European Commission

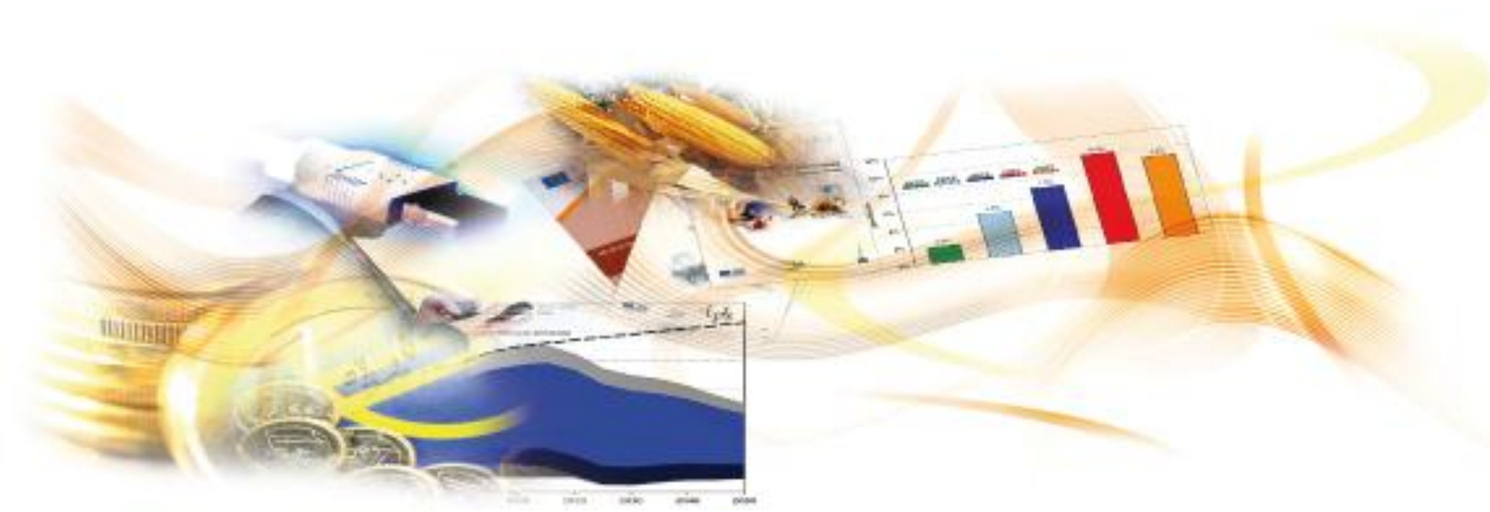
JRC TECHNICAL REPORTS

IPTS Working Papers on Corporate R&D and Innovation – No 02/2014

The hidden costs of R&D collaboration

Sara Amoroso

2014



Joint Research Centre

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

The **main author** of this paper is: Sara Amoroso (European Commission, JRC-IPTS, Knowledge for Growth Unit, Economics of Industrial Research and Innovation Action),
Acknowledgement: We wish to thank Bertrand Melenberg, Joseph Plasmans, Tobias Klein (University of Tilburg), Elena Huergo, Lourdes Moreno (Universidad Complutense Madrid), Dirk Czarnitzki (KULeuven), Mark Vancauteren (Hasselt University) and the participants to the Earie conference 2013, for their useful remarks on a previous version of this paper. We are responsible for all the remaining errors.

Contact information
Fernando Hervás Soriano
Address: European Commission, Joint Research Centre - Institute for Prospective Technological Studies
Edificio Expo. C/ Inca Garcilaso, 3 E-41092 Seville (Spain)
E-mail: jrc-ipts-kfg-secretariat@ec.europa.eu
Tel.: +34 95 448 84 63
Fax: +34 95 448 83 26
IPTS website: https://ec.europa.eu/jrc/en/institutes/ipts ; JRC website: https://ec.europa.eu/jrc/

This publication is a Technical Report by the Joint Research Centre of the European Commission.

JRC90991

ISSN 1831-9408 (online)

Spain: European Commission, Joint Research Centre, 2014

Abstract

Many European policy initiatives continue to promote R&D collaboration in view of its expected benefits. Despite the advantages of R&D cooperation, to benefit from it, firms must create a structure to support the efficient transfer of knowledge-based assets. In fact, the set-up and administration of common resources might be costly. This paper derives the distribution of the costs associated with R&D collaboration, as they could shape firms' R&D- related investments. To ascertain these costs, we model the expected benefits from R&D cooperation with a structural dynamic monopoly model. The modelling results show that the sunk costs of innovation are lower when collaborating with a research partner, and that a firm's probability of investing in R&D or innovation increases with the level of productivity, only when collaborating in R&D and innovation. We also find that the sunk costs of innovation are 1.5 to 3 times lower than the sunk costs of R&D. Additionally, it can be seen that the suggested structural framework of a firm's heterogeneity in cost functions used in our model can offer a straightforward extension to existing policy impact evaluation.

JEL classification: D22, D23, L14, L60, O32.

Keywords: R&D cooperation, transaction costs, dynamic structural model.

1 Introduction

Greater competition, the demand for higher-quality customised products and services, and fast delivery are just a few of the characteristics that shape the organisation of many industries. In this setting, rarely does a single company have the full range of expertise needed for prompt and cost-effective innovation. As a consequence, R&D alliances have become more common. Such alliances allow firms to access new technologies, realise economies of scale and scope in their R&D activities, and shorten development time and, therefore, the time to market. However, despite the evident advantages of cooperation in these alliances, to benefit from R&D collaboration, firms must create a structure to support the efficient transfer of knowledge-based assets and, at the same time, to minimise any unintended leakage of such valuable assets to potential competitors. From a transaction cost theory perspective, the organisational structures needed to form a R&D alliance are costly. The costs associated with the set-up and administration of common resources and know-how needed for the implementation of a collaboration are typically not recognised and are thus termed as "hidden cost". However these costs can be crucial, as firms are more likely to cooperate when the benefits from the enhanced knowledge transfer and control are greater than the additional set-up and administration costs (i.e. the hidden costs).

Therefore, in this paper we investigate the barriers to collaboration in terms of the hidden transaction costs, by first developing a model to estimate the distribution of operating costs (henceforth, referred to as "fixed costs") and sunk costs (costs that have already been incurred and that cannot be recovered) associated with firms' investment choices in R&D and innovation activities with and without a research partner. Indeed, as firms are assumed to be forward-looking and to take into account the implications of their decisions (and the associated costs) on their future pay-offs, decisions to invest in R&D and in innovation (with or without cooperation) are assumed to be costly to reverse and, therefore, more associated with sunk costs.

To ascertain both the fixed and sunk costs of R&D and innovation activities, with and without a research partner, we developed a structural dynamic monopoly model to quantify the linkages between R&D spending, innovation, and cooperation investment choices.

We consider R&D investment and (product or process) innovation (i.e. technological upgrades) as separate strategies, because they entail different levels of uncertainty and risk. It is well known that the processes of R&D and innovation are closely correlated with each other, but the two processes can differ according to e.g. the nature of the innovation. If the innovation is radical (i.e. groundbreaking innovation, creating a new market), then the risks associated with these types of innovation are higher, but so is the appropriability. In this case, the incentives to cooperate may differ and diverge from the incentives to cooperate that occur with only incremental innovations, i.e. those where the risks to be shared are typically smaller as the product/process is closer to commercialisation. To the best of our knowledge, this paper is the first reported attempt to explicitly model and derive the costs of cooperation by adapting a transaction costs perspective to a dynamic structural framework.

To estimate our model, we merge data on the sales, labour, physical capital, price indices for deflating total sales and material inputs of Dutch manufacturing firms extracted from the

Production Survey[1] (PS), and three waves of the Community Innovation Survey[2] (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 firms base their decisions to undertake an investment in R&D or in innovation with or without a research partner on past choices, firm-level total factor productivity, and an aggregate industry demand shifter.

When collaborating in R&D and innovation, the probability to invest in R&D or to innovate increases with the level of productivity. On costs, we found that the sunk costs of innovation are smaller when collaborating with a research partner. And, furthermore, that the sunk costs of innovation are 1.5 to 3 times smaller than the sunk costs of R&D, depending on whether the costs are shared or not, respectively.

In addition, simulating a reduction in the sunk costs of R&D cooperation and innovation can be thought of as an example of modelling an innovation policy intervention, such as a subsidy to start up R&D, or even an example of public procurement. Here, our results show that a 25% reduction in these sunk costs could increase the probability of investing in cooperative innovation, but not the probability to cooperate in R&D, where a costs reduction of up to 50% is needed, yet only increases the probability of cooperating from 0.9% to 6.5% (see Table 6.4. Therefore, the use of a structural framework to describe firms' heterogeneity in cost functions can provide a straightforward extension to a policy impact evaluation.

This paper is set out as follows. Following the introduction, in section 2 we briefly summarise the relevant literature about transaction cost theory concerning R&D cooperation and firms' heterogeneity. Section 3 then presents the model that we used to obtain information on both the fixed and sunk costs, and consequently on the optimal R&D, innovation, and co-operation decisions. Section 4 discusses the empirical strategy we used to obtain estimates of the static parameters of the model. Namely, we show how we obtained a measure of the firm-level productivity, the 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 pay-offs in its profits maximisation. Section 5 describes the steps of the algorithm developed by Imai et al. (2009) that we used to obtain the dynamic parameters estimates. Sections 6 and 7 describe the data and the results, respectively. In Section 8, we present the results of a policy simulation, and finally, in the last section we present our conclusions and put forward some suggestions for further work in this research area.

2 Firms' heterogeneity and transaction cost theories

This section provides an overview of the existing literature overarching the transaction cost theory perspective of R&D cooperation and firms' heterogeneity.

[1] Production statistics, Statistics Netherlands
[2] Community Innovation Survey, EUROSTAT.

This section provides an overview of the existing literature overarching the transaction cost theory perspective of R&D cooperation and firms' heterogeneity. According to transaction cost theory, a transaction can be internal or external to an organisation. The costs associated with the transaction occurs when a good or service is transferred from a provider to a user. The costs depend on how the transaction is organised: if the transaction occurs within an organisation, costs include the managing and the monitoring of the personnel and the acquisition of inputs; if buying the good or service (e.g. R&D) from an external provider, costs include the additional source selection, contract management, and performance management (Williamson, 1981, 1989). Within the transaction costs context, Brockhoff (1992) illustrates empirically that the success of a R&D collaboration depends on the perception of the transaction costs of the R&D cooperation, which is defined by the uncertainty of the project, the specificity of the assets employed, and the frequency of the transactions. These aspects, in turn, are determined by the exogenous independent variables, such as the formality of the agreement, previous experience with R&D cooperation, internal competence, the number of partners involved in the agreement, and the stage in the technological life cycle. In fact, cooperation in the early stages of a technological life cycle might involve a high degree of uncertainty, whereas cooperation in the later stages of a R&D project might be characterised by the use of more specific resources associated to a lower uncertainty.

Focusing more on the probability of vertical R&D collaboration, Oerlemans and Meeus (2001) extended the transaction cost models to incorporate the impact of a firm's resource base on the probability of R&D cooperation. While the transaction cost approach of Williamson (1985)[pp. 142-144] relates the emergence of governance structures (in support of a R&D alliance) to the technological innovation potential to realise cost savings at a firm's level, Oerlemans and Meeus (2001) empirically showed how the innovation potential rests on the underlying assumption of the presence of a knowledge base within the firm. Indeed, in order to exploit the gains from a collaboration, a company must be able to recognise the potential of a cost-saving innovation and must be capable of then following through with its development.

In the literature on industry dynamics, the seminal theoretical models Jovanovic (1982); Hopenhayn (1992); Ericson and Pakes (1995) explain the patterns behind individual firms' success and failure rates, and the overall evolution of the industry structure that was observed in a firm-level panel data set, given their stochastic productivity changes over time. In their models, firms are endowed with an exogenous level of productivity, randomly drawn from some distribution. The "lucky" firms, i.e. those with high productivity, survive and prosper, while the others fail and eventually exit the market.

More modern literature on industrial organization (IO) relaxes this exogeneity assumption, since an important source of the productivity differentials across firms is related to R&D and innovation activities[3] Consequently, a large number of empirical studies have estimated the effect of R&D investment on such growth, and have found that R&D spending has a significant 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. In

[3] Many authors have studied the connection between spending on R&D and productivity growth. Griliches (1980) provides an extensive survey of the empirical literature linking own-firm-R&D spending, R&D spillovers, and productivity growth.

particular, Doraszelski and Jaumandreu (2013) assumed productivity to be dependent on the investment decisions taken. The investment decisions they studied concerned past R&D expenditure (Doraszelski and Jaumandreu, 2013), or both R&D expenditure and export market participation (Aw et al., 2011).

However, any firm that wants to survive must not only be innovative, but must also be ready to outsource knowledge and to develop research networks. In fact, many firms increasingly rely on the external acquisition of new technological knowledge, as the institutional locations of such resources can be quite disparate or may not even exist in the firm. Although not the primary source of produced knowledge, R&D outsourcing[4] (so-called external R&D) has gained considerably in importance over recent years and now accounts for a substantial share of the total innovation expenditure in a large number of firms. In other studies dealing with R&D cooperation, alongside the transaction cost theory, the dimensions of the risks and the costs of innovation, as well as the need to exploit complementary resources, are considered as the main motives for cooperative behaviour, and therefore, this cooperative behaviour may be positively related to addressing a number of obstacles, such as high risk and the cost of innovation (Belderbos et al., 2004a,b; Carboni, 2012). R&D cooperation, in fact, allows firms to share costs and/or to reduce the risks of innovation.

In contrast to these studies, we propose a new methodology for deriving information on the barriers to collaboration, in terms of the hidden transaction costs. Our methodology derives the distribution of the costs associated with firms' investment choices in R&D and innovation activities with and without a research partner.

From our modelling, we hypothesise that cooperating in R&D could reduce both the fixed costs and the sunk costs of introducing an innovation to the market.

3 Structural Framework

The empirical model we used builds on the class of models developed in dynamic entry games in IO, where the dependent variable is the firm's decision to enter or not enter in to a market. In the same spirit, this paper defines the entry decision as actually the adoption of a set of discrete decisions: the decision to invest in R&D, to cooperate, to innovate, and to take part in innovation cooperation. These decisions are assumed to be costly to reverse and, therefore, are associated with sunk costs.[5] As firms are assumed to be forward-looking, it is presumed that they take into account the implications of their decisions (and the associated costs) on their future pay-offs. In our model, time is discrete and indexed by t. The single-agent dynamic

[4]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 organisations, as well as by some private non-profit and related hybrid organisations (Howells, 1999; Grimpe and Kaiser, 2010)

[5]We consider R&D investment and product or process innovation (i.e. technological upgrades) as separate strategies, because they entail different levels of uncertainty and risk. It is well known that the processes of R&D and innovation are closely correlated with each other, but still the two processes can differ according e.g. to the nature of the innovation. If the innovation is radical (i.e. groundbreaking innovation, creation of a new market), then the risks associated with these types of innovation are higher, but so is the appropriability. In this case, the incentives to cooperate may differ and diverge from the incentives to cooperate that occur with only incremental innovations, where the risks to be shared are typically smaller, i.e. those where the product/process is closer to commercialisation.

optimisation problem is solved for 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. In the dynamic dimension, a firm makes investment choices 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 a value of one if the firm does not engage in any activity other than operating in the market; rd_{it} takes a value of one if the firm decides to invest in R&D; choices c_{it} and d_{it} match firms' decisions to start a R&D collaboration or to invest in a technological upgrade or innovation, respectively; action cd_{it} marks the decision to both innovate and cooperate (e.g. with another firm, research institute, or supplier/customer).

3.1 Static decisions

In every time period, firms are competing on prices following the static Bertrand model pattern. 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 quantity demanded 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 function, with the gross output Q_{it} of firm i at time t a 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 the capital, L_{it} the labour, and M_{it} the 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, labour, 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 the sum of the mean productivity level across firms and over time, θ_0, and the productivity shock observable by the firm, but not by the econometrician (for example, managerial ability, quality of research), ω_{it}.

In line with the literature on imperfect competition in both product and labour markets (Bughin, 1993, 1996; Crépon et al., 2002; Dobbelaere, 2004; Abraham et al., 2009; Dobbelaere and Maires 2011; Amoroso, 2013)[chapter 2], we relax the conventional assumption of perfect competition in the labour market, allowing both firms and workers' unions to have some market power. In particular, the workers bargain with the firm over both the levels of employment L_{it}, and 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 (i.e. costs that do not vary with the quantity of output produced, and which are furthermore not irrevocably committed (Wang and Yang, 2001)), depending on capital, material, and innovation 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 the 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 maximises its profits only with respect to the variable costs, namely, the cost of labour. Amoroso (2013)[chapter 2] showed that by maximising with respect to labour, and taking into account the demand curve faced by the monopolistically competitive firm, the results could be expressed with the following equation for the elasticity of the labour input factor:

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

Amoroso (2013)[chapter 2] defined the bargained wage rate $\mu_{it}^W \equiv \frac{W_{it} - \bar{W}_{it}}{W_{it}}$ as the *wage mark-up*[6] From (4), after solving for L_{it} (see appendix B), we derive the following expression for labour:

$$L_{it} = \left[(\exp(\theta_0 + \omega_{it}) K_{it}^{\theta_K} M_{it}^{\theta_M})^{\frac{\eta+1}{\eta}} \frac{1}{1 - \mu^W} \frac{\eta + 1}{\eta} \frac{\theta_{iLt}}{W_{it}} \frac{P_t^j}{(Q_t^j)^{1/\eta}} (\exp(u_{it}^d))^{-1/\eta} \right]^{\eta/(\eta - \theta_{iLt}(\eta-1))} \tag{5}$$

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

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

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 $\delta \equiv \eta/(\eta - \theta_{iLt}(\eta - 1))$.

3.2 Dynamic decisions

In this study, we assume that the decisions to undertake R&D, to cooperate, or to innovate cannot be revoked, so we assume the costs associated with these actions to be sunk. We define the vector of fixed costs paid in the case of investment in R&D, cooperation, innovation, or both cooperation and innovation as $\theta_i^{FC} = (0, \theta_i^{FC}(rd), \theta_i^{FC}(c), \theta_i^{FC}(d), \theta_i^{FC}(cd))'$.

[6] In their paper, Amoroso (2013)[chapter 2] also shows how, maximising with respect to wages leads to an expression of the wage mark-up as a function of the bargaining parameter, ϕ_{it}, and the ratio between profits and cost of labour $\mu_{it}^W = \frac{\phi_{it}}{1-\phi_{it}} \frac{\Pi_{it}}{W_{it} L_{it}}$.

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 also sunk costs of finding an efficient research partner, or fixed costs of maintaining a research alliance, such as managing the contractual costs (i.e. transaction costs).

Given their level of productivity, capital, materials, and present and past knowledge investment decisions a_{it} and a_{it-1}, firms face 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 $FC(K_{it}, M_{it}, a_{it}) \equiv \tilde{FC}(K_{it}, M_{it}) - \theta_i'^{FC} a_{it-1}$

In order 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, i.e. one featuring imperfect competition only on the output market, and where capital and materials are assumed to be flexible inputs and 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-term profit function under these restrictions is derived in Appendix B and here 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 $\varphi \equiv -\frac{1}{1+\eta} \left(\frac{\eta}{1+\eta} \right)^{\eta}$.

3.2.1 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 a firm's past choices in R&D, innovation, and cooperation.[7]

[7] 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 determine the differences in and the evolution of productivity across firms and over time. Aw et al. (2011) take this a step further and assume that productivity evolves as a Markov process that depends on both investments in R&D and on export market participation.

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 given in (7) not only differs in their fixed costs intercepts, which depend on the set of choices a, 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 or innovation. Therefore, the variable ω_{it} associated with one choice might be different from that of an alternative investment choice.

Figure 1 schematically represents the profile of all the optimal strategies for firm i and the relative pay-offs at specific levels of productivity. 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 this might lead to higher profits than by them doing the R&D by themselves. In particular, cooperating can yield higher profits, as firms can reduce the costs and associated risks of R&D by sharing them. Firms with a level of productivity high enough to bear the sunk costs of introducing an innovation to market will typically invest in a product or process improvement that offers 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.

3.2.2 Value and policy functions

To ascertain information about the sunk costs of R&D, innovating, and cooperating, and to identify the evolution of the productivity states of firms depending on their 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 state points and let the firms' characteristics s_{it} be an element of S. To simplify the framework, but 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 by its past investment actions, a_{it-1}; therefore the state vector is summarized as $s_{it} = (\omega_{it}, \psi_t, a_{it-1})$. To fit the model to the data, we need to add unobserved heterogeneity. In particular, we introduce the vector of pay-off 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, we 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, though they are 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 cooperating 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 $\theta_\Pi \equiv \{\theta_{\Pi i}\}_{i=1,...,N}$ 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 vectorisation 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:

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

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

The problem then 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). 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:

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

where s' and ϵ' denote the next period state and shock. Therefore, when conditioning the value of the state and control variables, the optimal decisions of the firm do not depend on time t, but only on the current and next period state variables. The assumption of the existence of a state variable that is designed to capture the productive and competitive environments faced by the firm at each point might be quite restrictive in the context of technological innovation.

However, as in this paper, when we consider the dynamic optimisation problem of a single agent, the stationary dynamic programming framework can 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 mark-up, 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.

4 The estimation procedure

Estimation is achieved 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 the productivity evolution processes, $f(\psi_{t-1})$, and $\omega(\omega_{it-1},a_{it-1})$, respectively, and the structural parameters needed to construct the profit function as in (6). In the second step, we ascertain 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 a numerical approximation of the solution to the dynamic programming problem at trial parameters.

4.1 Step 1: Static parameters

The production function and the demand parameters are estimated with the method proposed by Amoroso (2013). Within the Cobb-Douglas production function framework, they relax the conventional assumption of perfect competition in the labour market, allowing both firms and workers' union to have some market power.

In their study, **?** 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 labour 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 labour 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, $\mu^W_{it} = 0$). For W_{it} tending to $\bar{W}_{it}$, the wage mark-up decreases, given that the elasticity and the share of labour are constant, which is inversely related to the output mark-up $\frac{\eta}{1+\eta}$.

Next to the labour market rigidities, Amoroso (2013) also corrected for the possible bias in the estimated coefficients from when the deflated gross output is used instead of the gross physical output. Defining the log deflated output as y_{it}, this can be rewritten as

$$y_{it} = q_{it} + (p_{it} - p^j_t),$$

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

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

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

$$y_{it} = \gamma_0 + \gamma_K k_{it} + \gamma_M m_{it} + (1-\mu^W_{it})s_{iLt}l_{it} - \frac{1}{\eta}q^j_t + \tilde{\omega}_{it} + \tilde{u}_{it} \tag{18}$$

where k_{it}, l_{it}, m_{it} are logs of deflated capital, labour, and deflated materials, respectively; q^j_t is the log of the production index in sector j. The composite error term, $\tilde{u}_{it} \equiv u^q_{it} + u^d_{it}$, includes the demand shock, $\tilde{u}^d_{it} \equiv -u^d_{it}/\eta$, and the measurement error, u^q_{it}. $\tilde{\omega}_{it} \equiv \omega_{it}(1+\eta)/\eta$ 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, $q^j_t = \sum_i^{N_j} ms_{it}y_{it}$. Both the intercept, $\gamma_0 \equiv \theta_0(1+\eta)/\eta$, and the factor elasticities of the capital and material, $\gamma_k \equiv \theta_k(1+\eta)/\eta, k = K, M$ are divided by the *output price mark-up* defined as $\equiv \eta/(1+\eta)$ for

$\eta < -1$. The elasticity of labour is defined as in (17).

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

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

Identification of all the structural parameters of the deflated revenue function in (18) is ensured by the presence of firm-specific wages. To estimate all the relevant parameters, Amoroso (2013) adopted a control function approach (Olley and Pakes, 1996) which consists of 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 being 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 input choices. 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, i.e. 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 the current values and one lag in the conditioning set.

The results of the estimation of the deflated revenue function under imperfect competition in both output and labour markets, (18), of the aggregate state transition function, (9), and of the nonlinearly persistent productivity process depending on the technology upgrade, (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) R&D and innovation.

4.2 Step 2: Profit function parameters

It is well known that the parameters of structural dynamic programming problems are not identified (Rust, 1994). Magnac and Thesmar (2002) sshowed that the utility functions of the firms can be identified though if the distribution function of the unobserved preference shocks, the discount rate, and the value function of one of the alternatives (normalisation) are fixed. Hence, it is theoretically possible to identify both the fixed and sunk costs of R&D and innovation. However, in practice the simultaneous identification of such costs requires sufficient variation in the observed R&D investment decisions. To circumvent this problem, we ascertain the fixed cost parameters within the static framework after stimulating the production function parameters. In particular, we consider the estimation of the fixed costs of innovative investments in 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 the fixed costs represent the alternative specific firm-level random coefficients associated with the research investment 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. The results of this estimation are reported in Section 5.

4.3 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). 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 dynamic programming (DP) problem, yielding the advantage of dealing with unobserved heterogeneity. The main idea behind their estimation approach is to avoid the computation of the full solution of the DP problem, i.e. by approximating the expected value function at a state-space point using the average of the 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. For an overview of their methodology and a summary of the algorithm steps, the reader is referred to Appendix A.

5 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 5.1 gives the mean, the standard deviation, and the 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. Labour (L_{it}) refers to the number of employees in each firm for each year,[8] collected in September of that year. The corresponding wages W_{it} include the gross wages plus the salaries and social contributions before taxes. The costs of the intermediate inputs ($Z_{it}M_{it}$) include the costs of energy, intermediate materials, and services. The unit user cost R_{it} (of capital stock K_{it}) is 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 classifi-

[8] 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)

cation.[9] For capital, we use a two-digit NACE deflator of fixed tangible assets calculated by Statistics Netherlands. The share of the cost of labour, material, and capital are denoted as s_{iLt}, s_{iMt}, and s_{iKt}, respectively. The share of the cost of labour 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 observations, after retaining only the respondents to the different waves of the Community Innovation Survey (CIS), is 8306. The CIS data-sets are the main data source for measuring innovation in Europe. The surveys are designed to provide an extensive description of the general structure of innovation 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.[10]

The middle part of Table 5.1 gives 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 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 is 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 5.1 displays the details of the control variable, namely the investment choice k. The furthermost 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 was 444 , whereas in 2008, the same matching exercise yielded a much larger number of firms, i.e. 2413. Our R&D investment and innovation variable is constructed as follows. The firm-specific choice na_{it} takes a value of one if the firm does not engage in any activity other than operating in the market; rd_{it} takes a value of one if the firm decides to spend in R&D; the investment decision c_{it} takes a 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} matches the firm's decision to invest in a technological upgrade; while action cd_{it} marks the decision to both innovate and cooperate. Concerning the type of investment, simple production without innovative or cooperation 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 was expanded so as to cover the whole reference

[9] 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.

[10] Community Innovation Survey, EUROSTAT.

Table 5.1: 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

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 of one for the whole time span.

6 Results

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

6.1 Static parameters

The point estimates of the output price mark-up and all the parameters used to construct the productivity evolution in (11) are reported in Table 6.1. The upper part of the table reports demand elasticity parameters, the aggregate state average, and the productivity level and growth.[11] The elasticity of the demand is found to be -2.8, with a corresponding output price mark-up of 55%. On average, the log productivity is 1.381 and its growth is 1.7%. The aggregate state, ψ_t, is constructed as the 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. We find the aggregate state to be 1.088, on

[11]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 Ackerberg et al. (2007).

Table 6.1: 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
	$\Delta\omega_{it}$	0.046	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)

average. In 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 the 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 labour and output markets seems to be able to reflect 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, $\hat{\sigma}_\epsilon = 0.114$.

With regards to the parameters of the productivity evolution, as in (11), we find evidence of a third order polynomial, and a fair dependence on innovation and cooperation. In particular, the estimated coefficient associated with the action of cooperating is significant at the 5% significance level, being 0.076, while that of innovating is 0.113. The coefficients associated with both cooperating and upgrading technology and the decision to do R&D are 0.062 and -0.011, respectively. In Figure 1, we present a schematic representation of what the profile of all the optimal strategies could look like given the levels of productivity. In Table 6.2, we report the estimated average levels and growth of firm productivity for each of the investment strategy. This shows that innovators (and innovators that cooperate) have the largest productivity growth, and that the firms that undertake collaborative research are the most productive (log of productivity is 0.203). Therefore, in a static scenario, we find that optimal strategy of a firm is to undertake collaborative innovation projects up to a certain level of productivity, ω^*_{it}, above which it is optimal to invest in collaborative R&D. The four means and standard errors of the posterior distributions of the fixed costs are reported in Table 6.3. 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 (EUR 3.0 million and EUR 3.5 million, respectively) are substantially higher than the per-period costs of maintaining an innovation (around EUR 460 000). Moreover, the fixed costs of maintaining an innovative activity, while sharing the costs of R&D, decreases the per-period costs (EUR 290 000). This confirms the rationale behind the cooperating strategies, i.e. the cost sharing motive (Cassiman and Veugelers, 2002; Lopez, 2008; **?**).

R&D cooperation, in fact, allows firms to share costs or to reduce the risks of innovation. The results for the fixed costs are comparable with those found by Aw et al. (2011) for the Taiwanese electronics industry, awho estimated these costs to be on average TWD 67.606 million (roughly EUR 1.8 million) 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, taking into consideration 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 of 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 2 displays the locally weighted scatterplot smoothing (lowess)[12] 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 a 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 for 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 externalities. These results, together with the evidence of the endogenous firm-level productivity, which is positively associated with the action of cooperating, suggest that an innovation policy aiming at encouraging research cooperation might result in a virtuous cycle. Indeed,

[12]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 6.2: Average productivity levels and growth per investment strategy

a	avg $\log \omega_{it}$	st.dvt.	$\Delta\omega_{it}$	st.dvt.
rd	0.092	0.279	0.027	0.228
c	0.203	0.299	0.038	0.213
d	0.137	0.276	0.051	0.186
cd	0.141	0.284	0.048	0.180

Table 6.3: 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	$\omega_{it} \leq 1.172$	$\omega_{it} \leq 1.347$	$\omega_{it} \leq 1.541$
$P_i(a = na\|s, \theta)$	0.408	0.409	0.415	0.414
$P_i(a = rd\|s, \theta)$	0.020	0.024	0.021	0.021
$P_i(a = c\|s, \theta)$	0.011	0.008	0.009	0.011
$P_i(a = d\|s, \theta)$	0.256	0.269	0.263	0.260
$P_i(a = cd\|s, \theta)$	0.304	0.288	0.289	0.293

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 3 plots the Markov Chain Monte Carlo (MCMC) draws of the fixed cost parameters. It appears that the the MCMC draws converge after 50 iterations.

6.2 Dynamic parameters

In this section, we present the results for the DDP model presented in (13). Once the fixed costs are estimated, we 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 ascertain both fixed and sunk costs of doing R&D or innovating with or without a research partner. Figure 3 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 6.4. 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 EUR 4 million for the average firm that undertakes

The values are estimated with the expected signs. Sunk costs are found to be EUR 4 million 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 those of research, but are 3 to 3.5 times higher than the fixed costs of innovating. Moreover, we find additional evidence for 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 (EUR 997 000 and EUR 1.4 million, respectively). The sunk costs parameters cannot be compared with the reported R&D expenditures. This is because the sunk costs may also be related to production factors, such as labour and/or capital that are allocated to research rather than to

Table 6.4: 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	"

production. For this reason, these costs will not appear in the balance sheets of the company (Santos, 2009).

Along with the estimation of the sunk cost parameters, we show the importance of the role played by these costs in shaping the probabilities of undertaking the different research investments. Table 6.4 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. The 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, $\theta_i^{SC}(rd)$ by 25% leads to an increase in the probability of undertaking R&D by 62.2%.

7 Conclusion and future work

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 of technologically upgrading. 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 decisions. 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, but 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 innovation activities, rather than for the just R&D investments.

We also find that the sunk costs of innovation are 1.5 to 3 times smaller than the sunk costs of R&D. Also, the sunk costs are found to be roughly 1.5 times larger than the fixed costs of research (both cooperative and done alone), 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 of 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 aimed at encouraging research cooperation. We show how these types 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 of engaging in both

R&D and innovation when these activities are shared with a research partner. Therefore, in elaborating their policies for innovation, governments should ensure to create frameworks that encourage the collaboration throughout the innovation process.

Future work could include differentiating between the types of innovations, as well as the types of cooperation. In particular, under future investigation, we would like to differentiate the sunk costs of combinations of types of innovation (product/process) with different types of cooperation partner. For example, we could include the sunk costs of product and process innovations and take the average expected benefits and costs of R&D and innovation for different cooperating strategies (see Peters et al. (2013) for the expected benefits of R&D), or the expected benefits and costs of cooperating for product versus process innovations.

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., Benkard, L., Berry, S., and Pakes, A. (2007). Econometric Tools for Analyzing Market Outcomes. *Handbook of Econometrics*, 6.

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. (2013). *Heterogeneity of Innovative, Collaborative, and Productive Firm-level Processes*. CentER dissertation. CentER, Tilburg University.

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.

Brockhoff, K. (1992). R&d cooperation between firms - a perceived transaction cost perspective. *Management Science*, 38(4):514–524.

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*.

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.

Doraszelski, U. and Jaumandreu, J. (2013). R&d and productivity: Estimating endogenous productivity*. *The Review of Economic Studies*.

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.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):pp. 1127–1150.

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.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, pages 649–670.

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.

Oerlemans, L. A. and Meeus, M. T. (2001). R&d cooperation in a transaction cost perspective. *Review of Industrial Organization*, 18(1):77–90.

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.

Peters, B., Roberts, M. J., Vuong, V. A., and Fryges, H. (2013). Estimating dynamic R&D demand: An analysis of costs and long-run benefits. Technical report.

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.

Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology*, 87(3):pp. 548–577.

Williamson, O. E. (1985). *The economic intstitutions of capitalism*. Simon and Schuster.

Williamson, O. E. (1989). Chapter 3 transaction cost economics. volume 1 of *Handbook of Industrial Organization*, pages 135 – 182. Elsevier.

A Imai et al. (2009) estimation algorithm

The direct way of obtaining continuation values as the one in (14) 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.[13] 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

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

(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. 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).

The Imai et al. (2009) 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.[14] 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

[14] 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."

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*

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

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

B 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 labour 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 (2013) definition of the wage mark-up $\mu_{it}^W \equiv \frac{W_{it}-\bar{W}_{it}}{W_{it}}$, and taking into account the demand as in (1), we can rewrite the cost of labour 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 labour 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 labour market, is derived from the following optimization problem for firm i:

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

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

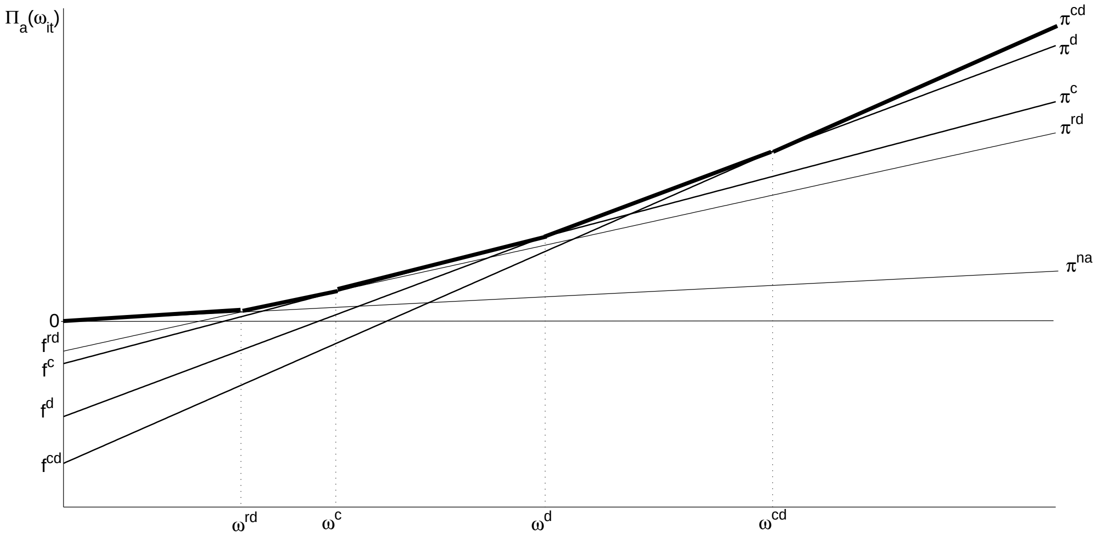
$$\Pi_{it} = -\frac{1}{\eta} P_{it} Q_{it}.$$

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

Figure 1: R&D, Cooperation and Innovation Choices



C Tables and Figures

Figure 2: Investment policy functions

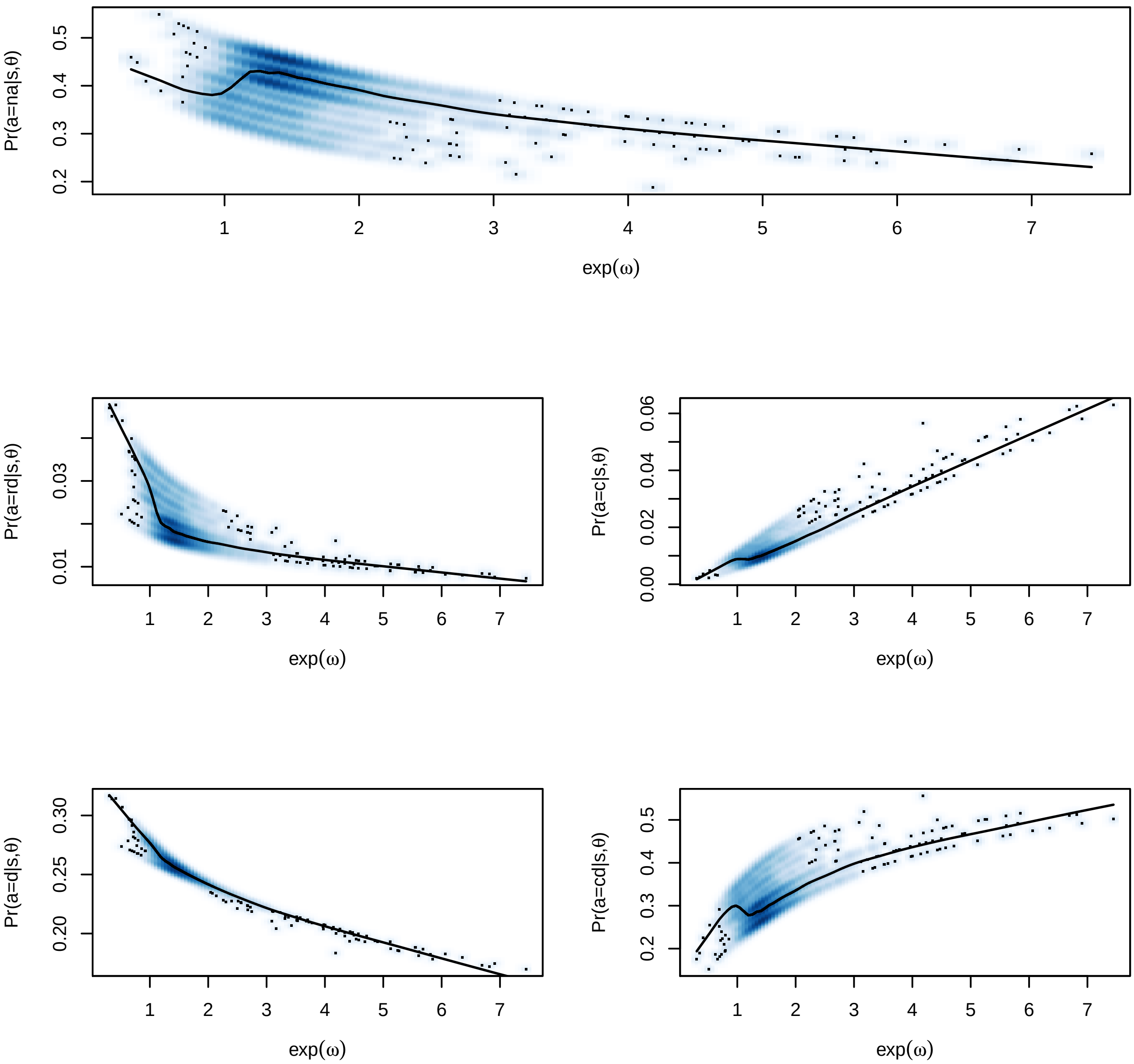
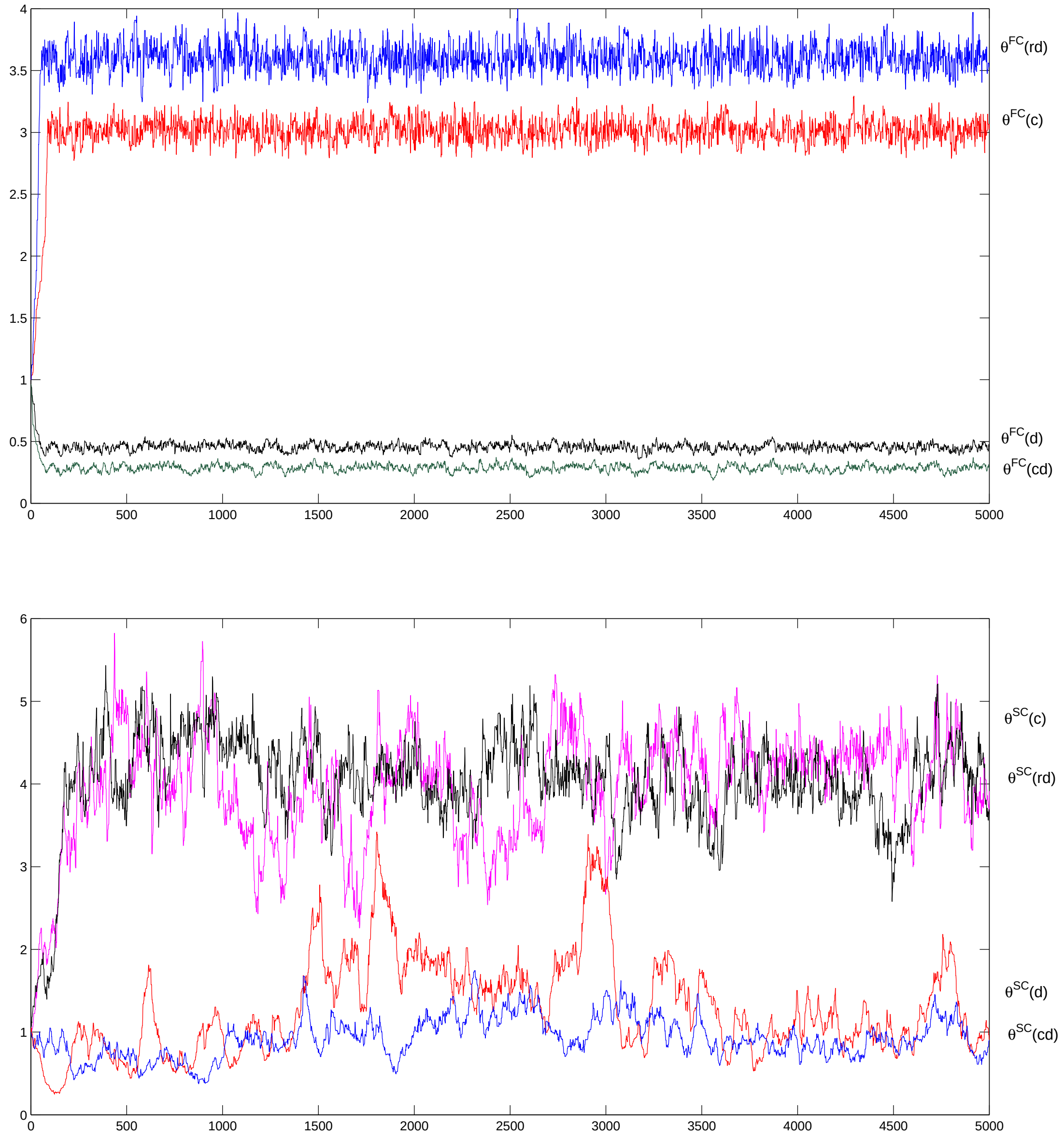


Figure 3: MCMC iterations of fixed and sunk cost parameters



Note: MCMC plots of θ^{FC} and θ^{SC}

European Commission
Joint Research Centre – Institute for Prospective Technological Studies

Title: **The hidden costs of R&D collaboration**

Author: Sara Amoroso

Spain: European Commission, Joint Research Centre

2014 – 34 pp. – 21.0 x 29.7 cm

Scientific and Technical Research series – ISSN 1831-9408 (online)

Abstract

Many European policy initiatives continue to promote R&D collaboration in view of its expected benefits. Despite the advantages of R&D cooperation, to benefit from it, firms must create a structure to support the efficient transfer of knowledge-based assets. In fact, the setup and administration of common resources might be costly. This paper derives the distribution of the costs associated with R&D collaboration, as they could shape firms' R&D- related investments. To ascertain these costs, we model the expected benefits from R&D cooperation with a structural dynamic monopoly model. The modelling results show that the sunk costs of innovation are lower when collaborating with a research partner, and that a firm's probability of investing in R&D or innovation increases with the level of productivity, only when collaborating in R&D and innovation. We also find that the sunk costs of innovation are 1.5 to 3 times lower than the sunk costs of R&D. Additionally, it can be seen that the suggested structural framework of a firm's heterogeneity in cost functions used in our model can offer a straightforward extension to existing policy impact evaluation.

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

Serving society
Stimulating innovation
Supporting legislation