (10) with the cumulative distribution function given by $F_\omega(\omega_{it+1}|\omega_{it}, a_{it})$. On the other hand, the stochastic evolution of the aggregate state is assumed to be independent from the research activities, and therefore can be expressed as $F_\psi(\psi_{t+1}|\psi_t)$. Moreover, since we do not know the firm–level production technology, we assume the sunk costs of R&D, innovating, and of partnering in research to be drawn from a known joint distribution $F_{SC}(\theta_i^{SC})$.

Let us define $\theta_{\Pi i} \equiv ((\theta_i^{FC})', (\theta^{SC})')'$, and $\theta_\Pi \equiv \{\theta_{\Pi i}\}_{i=1,\dots,N}$ as the matrix of choice- and firm-specific parameters that describe the profit function in (7). Finally, let $\theta = (vec(\theta_\Pi)', \theta_\omega', \theta_\psi', \theta_\epsilon', \beta)' \in \Theta$ be the vector of the parameters of interest, where $vec(\theta_\Pi)$ is the vectorization of the θ_Π matrix, and where θ_ω and θ_ψ are vectors of parameters that describe the transition probability functions F_ω and F_ψ, respectively, θ_ϵ represents the parameters in the distribution of F_ϵ, and β is the rate at which the firm discounts future profits.

Assuming that firms behave optimally, the value function of firm i corresponds to the maximum of the expected discounted sum of profits, conditional on the current level of productivity and market indexes:

$$V(s_{it}, \epsilon_{it}; \theta) \equiv \max_{a_{it}, a_{it+1}, \dots} E\left[\sum_{\tau=t}^{\infty} \beta^{\tau-t} \left(\Pi(a_{i\tau}, s_{i\tau}; \theta_{\Pi i}) + \epsilon_{i\tau}\right) | s_{it}, \epsilon_{it}\right] \tag{12}$$

where $\beta \in (0,1)$, and $\Pi(a_{it}, s_{it}; \theta_{\Pi i}) + \epsilon_{it}$ are the current profits of firm i with productivity level ω_{it}, in market aggregate condition ψ_t, choosing investment a_{it}.

The problem is to determine, for all N firms, the set of optimal stationary decision rules $\alpha = \{\alpha_i\}_{i=1}^N$, where $\alpha_i : S \to A_i$, that solves the stochastic/multiperiod optimization problem expressed in (12). The method of dynamic programming offers the advantage of translating the optimization problem in (12) into a sequence of simpler deterministic/static optimization problems, where for $\beta \in (0,1)$ and for bounded $\Pi(\cdot)$, the value of the objective function can be written (suppressing the subscript i) in the form of a Bellman equation:

$$\begin{aligned} V(a, s, \epsilon; \theta) &= \Pi(a, s; \theta_\Pi) + \epsilon + \beta E_{s',\epsilon'}\left[V(s'; \theta)|s, a\right] \\ V(s, \epsilon; \theta) &= \max_{a\in A} V(a, s, \epsilon; \theta) \end{aligned} \tag{13}$$

where s' and ϵ' denote the next period state and shock. Therefore, when conditioning on the value of the state and control variables, the optimal decisions of the firm do not depend on time t, but only on current and next period state variables. The assumption of the existence of a state variable that is designed to capture the productive and competitive environment faced by the firm at each point might be quite restrictive in the context of technological innovation. However, as in this paper, we consider the dynamic optimization problem of a single agent, the stationary dynamic programming framework could still capture the salient features of such a structural model.

The expected value function for next period is equal to:

$$E_{s',\epsilon'}\left[V(s',\epsilon';\theta)|s,a\right] = \int_{s'}\int_{\epsilon'} V(s',\epsilon';\theta)dF_\epsilon(\epsilon';\theta_\epsilon)dF_s(s'|s,a;\theta), \tag{14}$$

where $dF_s(s'|s,a;\theta) \equiv dF_\omega(\omega'|\omega,a;\theta_\omega)dF_\psi(\psi'|\psi;\theta_\psi)$. Given that the optimal strategy, $\alpha(s,\epsilon)$, satisfies

$$\alpha(s,\epsilon) = \arg\max_{a\in A} V(a,s,\epsilon;\theta),$$

and observing data $(\mathbf{a},\boldsymbol{\omega},\boldsymbol{\psi}) \equiv \{\{a_{it},\omega_{it}\}_{i=1}^N,\psi_t\}_{t=1}^T$, in order to estimate θ, we construct the likelihood as the product of firms' conditional choice probabilities (CCPs), $P_{it}(a_{it}|s_{it};\theta)$, as

$$\begin{aligned} P_{it}(a_{it}|s_{it};\theta) &\equiv Pr(\epsilon : V(a_{it},s_{it};\theta) \geq V(\tilde{a}_{it},s_{it};\theta)), \quad \forall \tilde{a}_{it} \\ &= Pr(\epsilon : a_{it} = \alpha(s_{it},\epsilon_{it})) \\ &= \int \mathbb{1}\{a_{it} = \alpha(s_{it},\epsilon_{it})\}dF_\epsilon. \end{aligned}$$

The joint likelihood of the observed data is then:

$$L(\mathbf{a}|\mathbf{s};\theta) = \prod_i\prod_t P_{it}(a_{it}|s_{it};\theta). \tag{15}$$

Moreover, since ϵ follows a joint Gumbel (extreme value type I) distribution, independent across alternatives k, the likelihood increment for firm i is

$$P_{it}(a_{it}|s_{it};\theta) = \frac{\exp\{V(\tilde{a}_{it},s_{it};\theta)\}}{\sum_{a_{it}\neq\tilde{a}_{it}}\exp\{V(a_{it},s_{it};\theta)\}}. \tag{16}$$

In the next section, we discuss the empirical strategy to estimate the static structural parameters, namely, the demand elasticity, the wage markup, the aggregate state proxying the industry competitive environment, the productivity evolution parameters, the fixed costs, and the dynamic parameters, i.e., the sunk costs, and the discount factor.

3 The estimation procedure

Estimation is done in three steps. In the first step, we estimate a production function that allows us to retrieve estimates of the firm-level productivity, ω_{it}, the parameters describing the aggregate state and productivity evolution processes, $f(\psi_{t-1})$, and $\omega(\omega_{it-1},a_{it-1})$, repectively, and the structural parameters needed to construct the profit function as in (6). In the second step, we retrieve the management costs concerning the research activity adopted by the firm. In the last step, we obtain estimates of the dynamic structural parameters, $\theta_\Pi,\theta_\omega,\theta_\psi,\theta_\epsilon$, by numerical approximation of the solution to the dynamic programming problem at trial parameters.

3.1 Step 1: Static parameters

The production function and the demand parameters are estimated with the method proposed by Amoroso et al. (2012). Within the Cobb-Douglas production function framework, they relax the conventional assumption of perfect competition in the labor market, allowing both firms and workers' union to have some market power.

In their study, Amoroso et al. (2012) report empirical evidence of the underestimation of the true level of price–cost margins caused by the omission of direct effects of the wage bill on marginal costs. In fact, the exclusion of frictions in the labor market (i.e., $\phi_{it} = 0$ or $W_{it} = \bar{W}_{it}$) might lead to misestimating the firm's market power. When there is no imperfect competition in the labor market, firms set the wage at the lowest value possible, ultimately equal to the competitive wage, i.e., $W_{it} = \bar{W}_{it}$ (and, therefore, $\mu_{it}^{W} = 0$). For W_{it} that tends to $\bar{W}_{it}$, the wage markup decreases, given that the elasticity and the share of labor are constant, which is inversely related to the output markup $\frac{\eta}{1+\eta}$.

Next to the labor market rigidities, Amoroso et al. (2012) also correct for the possible bias in the estimated coefficients when deflated gross output is used instead of gross physical output. Defining the log deflated output as y_{it}, this can be rewritten as

$$y_{it} = q_{it} + (p_{it} - p_t^j),$$

where p_t^j is the log industry price index. The firm-level price deviations $(p_{it} - p_t^j)$ will enter the production function as an extra error component, introducing potential correlation with the input choices. Substituting p_{it} with the inverse Dixit-Stiglitz demand function, and taking into account the labor input elasticity under imperfect competition in the labor market,

$$\theta_{iLt}\left(\frac{\eta+1}{\eta}\right) \equiv \gamma_{iLt} = s_{iLt}(1 - \mu_{it}^{W}), \tag{17}$$

where s_{iLt} is the share of labor and it is defined as the ratio between cost of labor and total sales $\left(\frac{W_{it}L_{it}}{P_{it}(Q_{it})Q_{it}}\right)$, they estimate a log deflated revenue function that features both labor and output market distortions:

$$y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1 - \mu_{it}^{W}) s_{iLt} l_{it} - \frac{1}{\eta} q_t^j + \tilde{\omega}_{it} + \tilde{u}_{it} \tag{18}$$

where k_{it}, l_{it}, m_{it} are logs of deflated capital, labor, and deflated materials, respectively; q_t^j is the log of the production index in sector j. The composite error term, $\tilde{u}_{it} \equiv u_{it}^q + u_{it}^d$, includes the demand shock, $\tilde{u}_{it}^d \equiv -u_{it}^d/\eta$, and the measurement error, u_{it}^q. $\tilde{\omega}_{it} \equiv \omega_{it}(1+\eta)/\eta$ is the productivity.

The production index is constructed as in De Loecker (2011), by proxying the total demand for a sector j with a (market share) weighted average of deflated revenue, $q_t^j = \sum_i^{N_j} ms_{it} y_{it}$. Both the intercept, $\gamma_0 \equiv \theta_0(1+\eta)/\eta$, and the factor elasticities of capital and material, $\gamma_k \equiv \theta_k(1+\eta)/\eta, k = K, M$ are divided by the *output price markup* defined as $\equiv \eta/(1+\eta)$ for $\eta < -1$. The elasticity of labor is defined as in (17).

The firm–level productivity ω_{it} is estimated as

$$\hat{\omega}_{it} = \hat{\eta}/(1+\hat{\eta})\tilde{\omega}_{it} = \hat{\eta}/(1+\hat{\eta})\left[y_{it} - \left(\hat{\gamma}_0 + \hat{\gamma}_K k_{it} + \hat{\gamma}_M m_{it} + (1 - \hat{\mu}_{it}^{W}) s_{iLt} l_{it} - \frac{1}{\hat{\eta}} q_t^j\right)\right].$$

Identification of all the structural parameter of the deflated revenue function in (18) is ensured by the presence of firm specific wages. To estimate all the relevant parameters, they adopt a control function approach (Olley and Pakes, 1996) which consists in including additional regressors to capture the endogenous part of the unobserved productivity. In particular, the productivity $\tilde{\omega}_{it}$ can be approximated by a third–degree polynomial (Levinsohn and Petrin, 2003) in all three factor inputs k_{it}, l_{it}, m_{it}. The productivity is also assumed to evolve over time as a Markov process that depends on the firms' investment choices, as in (11). The replacement function approach allows for dynamics in the productivity process, but restricts the investment function, and consequently the productivity process, to be homogeneous across firms. On the other hand, the instrumental variables approach comes at the cost of not allowing for the possibility that the unobserved productivity could be correlated with past choices of inputs. Therefore, for the problem at hand, we rely on the control function approach to identify the deflated revenue function parameter, and our object of interest, the firm level productivity. The estimation of (18) requires the following moment restrictions

$$E(\xi_{it} + \tilde{u}_{it} | m_{it}, k_{it}, l_{it-1}, m_{it-1}, k_{it-1}, \ldots, l_{i1}, m_{i1}, k_{i1}) = 0,$$

however, identification could hold with just current values and one lag in the conditioning set.

Results of the estimation of the deflated revenue function under imperfect competition in both output and labor markets (18), of the aggregate state transition function (9), and of the nonlinearly persistent productivity process depending on technology upgrading (11) are reported and discussed in Section 4. In the following subsection, we discuss the second step of our estimation strategy, namely, how to retrieve the fixed costs of (cooperative) research and innovation.

3.2 Step 2: Profit function parameters

It is well-known that, in general, the parameters of structural dynamic programming problems are not identified (Rust, 1994). Magnac and Thesmar (2002) show that the utility functions of the firms can be identified if the distribution function of the unobserved preference shocks, the discount rate, and the value function of one the alternatives (normalization) are fixed. Hence, it is theoretically possible to identify both fixed and sunk costs of R&D and innovation. However, in practice the simultaneous identification of such costs requires enough variation in the observed R&D investment decisions. To circumvent this problem, we recover the fixed cost parameters within the static framework, after having estimated the production function parameters. In particular, we consider the estimation of the fixed costs of innovative investments as a random utility model (multinomial mixed logit model), where the alternative-specific utility function of firm i is associated with the level of productivity and fixed costs represent the alternative-specific firm-level random coefficients associated with the research ivestment k, i.e.,

$$V(a_{it}, s_{it}, \zeta_{it}; \theta^{FC}) = \varphi \psi_t^{\eta} \exp(\omega_{it})^{-(1+\eta)} - \theta_i'^{FC} a_{it} + \zeta_{it}.$$

The error term ζ_{it} is a random term assumed to be iid extreme value distributed. To identify θ_i^{FC}, we assume that the additive separable utility shock ζ_{it} is exogenous. Results of this estimation are reported in Section 5.

3.3 Step 3: Dynamic parameters

The main limiting factor in estimating dynamic discrete choice (DDC) models is the computational complexity resulting form the need to compute the continuation values as in (14). The direct way of obtaining such continuation values has been to compute them as the fixed point of a functional equation. For example, Rust (1987) proposes a computational strategy named the nested fixed point (NFXP) algorithm, which is a gradient iterative search method to obtain the maximum likelihood estimator of the structural parameters. Unfortunately, the NFXP algorithm is computationally demanding because it requires to obtain the fixed point of a Bellman operator (hence, it must run successive iterations of the value functions until convergence) for each point in the state space of the structural parameters. Additionally, the number of state points grows exponentially with the dimensionality of the state space. This concern about the computational burden of implementing the NFXP algorithm, and the curse of dimensionality, have led to a number of estimators that are computationally faster (Bajari et al., 2007; Pakes et al., 2007). For example, the two-step estimator by Hotz and Miller (1993), using nonparametric estimates of choice and state transition probabilities, yields a simple representation of the choice-specific value functions for values in a neighborhood of the true vector of structural parameters.[5] The main advantage of this two-step estimator is its computational simplicity. The first step is a nonparametric regression to obtain the productivity and the aggregate state transition functions, the second step is the estimation of a standard discrete choice model (the policy functions) with a criterion function that is globally concave (e.g., such as the likelihood of a multinomial logit model in our investment choice study case). Thus, the agent's continuation values can be obtained nonparametrically by first estimating the agent's choice probabilities at each state, and then inverting the choice problem to obtain the corresponding continuation values. However, as with other approaches, there are limitations. First, since the two-step empirical strategy involves the (nonparametric) estimation of the CCPs, the continuation values are estimated rather than computed, and therefore they contain sampling error. This sampling error might be significant if the state space of the model is large relative to the available data. The second limitation comes from the formal requirements of the limit properties of the estimator. As a matter of fact, to obtain an estimator with desirable properties, the data must visit a subset of the points repeatedly. More precisely, all the states in some recurrent class $\Re \subseteq S$ must be visited infinitely often, and the equilibrium strategies must be the same every time each point of $\Re$ is visited. Simply put, the two-step approach requires the assumption of stationarity. To give an example, when forecasting the CCPs of a firm observed in year t when being active on the market in year $t + \tau$, it is assumed that the firm at time t would face the same decision-making environment observed in year $t + \tau$. Moreover, it must also be assumed that there is no permanent unobserved heterogeneity, otherwise, it would be impossible to match the actions of the firm at time t with the action at time $t + \tau$.

To correct for the finite sample bias, Aguirregabiria and Mira (2002) propose a nested pseudo-likelihood algorithm (NPL) for the estimation of the class of discrete Markov decision models with the conditional independence assumption. In particular, their method considers a K-step extension of the Hotz and Miller (1993) estimator. In fact, Aguirregabiria and Mira (2002) obtain a new estimate of the CCPs given the two-step estimator and an initial nonparametric estimator of the CCPs. Successive iterations return a sequence of estimators of the structural parameters and CCPs that are asymptotically equivalent to the partial MLE and to the two-step PML (Aguirregabiria and Mira, 2002, Proposition 4). Moreover, Aguirregabiria and Mira (2002) report results from Monte Carlo experiments that illustrate how iterating in this procedure does in fact produce significant reductions in finite sample bias. However, their estimation algorithm have difficulties dealing with unobserved heterogeneity. Extensions to accommodate unobserved heterogeneity via finite mixture distributions into CCP estimation are attributable to Arcidiacono and Miller (2011).

Given these recent extensions, there is still one main limiting factor in estimating DP models, which is the computational burden associated with the iterative process.

[5] For an exhaustive, but self-contained review and description of Hotz and Miller (1993) two-step estimator and extensions, see Aguirregabiria and Mira (2010).

Therefore, it is not surprising that there have been continuing efforts to reduce the computational burden of estimating DP models. Recently, computationally practical Bayesian approaches that rely on Markov Chain Monte Carlo (MCMC) methods have been developed by Imai et al. (2009) and Norets (2009).

In this paper, we adopt the estimation method proposed by Imai et al. (2009). Their algorithm is related to the one proposed by Aguirregabiria and Mira (2002), but it is based on the full solution of the DP problem, yielding the advantage of dealing with unobserved heterogeneity. The main idea of their estimation approach is to avoid the computation of the full solution of the DP problem, by approximating the expected value function at a state space point using the average of value functions at past iterations in which the parameter vector is close to the current parameter vector and the state variables are close to the current state variables.[6] In the conventional NFXP algorithm, most of the information obtained in the past iterations remains unused in the current iteration.

The Imai et al. (2009) algorithm consists of two loops:

1. **The outer loop (Metropolis-Hasting Algorithm)**
   The outer loop performs a M-H (Metropolis-Hasting) algorithm. First, we draw a candidate parameter vector from a proposal density, then we evaluate the likelihood, conditional on the candidate parameter vector and on the previous iteration parameter vector, to compute the acceptance probability, with which we can decide whether or not to accept the candidate parameter vector.

   In our setting, we allow for the parameters of the profit function, θ_Π, to take different values for each firm. In particular, we assume that the vector of firm-specific parameters $\theta_{\Pi i}$ follows the density function:

   $$\theta_{\Pi i} \sim g(\theta_{\Pi i}(a); \mu),$$

   where $\mu = (\bar{\theta}_\Pi, \sigma_\Pi)'$ is the hyperparameter vector for this density. In particular, we assume g is a normal distribution and μ includes parameters for means, $\bar{\theta}_\Pi$, and standard deviations, σ_Π. Assuming that the prior of the mean parameters is normal and that of the standard deviation parameters is inverted Gamma, the posterior distribution for the mean parameters is normal and that for the standard deviation parameter is inverted Gamma. To simplify the framework, without losing the generality of the structural model, we assume that the priors are independent across investment alternatives.

   The entire parameter vector consists now of $\theta = (\mu', vec(\theta_\Pi)', \theta'_\omega, \theta'_\psi, \theta'_\epsilon, \beta)'$. Following Ching et al. (2012), let us rewrite this vector as $\theta = (\mu', vec(\theta_\Pi)', (\theta_c)')'$, where $\theta_c = (\theta'_\omega, \theta'_\psi, \theta'_\epsilon, \beta)'$ is the vector of parameters common across firms. As for the prior on θ_c, we use independent flat priors. Suppose we are at iteration r, with parameter estimates being $(\mu^r, vec(\theta_\Pi), \theta_c)$, then the outer loop iteration for drawing a parameter vector from the posterior distribution can be divided into three steps:

   1.1 *Hyperparameter updating step*

[6] Ching et al. (2012) claim that the practical Bayesian approach developed by Imai et al. (2009)

> "...is potentially superior to prior methods because (1) it could significantly reduce the computational burden of solving for the DDP model in each iteration, and (2) it produces the posterior distribution of parameter vectors, and the corresponding solutions for the DDP model–this avoids the need to search for the global maximum of a complicated likelihood function."

Draw μ^r. That is, given θ_{Π}^{r-1}, for all alternative $a \in A$, draw $\bar{\theta}_{\Pi} \sim f_\theta(\cdot|\sigma_{\theta_\Pi}^{r-1}, \{\theta_{\Pi i}^{r-1}\}_{i=1}^N)$ and $\sigma_{\Pi(a)}^r \sim f_\sigma(\cdot|\bar{\theta}_{\Pi}^r, \{\theta_{\Pi i}^{r-1}\}_{i=1}^N)$, where f_θ and f_σ are the conditional posterior distributions.

1.2 *Data augmentation step*

Now that we have effectively constructed the prior for $\theta_{\Pi i}$, we draw, for each alternative a, a candidate parameter from the proposal density, which we assume to be a normal density,

$$\theta_{\Pi i}^{*r} \sim q(\bar{\theta}_{\Pi}^{r-1}, \sigma_{\theta_\Pi}^{r-1}).$$

Then, accept $\theta_{\Pi i}^{*r}$ with probability λ, where

$$\lambda = \min\left\{ \frac{g(\theta_{\Pi i}^{*r};\mu^r) P_i^r(a_i|\omega_i,\psi;\theta_{\Pi i}^{*r},\theta_c^{r-1}) q(\cdot|\theta_{\Pi i}^{*r},\mu^r)}{g(\theta_{\Pi i}^{r-1};\mu^r) P_i^r(a_i|\omega_i,\psi;\theta_{\Pi i}^{r-1},\theta_c^{r-1}) q(\cdot|\theta_{\Pi i}^{r-1},\mu^r)}, 1 \right\}.$$

The computation of the firm-specific likelihood component P_i^r, as defined in (16), requires the computation of the expected value function for the firm, which happens in the inner loop.

1.3 *Common parameters drawing step*

We draw a candidate parameter form the proposal density $\theta_c^{*r} \sim q(\theta_c^{*r}|\theta_c^{r-1})$, then accept θ_c^{*r} with probability λ, where

$$\lambda = \min\left\{ \frac{\pi(\theta_c^{*r}) L^r(\mathbf{a}|\boldsymbol{\omega},\psi;\theta_{\Pi}^r,\theta_c^{*r}) q(\cdot|\theta_c^{*r})}{\pi(\theta_c^{r-1}) L^r(\mathbf{a}|\boldsymbol{\omega},\psi;\theta_{\Pi}^r,\theta_c^{r-1}) q(\cdot|\theta_c^{r-1})}, 1 \right\},$$

where $(\mathbf{a},\boldsymbol{\omega}) \equiv \{a_i,\omega_i\}_{i=1}^N$, and L^r is the joint likelihood defined in (15).

2. **The inner loop**
The inner loop computes and updates the alternative specific value function by applying the Bellman operator once. Imai et al. (2009) propose to approximate the expected value functions by storing and using information from earlier iterations of the algorithm. In particular, storing up to M past accepted draws of parameters and value functions, $\{\theta^{*l}, s^l, V^l(s^l,\epsilon^l;\theta^{*l})\}_{l=r-M}^{r-1}$, Imai et al. (2009) propose to construct the expected value function in iteration r as,

$$E_{\epsilon'}^r\left[V(s',\epsilon';\theta^{*r}|s,a)\right] = \sum_{l=r-M}^{r-1} V^l(s^l,\epsilon^l;\theta^{*l}) \chi(\theta^{*l},\theta^{*r};s^l,s|a), \tag{19}$$

where

$$\chi(\theta^{*l},\theta^{*r};s^l,s|a) = \frac{K_{h_\theta}(\theta^{*l},\theta^{*r}) K_{h_s}(s^l,s|a)}{\sum_{k=r-M}^{r-1} K_{h_\theta}(\theta^{*k},\theta^{*k}) K_{h_s}(s^k,s|a)},$$

so as to assign higher weights to past parameters that are closer the current iteration one, and higher weights to states s' that have higher transition density from states s. $K_{h_\theta}(\theta^{*k},\theta^{*k})$ and $K_{h_s}(s^k,s|a)$ are kernel function with bandwidth h_θ, and h_s, for the parameter vector, θ, and the state variable s, respectively. The value function obtained from (19) is used to construct the choice specific value function,

$$V^r(a,s,\epsilon;\theta^{*r}) = \Pi(a,s;\theta_{\Pi}^{*r}) + \epsilon + \beta E_{\epsilon'}^r\left[V(s',\epsilon';\theta^{*r})|s,a\right]. \tag{20}$$

The value function in (20) is used to construct the likelihood as in (16). Note that the integration over the continuous state variables is already incorporated into the computation of the weighted average of past value functions. This approach has the advantage, compared to Rust's random grid approximation, of avoiding to compute the value function at N_{grid} random points of the state variables state in each iteration.

Finally, given the assumption of *iid* extreme value distributed ϵ's, we have that

$$V^r(s,\epsilon;\theta^{*r}) = \max_{a\in A} V(a,s,\epsilon;\theta^{*r}) = \ln\left[\sum_a \exp(V(a,s;\theta^{*r}))\right]$$

4 Data

In this section, we report the summary statistics of all the variables used to estimate the static and the dynamic structural models. In particular, the upper part of Table 4.1 displays mean, standard deviation, and number of observation of the variables extracted from the PS (Production Survey, Statistics Netherlands) for the years 2002-2008. To estimate the deflated revenue function as in (18), we use the deflated value of gross output Y_{it} ($\equiv \frac{P_{it}Q_{it}}{\tilde{P}_t^j}$) of each firm i in sector j in period t, where $P_{it}Q_{it}$ are the firm's revenues, and $\tilde{P}_t^j$ is the sector j price deflator. Labor (L_{it}) refers to the number of employees in each firm for each year,[7] collected in September of that year. The corresponding wages W_{it} include gross wages plus salaries and social contributions before taxes. The costs of intermediate inputs ($Z_{it}M_{it}$) include costs of energy, intermediate materials, and services. The unit user costs R_{it} (of capital stock K_{it}) are calculated as the sum of the depreciation of fixed assets and the interest charges. Q_t^j indicates the sector-specific production index.

The nominal gross output and intermediate inputs are deflated with the appropriate price indices from the input-output tables available at the NACE rev. 1 two-digits sector classification.[8] For capital, we use a two-digit NACE deflator of fixed tangible assets calculated by Statistics Netherlands. The share of the cost of labor, material, and capital are denoted as s_{iLt}, s_{iMt}, and s_{iKt}, respectively. The share of the cost of labor constitutes 24.2 percent of the gross production value, while materials account for 65.7 percent of gross output, and capital for 4 percent.

The total number of observation, after retaining only the respondents to the different waves of the Community Innovation Survey, is 8306. The CIS datasets are the main data source for measuring innovation in Europe. The surveys are designed to provide an extensive description of the general structure of innovative activities at the sectoral, regional, and country levels, including basic information of the enterprise, product and process innovation, innovation activity and expenditure, effects of innovation, innovation cooperation, public finding of innovation, source of information for innovation patents, and so forth.[9]

The middle part of Table 4.1 reports descriptive statistics for the different types of R&D expenditure extracted from three waves of the Community Innovation Survey (CIS), carried out by Statistics Netherlands. In particular, we constructed an unbalanced panel

[7] For each enterprise, jobs are added and adjusted for part-time and duration factors, resulting in number of men/years expressed as Full Time Equivalents (FTEs)(*Source:* Statistics Netherlands)

[8] NACE Rev. 1 is a 2-digit activity classification which was drawn up in 1989. It is a revision of the General Industrial Classification of Economic Activities within the European Communities, known by the acronym NACE and originally published by Eurostat in 1970.

[9] Community Innovation Survey, EUROSTAT.

of survey respondents, merging the CIS 4 (reference period 2002-2004), the CIS 2006 (reference period 2004-2006), and the CIS 2008 surveys (2006-2008). The R&D expenditures are expressed in thousands of Euros. The intramural expenditure are more than three times larger than the extramural. The average total amount of research expenditure is roughly 3 million Euros. The number of firms that reported R&D spending is 2171 out the total sample of 3565 (unevenly distributed over the period 2002-2008). The last part of Table 4.1 displays the details of the control variable, namely the investment choice k. The most right column reports the total number of firms for each year. For example, in 2002, the number of enterprises that participated to the CIS and that were matched with the PS is 444 , whereas in 2008, the same matching exercise yields a much larger number of firms, i.e., 2413. Our R&D investment variable is constructed as follows. The firm-specific choice na_{it} takes value one if the firm does not engage in any activity other than operating in the market; rd_{it} takes value one if the firm decides to spend in R&D; the investment decision c_{it} takes value of one if the firm has at least one cooperative agreement (with either a firm, a supplier, a customer, or a public (private) research institute); d_{it} match firms' decision to invest in a technological upgrade; action cd_{it} tags the decision to both innovate and cooperate. Concerning the type of investment, the simple production without innovative or cooperative activities is the most frequent, with a total of 3389 observations ($k = na$). Introducing an innovation (product or process, $k = d$), and both innovating and cooperating with either another firm ($k = cd$), or with a research institute are also very frequent answers (2129 and 2530 observation, respectively). On the other hand, the number of firms engaging in only R&D ($k = rd$) or only research alliances ($k = c$) is quite small, with an average of 23 and 13 firms for the rd and c investment choices, respectively.

The cross-sectional data from each wave is expanded so as to cover the whole reference period (there is a one-year overlap between the three waves). For example, if the firm has reported to have introduced an innovation during the reference period, and the same firm has not abandoned the innovation project, then we impute the value 1 for the whole time span.

5 Results

In this section we first present the parameter estimates of the deflated revenue function under imperfect competition in both output and labor markets, (18), and of the state variables evolution, (9), and (11). We then use the estimates of the static parameters to present the results of the dynamic discrete choice model.

5.1 Static parameters

The point estimates of the output price markup and all the parameters used to construct the productivity evolution as in (11) are reported in Table 5.1. The upper part of the table reports demand elasticity parameters, the aggregate state average, and the productivity level and growth.[10] The elasticity of the demand is found to be equal to -2.8, with a corresponding output price markup of 55%. On average, the log productivity is equal to 1.381 and its growth is equal to 1.7%. The aggregate state, ψ_t, is constructed as weighted deflated total industry revenues, $\psi_t \equiv \sum_j \tilde{p}_t^j / N^j (q_t^j)^{1/\eta}$, where $\tilde{p}_t^j$ is the price deflator for industry j at time t, and q_t^j is the weighted average of deflated revenues per industry.

[10] For a complete discussion on the factor input elasticities and the implication of the rent-sharing parameter on productivity growth, we would refer the reader to the paper of Amoroso et al. (2012).

Table 4.1: Summary Statistics

	mean	sd	median	1^{st} quartile	3^{rd} quartile	N. obs
$P_{it}Q_{it}$	63323.97(K Euros)	318679	14881.500	5838.000	39925	8306
L_{it}	152.657	347.055	75	36	152	8306
$Z_{it}M_{it}$	48353.050(K Euros)	280848	9868	3539	27120	8306
$R_{it}K_{it}$	2255.667(K Euros)	26330	359	117	1156	8257
s_{iLt}	0.242	0.124	0.228	0.154	0.310	
s_{iMt}	0.657	0.149	0.663	0.567	0.758	
s_{iKt}	0.040	0.223	0.027	0.013	0.048	
Q_t^j	73.080	10.465	73.498	63.648	80.889	8306
Intramural R&D	1806.574(K Euros)	18396.654	100	10	400	4937
Extramural R&D	612.855(K Euros)	7243.232	0	0	50	4937
R&D Expenditure	3038.461(K Euros)	26356.650	255	63	846	4937
	$k=na$	$k=rd$	$k=c$	$k=d$	$k=cd$	N_t
N_{2002}	153	22	9	136	124	444
N_{2003}	133	9	7	102	167	418
N_{2004}	175	13	6	131	221	546
N_{2005}	179	8	3	130	184	504
N_{2006}	769	28	15	471	557	1840
N_{2007}	907	38	26	553	617	2141
N_{2008}	1073	46	28	606	660	2413
Tot.	3389	164	94	2129	2530	8306

We find the aggregate state to be equal to 1.088, on average. Analyzing the evolution of the aggregate state over the years, we find that the market conditions were stable until 2006 and start worsening in 2007 and 2008. The same pattern is followed by the total factor productivity (TFP) growth. The correlation between ψ_t and productivity is 0.922 (significant at 0.001 significance level). These results confirm that, at an aggregate level, the TFP growth estimated under the assumptions of imperfect competition in both labor and output markets seems to pick up the actual features of the Dutch manufacturing industry.

The aggregate state transition of (9) is specified by the three estimated parameters, the mean, $\hat{\mu}_0 = 0.853$, the autocorrelation, $\hat{\rho} = 0.241$, and the variance, $\hat{\sigma}_\epsilon = 0.114$.

Concerning the parameters of the productivity evolution as in (11), we find evidence of a third order polynomial, and fair dependence on innovation and cooperation. In particular, the estimated coefficient associated with the action of cooperating is significant at the 5% significance level, and equal to 0.076, and that of innovating is equal to 0.113. The coefficient associated with both cooperating and upgrading technology, and the decision to do R&D, are equal to 0.062 and -0.011, respectively. The four means and standard errors of the posterior distributions of the fixed costs are reported in Table 5.2. Assuming that all firms face the same log-normal distribution for all four fixed costs, we find that the fixed costs of R&D and cooperating in R&D (3.0 and 3.5 million Euro, respectively) are substantially higher than the per-period costs of maintaining an innovation (460 thousand Euro). Moreover, the fixed costs of maintaining an innovative activity, while sharing the costs of R&D, decreases the per-period costs (290 thousand Euro). This confirms the rationale behind the cooperating strategies, i.e., the cost sharing motive (Cassiman and Veugelers, 2002; Lopez, 2008; Amoroso, 2011). R&D cooperations, in fact, allow firms to share costs or to reduce risks of innovation. The results for the fixed costs are comparable with those found by Aw et al. (2011) for the Taiwanese elec-

Table 5.1: Demand and productivity evolution parameters

	parameter	estimate	(st.err.)/st.dvt.
Eq. (18)	θ_L	0.266	(0.036)
	θ_M	1.206	(0.114)
	θ_K	0.044	(0.010)
	η	-2.800	(0.428)
	$\eta/\eta+1$	1.555	(0.132)
	μ^W	0.311	(0.050)
	φ	0.332	(0.000)
	ψ_t	1.088	0.178
	ω_{it}	1.381	0.327
	$\Delta\omega_{it}$	0.017	0.225
Eq. (9)	μ_0	0.853	(0.022)
	ρ	0.241	(0.020)
	σ_ϵ	0.114	(0.001)
Eq. (11)	β_0	1.650	(0.020)
	β_1	0.581	(0.043)
	β_2	-0.002	(0.002)
	β_3	0.001	(0.000)
	β_4	0.076	(0.043)
	β_5	0.113	(0.205)
	β_6	0.062	(0.056)
	β_7	-0.011	(0.077)

tronics industry, as they estimate these costs to be on average 67.606 million TW dollars (roughly 1.8 million Euro) Below the posterior means and standard deviation of the fixed costs relative to each innovative activity, we report the probabilities of undertaking the different investments, given the level of productivity and the market conditions. On average, the probability to not engage in any activity is the highest (0.41), followed by the probability to simultaneously cooperate and innovate (0.30), and by the probability to introduce an innovation (0.26). Next to the averages of the probability of choosing action k, we report the same probabilities for the levels of the log productivity at each quartile. As we are interested in understanding the relation between the level of productivity and the probability of undertaking an activity, Figure 1 displays the locally weighted scatterplot smoothing (lowess)[11] curves fitting the relationships between the probabilities to undertake action a and the level of productivity, $\exp(\omega_{it})$. The darker areas of the smoothed scatterplots represent higher density of the data points. The plot at the top reports the curve fitting the relation between the probability of taking no action and the level of productivity. The probability of remaining inactive in research and innovation is inversely related to the productivity. We find the same pattern for the probability of doing R&D and the probability of introducing an innovation. Simply put, the higher the firm level productivity, the smaller the probability of investing in R&D, or innovating. However, the situation is reversed when the investment in R&D or in a new product or process is shared with a partner. Indeed, when cooperating, the probabilities of doing research, $Pr(a = c|s, \theta)$, and innovating, $Pr(a = cd|s, \theta)$, are (non monotonic) increasing functions of productivity. This pattern could point to the presence of knowledge external-

[11]Locally weighted regression fitting techniques provide a generally smooth curve, the value of which at a particular location along the x-axis is determined only by the points in that vicinity. The method consequently makes no assumptions about the form of the relationship, and allows the form to be discovered using the data itself.

Table 5.2: Fixed costs

	posterior mean($\times 1mln$)		std error	
$\theta_i^{FC}(rd)$	3.025		0.082	
$\theta_i^{FC}(c)$	3.528		0.100	
$\theta_i^{FC}(d)$	0.459		0.025	
$\theta_i^{FC}(cd)$	0.286		0.025	
	mean	$\omega_{it} \leq 1.172$	$\omega_{it} \leq 1.347$	$\omega_{it} \leq 1.541$
$P_i(a = na \mid s, \theta)$	0.408	0.409	0.415	0.414
$P_i(a = rd \mid s, \theta)$	0.020	0.024	0.021	0.021
$P_i(a = c \mid s, \theta)$	0.011	0.008	0.009	0.011
$P_i(a = d \mid s, \theta)$	0.256	0.269	0.263	0.260
$P_i(a = cd \mid s, \theta)$	0.304	0.288	0.289	0.293

ities. These results, together with the evidence of the endogenous firm-level productivity, positively associated with the action of cooperating, suggest that an innovation policy aiming at encouraging research cooperation might result in a virtuous cycle. Indeed, past investments in cooperative research have a positive impact on current productivity, which, in turn, positively influence the probability to engage in both R&D and innovation when these activities are shared with a research partner. Figure 2 plots the MCMC draws of the fixed cost parameters. It appears that the the MCMC draws converge after 50 iterations.

5.2 Dynamic parameters

In this section we present the results for the DDP model presented in (13). Once the fixed costs are estimated, we can subtract them from the profit function as in (7). For simplicity, we estimate the model without unobserved heterogeneity. Therefore, the standard deviations σ_Π are set equal to zero. The discount factor is fixed at 0.93. During this stage we are able to recover both fixed and sunk costs of doing R&D or innovating with or without a research partner. Figure 2 shows that the sunk cost parameters converge at different rates, and, in general, much slower than the fixed costs.

The estimated coefficients are reported in Table 5.3. Next to the mean values of the sunk costs, we report the standard deviations of the MCMC draws. The values are estimated with the expected signs. Sunk costs are found to be 4 millions for the average firm that undertakes R&D with or without a partner, 14 to 33% higher than the fixed costs. The sunk costs of innovation are still much smaller than the ones of research, but 3 to 3.5 times higher than the fixed costs of innovating. Moreover, we find additional evidence of the risk-sharing motive behind the decision to introduce an innovation. In fact, the average sunk costs of producing an innovation with a research partner is almost one third smaller than the average sunk costs of undertaking the same project without an alliance (997,000 Euro and 1.4 million Euro, respectively). The sunk costs parameters cannot be compared with the reported R&D expenditures. This is because the sunk costs can be related to productive factors, such as labor and/or capital that are allocated to research rather than to production. For this reason, these costs will not appear in the balance sheets of the company (Santos, 2009).

Next to the estimation of the sunk cost parameters, we show the importance of the role played by these costs in shaping the probabilities if undertaking the different research

Table 5.3: Dynamic Parameter Estimates

	posterior mean	std error
$\theta_i^{SC}(rd)$	3.984	0.570
$\theta_i^{SC}(c)$	4.046	0.216
$\theta_i^{SC}(d)$	1.433	0.560
$\theta_i^{SC}(cd)$	0.997	0.216

$\theta_i^{SC}(rd)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.141	0.264	0.367
$P_i(a = rd\|s, \theta)$	0.049	0.630	0.008
$P_i(a = c\|s, \theta)$	0.037	0.006	0.009
$P_i(a = d\|s, \theta)$	0.002	0.048	0.390
$P_i(a = cd\|s, \theta)$	0.770	0.051	0.226
$\theta_i^{SC}(c)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.484	0.693	”
$P_i(a = rd\|s, \theta)$	0.006	0.002	”
$P_i(a = c\|s, \theta)$	0.065	0.008	”
$P_i(a = d\|s, \theta)$	0.263	0.109	”
$P_i(a = cd\|s, \theta)$	0.182	0.188	”
$\theta_i^{SC}(d)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.668	0.433	”
$P_i(a = rd\|s, \theta)$	0.002	0.005	”
$P_i(a = c\|s, \theta)$	0.002	0.007	”
$P_i(a = d\|s, \theta)$	0.225	0.407	”
$P_i(a = cd\|s, \theta)$	0.103	0.147	”
$\theta_i^{SC}(cd)$	-50%	-25%	0%
$P_i(a = na\|s, \theta)$	0.616	0.541	”
$P_i(a = rd\|s, \theta)$	0.007	0.003	”
$P_i(a = c\|s, \theta)$	0.004	0.004	”
$P_i(a = d\|s, \theta)$	0.129	0.162	”
$P_i(a = cd\|s, \theta)$	0.244	0.290	”

investments. Table 5.3 also reports the changes in probabilities associated with 50% and 25% reductions in the costs of engaging in research and/or innovating. A reduction in the sunk costs of R&D, cooperating, and innovating can be thought of as an example of an innovation policy, such as a subsidy to R&D start up, or public procurement. Results show that a 25% reduction in these costs is expected to increase the probability of undertaking the corresponding activity. For example, reducing the costs of R&D, $\theta_i^{SC}(rd)$ of 25% leads to an increase of probability of doing R&D of 62.2%.

6 Conclusion

In this paper, we present empirical evidence of the fixed and sunk costs of investments in research activities, and quantify the linkages between the cost structure, firm-level productivity, and the probabilities to technologically upgrade. In particular, we propose and estimate a structural model with endogenous choices of technological upgrade for the Dutch manufacturing industry. The model describes a firm's dynamic decision process for undertaking different research activities, namely, innovating and conducting R&D, with or without a research partner. The R&D investment choices are endogenous, as they depend on the firm's level of productivity, an aggregate measure of industry competition, fixed and sunk costs of R&D, and past research choices. To our knowledge, none of the existing studies proposes and estimates a dynamic structural model to derive the total cost function of firms engaging in technological activities.

We find that the firm's probability to do R&D or to introduce an innovation increases with the level of productivity, only when this activity is shared with a research partner. Moreover, according to the literature on R&D cooperation, the costs of innovating are smaller when cooperating. In fact, given the higher risks associated with the uncertainty of the market demand for new products or processes, the firm might allocate more importance to the cost/risk sharing rationale for this type of innovative activities, rather than for the sheer research investments.

Sunk costs are found to be roughly 1.5 times larger than the fixed costs of research (both cooperative and private), and 3 to 3.5 times larger than the fixed costs of innovating. Moreover, we show the importance of the role played by these costs in shaping the probabilities if undertaking the different research investments. In general, a reduction in the sunk costs of R&D, cooperating, and innovating increases the probability of undertaking the corresponding activity.

Additionally, we present some preliminary conclusions on innovation policies aiming at encouraging research cooperation. We show how these type of policy interventions might result in a virtuous cycle. Indeed, past investments in cooperative research have a positive impact on current productivity, which, in turn, positively influences the probability to engage in both R&D and innovation when these activities are shared with a research partner. Therefore, in elaborating their policies for innovation, governments must ensure to create frameworks that encourage the collaboration throughout the innovation process.

References

Abraham, F., Konings, J., and Vanormelingen, S. (2009). The effect of globalization on union bargaining and price–cost margins of firms. *Review of World Economics*, 145:13–36.

Ackerberg, D., Caves, K., and Frazer, G. (2006). Structural Identification of Production Functions. Technical report, UCLA mimeo.

Aguirregabiria, V. and Mira, P. (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica*, 70(4):pp. 1519–1543.

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, 156(1):38–67.

Amoroso, S. (2011). Bayesian analysis of r&d cooperation determinants. *Working Paper*.

Amoroso, S., Melenberg, B., Plasmans, J., and Vancauteren, M. (2012). Firm level productivity under imperfect competition in output and labor markets. *Working Paper*.

Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867.

Aw, B. Y., Roberts, M. J., and Xu, D. Y. (2011). R&d investment, exporting, and productivity dynamics. *American Economic Review*, 101(4):1312–44.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Belderbos, R., Carree, M., Diederen, B., Lokshin, B., and Veugelers, R. (2004a). Heterogeneity in R&D cooperation strategies. *International Journal of Industrial Organization*, 22(8-9):1237–1263.

Belderbos, R., Carree, M., and Lokshin, B. (2004b). Cooperative r&d and firm performance. *Research Policy*, 33(10):1477–1492.

Bughin, J. (1993). Union-firm efficient bargaining and test of oligopolistic conduct. *The Review of Economics and Statistics*, 75(3):563–567.

Bughin, J. (1996). Trade unions and firms' product market power. *The Journal of Industrial Economics*, 44(3):pp. 289–307.

Carboni, O. A. (2012). An empirical investigation of the determinants of r&d cooperation: An application of the inverse hyperbolic sine transformation. *Research in Economics*, 66(2):131 – 141.

Cassiman, B. and Veugelers, R. (2002). R&d cooperation and spillovers: Some empirical evidence from belgium. Open access publications from katholieke universiteit leuven, Katholieke Universiteit Leuven.

Ching, A. T., Imai, S., Ishihara, M., and Jain, N. (2012). A practitioner's guide to bayesian estimation of discrete choice dynamic programming models. *Quantitative Marketing and Economics*, 10:151–196.

Crépon, B., Desplatz, R., and Mairesse, J. (2002). Price-cost margins and rent sharing: Evidence from a panel of french manufacturing firms. *Revised version of CREST Working Paper No. G9917*.

Crépon, B., Duguet, E., and Mairessec, J. (1998). Research, innovation and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technology*, 7(2):115–158.

De Loecker, J. (2011). Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity. *Econometrica*, 79(5):1407–1451.

Dobbelaere, S. (2004). Estimation of price-cost margins and union bargaining power for belgian manufacturing. *International Journal of Industrial Organization*, 22:1381–1398.

Dobbelaere, S. and Mairesse, J. (2011). Panel data estimates of the production function and product and labor market imperfections. *Journal of Applied Econometrics*.

Doraszelski, U. and Jaumandreu, J. (2008). R&d and productivity: Estimating production functions when productivity is endogenous. CEPR Discussion Papers 6636, C.E.P.R. Discussion Papers.

Ericson, R. and Pakes, A. (1995). Markov-perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, 62(1):pp. 53–82.

Griliches, Z. (1980). R & d and the productivity slowdown. *The American Economic Review*, 70(2):pp. 343–348.

Grimpe, C. and Kaiser, U. (2010). Balancing internal and external knowledge acquisition: The gains and pains from r&d outsourcing. *Journal of Management Studies*, 47(8):1483–1509.

Hall, B. H. and Mairesse, J. (1995). Exploring the relationship between r&d and productivity in french manufacturing firms. *Journal of Econometrics*, 65(1):263–293.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):pp. 497–529.

Howells, J. (1999). Research and technology outsourcing. *Technology Analysis & Strategic Management*, 11(1):17–29.

Imai, S., Jain, N., and Ching, A. (2009). Bayesian estimation of dynamic discrete choice models. *Econometrica*, 77(6):1865–1899.

Jones, C. I. and Williams, J. C. (1998). Measuring the social return to r&d. *The Quarterly Journal of Economics*, 113(4):1119–1135.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–341.

Lopez, A. (2008). Determinants of r&d cooperation: Evidence from spanish manufacturing firms. *International Journal of Industrial Organization*, 26(1):113–136.

Magnac, T. and Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, 70(2):801–816.

Norets, A. (2009). Inference in dynamic discrete choice models with serially orrelated unobserved state variables. *Econometrica*, 77(5):1665–1682.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):pp. 1263–1297.

Pakes, A., Ostrovsky, M., and Berry, S. (2007). Simple estimators for the parameters of discrete dynamic games (with entry/exit examples). *The RAND Journal of Economics*, 38(2):373–399.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica*, 55(5):999–1033.

Rust, J. (1994). Chapter 51 structural estimation of markov decision processes. In Engle, R. F. and McFadden, D. L., editors, *Handbook of Econometrics*, volume 4 of *Handbook of Econometrics*, pages 3081 – 3143. Elsevier.

Santos, C. D. (2009). Recovering the sunk costs of r&d: the moulds industry case. CEP Discussion Papers dp0958, Centre for Economic Performance, LSE.

Wang, X. H. and Yang, B. Z. (2001). Fixed and sunk costs revisited. *Journal of Economic Education*, 32(2):178–185.

A Profit function

Given the following maximization problem

$$\max_{L_{it},W_{it}} \left[\phi_{it}\log(U_{it}(W_{it},L_{it})) + (1-\phi_{it})\log\Pi_{it}\right],$$

the first order conditions can be written as:

$$w.r.t.\quad L_{it} \quad\rightarrow\quad (1-\phi_{it})\frac{W_{it} - \left(1+\frac{1}{\eta}\right)P_{it}(Q_{it})\frac{\partial Q_{it}}{\partial L_{it}}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}, \tag{21}$$

$$w.r.t.\quad W_{it} \quad\rightarrow\quad (1-\phi_{it})\frac{W_{it}-\bar{W}_{it}}{\Pi_{it}} = \frac{\phi_{it}}{L_{it}}. \tag{22}$$

Combining equations (21) and (22), the marginal revenue product of labor is

$$\left(\frac{\eta+1}{\eta}\right)P_{it}(Q_{it})\frac{\partial Q_{it}}{\partial L_{it}} = \bar{W}_{it}. \tag{23}$$

Therefore, by multiplying both sides of (23) by $\frac{L_{it}}{Q_{it}}$, we have

$$\frac{\eta+1}{\eta}\theta_{iLt} = \frac{\bar{W}_{it}L_{it}}{P_{it}(Q_{it})Q_{it}} = \frac{\bar{W}_{it}}{W_{it}}\frac{W_{it}L_{it}}{P_{it}(Q_{it})Q_{it}}.$$

Using Amoroso et al. (2012) definition of the wage markup $\mu_{it}^W \equiv \frac{W_{it}-\bar{W}_{it}}{W_{it}}$, and taking into account the demand as in (1), we can rewrite the cost of labor as

$$W_{it}L_{it} = \frac{1+\eta}{\eta}\theta_{iLt}\frac{1}{1-\mu_{it}^W}(Q_{it})^{\frac{1+\eta}{\eta}}\frac{P_t^j}{(Q_t^j)^{1/\eta}}\exp(-u_{it}^d/\eta).$$

Replacing Q_{it} with the Cobb-Douglas function as in (2), and solving for L_{it}, we get

$$L_{it} = \left[(\exp(\theta_0+\omega_{it})K_{it}^{\theta_K}M_{it}^{\theta_M})^{\frac{\eta+1}{\eta}}\frac{1}{1-\mu^W}\frac{\eta+1}{\eta}\frac{\theta_{iLt}}{W_{it}}\frac{P_t^j}{(Q_t^j)^{1/\eta}}(\exp(-u_{it}^d/\eta))\right]^{\eta/(\eta-\theta_{iLt}(\eta-1))}. \tag{24}$$

The short-run profits, $P_{it}Q_{it} - W_{it}Lit$, can be rewritten as

$$\Pi^{SR}(\omega_{it},W_{it},K_{it},M_{it},\psi_t) = (\exp(\theta_0+\omega_{it})K_{it}^{\theta_K}L_{it}^{\theta_L}M_{it}^{\theta_M})^{\frac{1+\eta}{\eta}}\frac{P_t^j}{(Q_t^j)^{1/\eta}}\exp(-u_{it}^d/\eta)\left[1-\frac{1+\eta}{\eta}\theta_{iLt}\frac{1}{1-\mu_{it}^W}\right].$$

Replacing the labor demand with (24), we get the final profit function:

$$\Pi^{SR}(\omega_{it},W_{it},K_{it},M_{it},\psi_t) = \left(\frac{1-\gamma}{\gamma^{1-\delta}}\right)W_{it}^{1-\delta}\left[\left(\exp(\theta_0+\omega_{it})K_{it}^{\theta_K}M_{it}^{\theta_M}\right)^{\frac{\eta+1}{\eta}}\left(\psi_t(\exp(u_{it}^d))^{-1/\eta}\right)\right]^{\delta}$$

where $\psi_t \equiv \frac{P_t^j}{(Q_t^j)^{1/\eta}}$, $\gamma \equiv \theta_L\frac{\eta+1}{\eta}\frac{1}{1-\mu^W}$, and $\delta \equiv \eta/(\eta-\theta_{iLt}(\eta-1))$.

The short-run profit function as in (8), assuming no imperfect competition on the labor market, is derived from the following optimization problem for firm i:

$$\max_{X_{it}} \left\{P_{it}Q_{it} - V_{it}'X_{it} \mid A_{it}F(X_{it}) \geq Q_{it}\right\}, \tag{25}$$

where $X_{it} \equiv (X_{i1t}, X_{i2t}, \ldots, X_{irt})'$ denotes the vector of r factor inputs, $F(.)$ is production function, and $V_{it} \equiv (V_{i1t}, V_{i2t}, \ldots, V_{irt})'$ is the vector of r input prices. Taking into account the demand as in (1), the FOC is:

$$\frac{\eta+1}{\eta} P_{it} \frac{\partial Q_{it}}{\partial X_{it}} = V_{it},$$

since $MC_{it}^X = V_{it} \frac{\partial X_{it}}{\partial Q_{it}}$ is defined as the marginal cost of X_{it}, we have that

$$\frac{P_{it} - MC_{it}^X}{P_{it}} = -\frac{1}{\eta}. \tag{26}$$

Assuming that the marginal cost of X_{it} are an inverse function of the firm-level productivity such as

$$MC_{it}^X \equiv \frac{1}{\exp(\omega_{it})},$$

the price can be expressed as a function of the demand elasticity and the productivity,

$$P_{it} = \frac{\eta}{\eta+1} \frac{1}{\exp(\omega_{it})}. \tag{27}$$

Multiplying (26) by $P_{it} Q_{it}$, we obtain the profits, therefore the profit function can be written as

$$\Pi_{it} = -\frac{1}{\eta} P_{it} Q_{it}.$$

Substituting Q_{it} with (1) and P_{it} with (27), we obtain the following short-run profit function:

$$\Pi(\omega_{it}, \psi_t) = \varphi \psi_t \exp(\omega_{it})^{-(1+\eta)},$$

where $\varphi \equiv -\frac{1}{1+\eta} \left(\frac{\eta}{1+\eta}\right)^{\eta}$.

B Tables and Figures

Figure 1: Investment policy functions

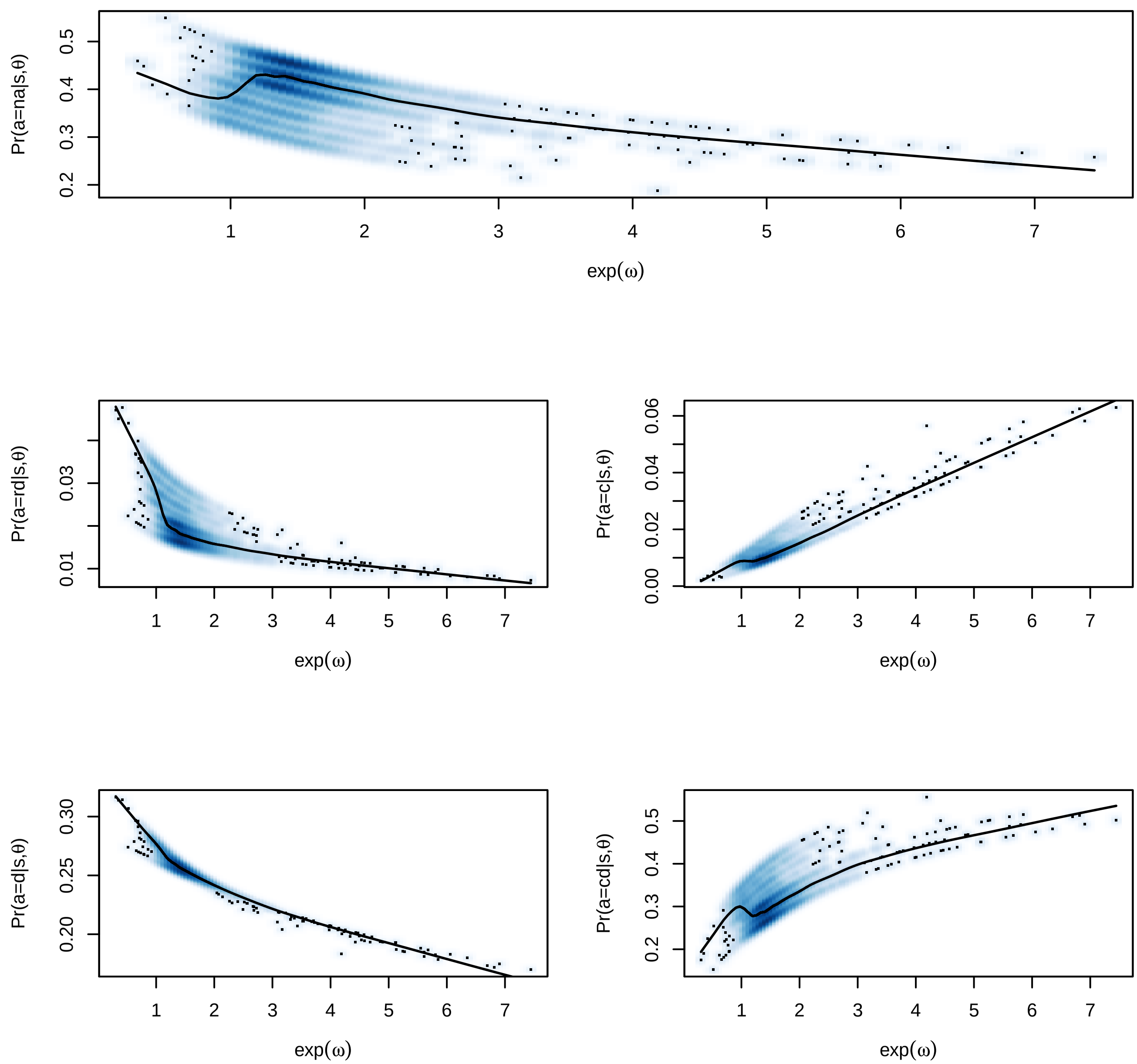
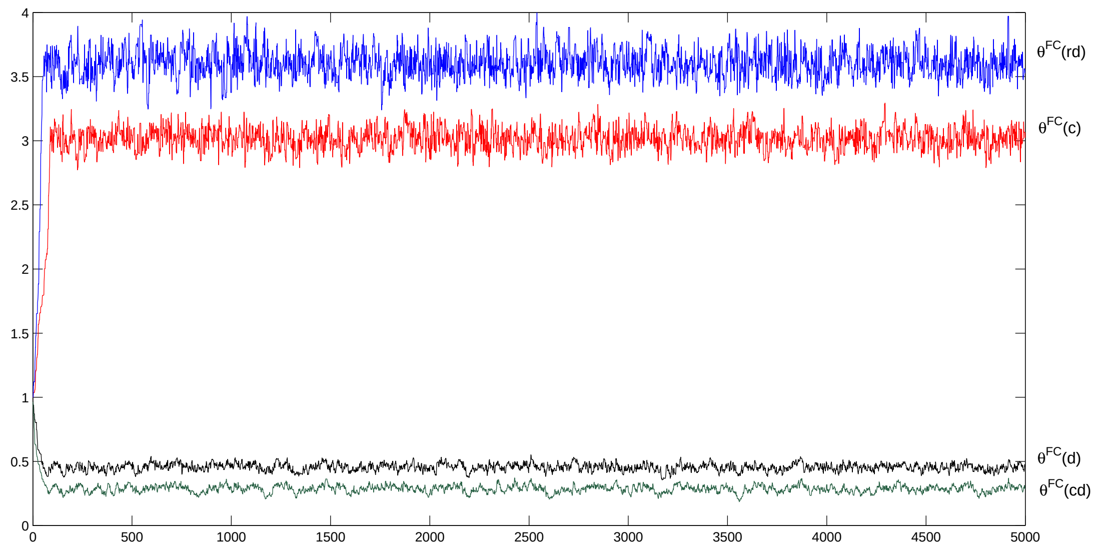
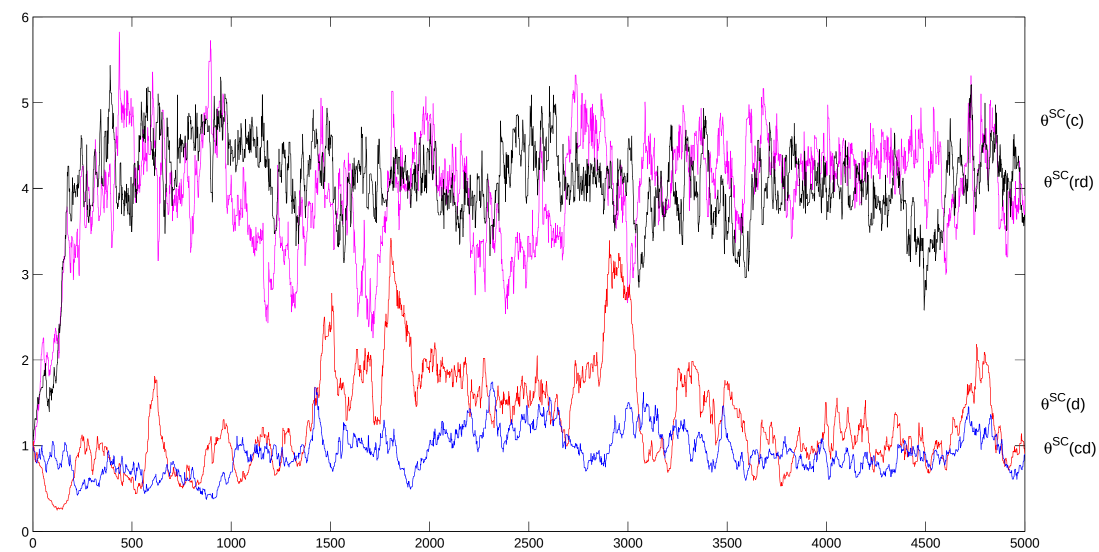


Figure 2: MCMC iterations of fixed and sunk cost parameters





Note: MCMC plots of θ^{FC} and θ^{SC}

Figure 3: R&D, Cooperation and Innovation Choices