UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 02-62 (18)
Statistics and Econometrics Series
December 2002

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

PSEUDO-MAXIMUM LIKELIHOOD ESTIMATION OF A DYNAMIC STRUCTURAL INVESTMENT MODEL

Rocío Sánchez-Mangas*

Abstract

This paper belongs to the recent investment literature focused on the modelling of microeconomic investment decisions. The increasing concern about this topic is related to the growing availability of microeconomic datasets which show the investment behavior taking place at the firm level. This behavior is far from the smooth capital adjustment pattern derived from the traditional investment models. Rather it is characterized by infrequent and lumpy adjustment. New investment models must be considered to capture this behavior. In this paper we formulate a dynamic structural investment model with irreversibility and nonconvex adjustment costs and try to stress the importance of these costs in the firms' investment decisions. From the methodological point of view, we set the investment decision on the dynamic programming framework. More specifically, we consider a discrete choice dynamic programming problem in which firms decide to invest or not to invest. The estimation strategy we adopt is the Nested Pseudo-Likelihood (NPL) algorithm recently proposed by Aguirregabiria and Mira (2002). It is an estimation method which has clear advantages over previous techniques proposed in this context. Up to our knowledge, this paper constitutes the first empirical application of this estimation method.

Keywords: Capital adjustment costs, Irreversible investment, Structural estimation, Discrete choice models, Dynamic programming, Nested algorithms

*Sánchez-Mangas, Dept. of Statistics and Econometrics, Universidad Carlos III de Madrid, C/Madrid, 126, 28903 Getafe (Madrid). Spain. E-mail: rsman@est-econ.uc3m.es. Tel: +34 916249826. I am very grateful to César Alonso, Victor Aguirregabiria, Pedro Mira, Alfonso R. Sánchez and seminar participants at ESEM 2002, Universidad Carlos III de Madrid, Universitat Autonoma de Barcelona, Universidad de Navarra, Universidad de Vigo and Universitat de les Illes Balears for very helpful comments. Financial support from Spanish DGI, grant BEC2000-0170, is acknowledged. All remaining errors are my own.

1 Introduction

In the last years investment literature has shown an increasing concern about the modelling of microeconomic investment decisions. This increasing concern is related to the growing availability of microeconomic datasets which show the investment behavior taking place at the firm level. The analysis of this behavior highlights the importance of some phenomena that are masked when analyzing aggregate investment data. Traditional investment models, characterized by strictly convex capital adjustment costs functions, seemed to be appropiated to capture the investment behaviour observed in the aggregate data. This behavior is associated with a pattern of smooth capital adjustment. However, in the analysis of firm-level data, we find that there are periods in which some firms decide not to invest, and periods in which the investment carried out involves a high percentage of the installed firm capital stock. These phenomena of infrequent and lumpy adjustment are far away from the smooth adjustment pattern that characterized the investment literature until the last decade. This fact has given rise to a new generation of investment models that moves away from the convexity structure. Investment models taking into account irreversibilities and adjustment costs structures with nonconvex components have began to be considered.

In this paper we propose and estimate a dynamic structural model of fixed capital investment at the firm level. Our dataset consists of an unbalanced panel of Spanish manufacturing firms. The presence of infrequent and lumpy investment seems to be important in these data. Based on this empirical fact we consider a dynamic model of irreversible investment with a general specification of adjustment costs including convex and nonconvex components. We try to get insight about the investment behaviour taking place at the firm level and the importance of different types of adjustment costs in the frims' investment decisions.

From the methodological point of view, we set the firm's investment decision problem in the dynamic programming framework. More specifically, we formulate a discrete choice dynamic programming problem in which firms decide each period between not to invest or to undertake an investment project. Until the last years, the usual approach to handle with the estimation of these models was some kind of solution-estimation algorithm in the spirit of the Nested Fixed Point (NFXP) algorithm (Rust, 1987). This estimation method consists of a nested algorithm in which the dynamic programming problem must be solved in each iteration in the search for the parameter estimates. The number of empirical works using this method is limited, because the high computational cost associated with it obligues to consider very parsimonious specifications for the objective function. In the last decade some alternative methods have appeared. Hotz and Miller (1993) proposed an estimation method, the Conditional Choice Probability (CCP) estimator that does not require to solve the dynamic programming problem to obtain estimates of the structural parameters. From the computational point of view, this method has clear advantages over the solution-estimation tecniques. However, these computational gains are obtained at the expense of efficiency. Thus, there is a clear trade-off between these two aspects. In Sánchez-Mangas (2002) we apply this method to the estimation of a dynamic structural model of irreversible investment for Spanish firms. Recently, Aguirregabiria and Mira (2002) have proposed an estimation method, the Nested Pseudo-Likelihood (NPL) algorithm, that bridges the gap between the two estimation strategies mentioned above. It does not require the solution of the dynamic programming problem. Rather, it is based on a representation of the solution of that problem in the space of conditional choice probabilities. Successive iterations in the algorithm return a sequence of estimators that includes as extreme cases the Hotz and Miller's CCP estimator and Rust's NFXP estimators. Furthermore, all the estimators in the sequence are distributed asymptotically like the maximum likelihood estimator.

In this paper, we apply this strategy to the estimation of our dynamic discrete choice model of irreversible investment. Up to our knowledge, this is the first exercise of application of this new and, in our opinion, promising estimation strategy.

The rest of the paper is organized as follows. In Section 2 we describe the dataset used in this study. Section 3 formulates a dynamic structural model of irreversible investment with nonconvex adjustment costs. In Section 4 we describe the estimation strategy we adopt. Section 5 reports the estimation results and Section 6 concludes.

2 Evidence from the data

This section stresses the stylized facts which are present in Spanish manufacturing firms. This is a revised and extended version of the analysis carried out in Sánchez-Mangas (2002). The dataset we use has been taken from the Encuesta sobre Estrategias Empresariales (ESEE) conducted by the Spanish Ministry of Industry and Energy. It contains annual information of the balance sheet and other economic variables. Our sample is an unbalanced panel of 1592 firms between 1990 and 1997. We concentrate on capital stock and gross expenditure on capital goods. The investment rate for period t has been constructed as the ratio between gross expenditure in that period and the capital stock at the beginning of the period.

We analyze the investment rate for the whole dataset and for firms of different size. The recent empirical studies carried out describing firm investment behavior in different countries have highlighted the importance of two fundamental features: firms do not adjust their capital stock smoothly. Rather a relevant percentage of firms decide not to invest during some periods, and, when they decide to invest, the amount of investment represents a high percentage of installed capital. This evidence can be found in Barnett and Sakellaris (1995), Doms and Dunne (1998) or Nielsen and Schiantarelli (1998), among others.

The presence of these two phenomena, inaction and investment spikes, has important effects from the point of view of the modelization of the investment decision. In the following figures and tables, we explore the importance of these phenomena in our dataset.

Figure 1 depicts a histogram of annual firm-level gross investment rates. The distribution is strongly skewed to the right. Around 30% of the observations have investment rates that are zero or close to zero (less than 0.033 gross investment rate), which reflects the fact that many firm-year observations involve little or no investment. The long right tail ilustrates the fact that a fraction of plants experiment a large investment episode in any given year. The last bar accounts for the observations having an investment rate greater than 0.98.

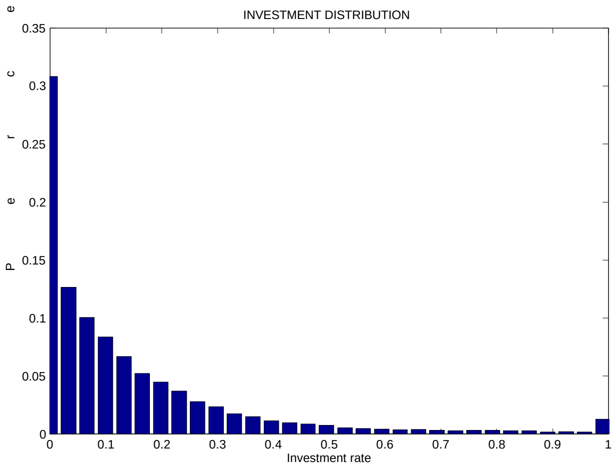

Figure 1

In the following table we can see the evidence of inaction and lumpiness by year. The first column shows the percentage of observations that experiment zero investment in a given year, while the second one shows the percentage that experiment an investment rate greater than 20%, where i stands for investment rate.

	Inaction	Lumpiness
Year	(% obs. with $i = 0$)	(% obs. with $i > 0.2$)
1991	17.46	30.81
1992	18.61	26.59
1993	23.46	18.52
1994	20.28	21.54
1995	18.39	26.65
1996	16.52	23.75
1997	16.24	26.19
Total	18.05	24.70

Table 1: Evidence of inaction and lumpiness.

Even in the year of lowest percentage of observations with zero investment, this percentage is quite high,above 16%. On the other hand, investment rates higher than 20% arise in more than 20% of the observations in almost every year, reaching 30% in one of them. Furhtermore, we can observe in this table a cyclical behavior. Inaction is a countercyclical phenomenon, since the highest percentage of observation with zero investment occurred in 1993, year in which the GDP and the gross formation of fixed capital underwent, respectively, a decrease of 0.68% and 11.72% with respect to 1992. On the contrary, investment spikes are a cyclical phenomena.

We also report this evidence of infrequent and lumpy capital adjustment distinguishing three categories of firms: small, medium and large firms. We follow the classification criterion established by the European Commission. According to this criterion, small firms are those with no more than 50 employees and no more than 7 million euro of annual turnover. Medium firms are those with more than 50 and no more than 250 employees and an annual turnover greater than 7 million euro and lesser than 40 million euro. Large firms are those with more than 250 employees and an anual turnover greater than 40 million euro. The following table shows the distribution of firms in the sample in the first column, the percentage of observations with zero investment (inaction) in the second column and the percentage of observations with an investment rate greater than 20% of the installed capital (lumpiness) in the third colum.

Type of firm	% obs.	Inaction (% obs. with $i = 0$)	Lumpiness (% obs. with $i > 0.2$)
Small firms	57.22	29.24	24.52
Medium firms	25.67	4.51	23.78
Large firms	17.12	0.93	26.65
Total	100	18.05	24.70

Table 2: Evidence of inaction and lumpiness by categories of firms

As we can see, more than half of the firms in our dataset are small firms. The percentage of observations with zero investment if very different for small, medium and large firms. While there are around 30% of observations accounting for zero investment in the group of small firms, this percentage is only 4% for medium firms and almost insignificant for large firms. However, in the three categories considered, the percentage of observations with investment rates greater than 20% of installed capital is quite similar, around 23%.

Figure 2 mimics Figure 2a in Doms and Dunne (1998). For each firm in our dataset we have ranked its annual investment rate in descending order. The figure shows the mean and the median investment rate in each rank.

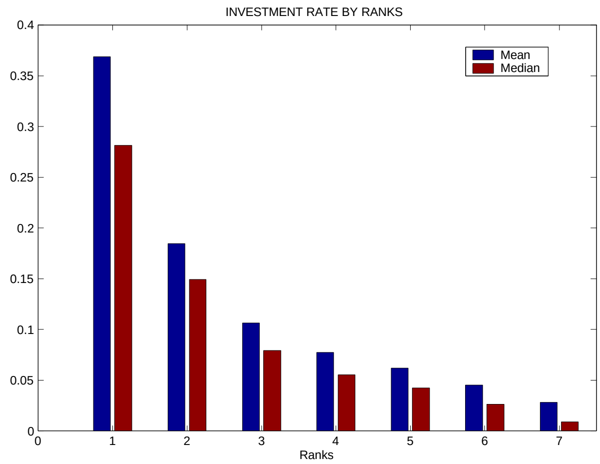

Figure 2

As we can see, the first bar, corresponding to the highest mean investment rate, exceeds 35%, while for the second rank is below 20%, and for subsequent ranks is even less than 10%. That is, the means drop off significantly after ranks 1 and 2, meaning that many firms experiment one or two periods of intense investment, while the rest of the periods are characterized by moderate investment. The median is always below the mean, reflecting the skewness to the right of the investment rate distribution.

In Figure 3, the ranks of investment have been constructed as in Figure 2. We present, for each rank, the percentage that the average investment in this rank represent over the investment carried out in the whole period. It can be seen that, on average, more than the third part of investment in the whole period has been carried out in only one year and almost 60% in two years. Table 3 shows the percentages by ranks

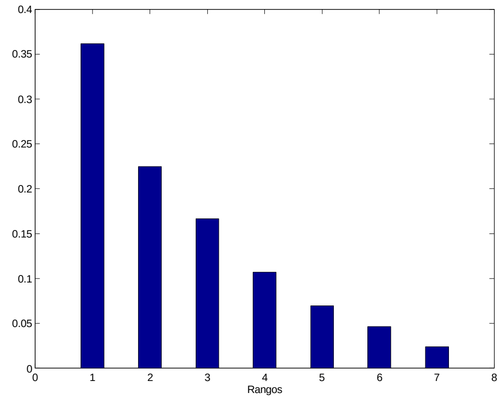

Figure 3

Rank	% investment over total investment	% acumulated
1	0.3617	0.3617
2	0.2247	0.5864
3	0.1667	0.7531
4	0.1068	0.8599
5	0.0695	0.9295
6	0.0466	0.9761
7	0.0238	1.000

Table 3: Percentaje of average investment in each rank over investment in the whole period

Figure 4 gives an insight about the importance of large investment episodes on the time series fluctuations of investment.

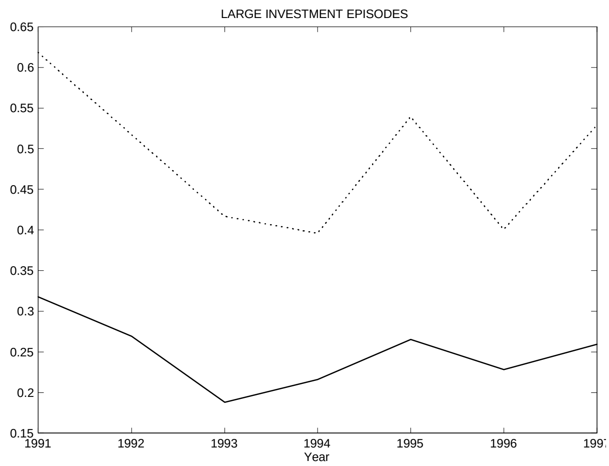

Figure 4

The solid line reflects the percentage of observations with an investment rate greater than 20%. The dotted line represents the percentage of investment accounted by observations having these large investment episodes. Observations with large investment episodes constitute around 25% of the total, but account for approximately 50% of gross investment. That is, around half of the total gross investment is related with lumpiness and half of it with smooth adjustments. Similar evidence has been reported in Cooper, Haltiwanger and Power (1999) for a large set of US manufacturing firms.

This descriptive analysis of investment behavior in Spanish manufacturing firms highlights the importance of inaction and investment spikes. These phenomena are far away from the pattern of smooth capital adjustment derived from the investment models proposed in the literature until recent years. Furthermore, these empirical findings clearly support the convenience of focusing on investment models able to capture this behavior. Thus, we will formulate an investment model which accounts for irreversibilities and nonconvex capital adjustment costs.

3 A dynamic structural model of fixed capital investment

3.1 Framework and basic assumptions

This section builds heavily on Section 3 in Sánchez-Mangas (2002). Consider a risk neutral firm that produces an homogeneous good using as inputs labor and capital equipment with some firm-specific characteristics. At each period the firm decides hirings and dismissals of workers and purchases of new capital in order to maximize the expected discounted stream of current and future profits over an infinite time horizon. The firm operates in competitive product and input markets, and its profit at period t, in output units is given by:

$$\Pi_t = Y_t(K_t, L_t, a_t) - w_t L_t - p_t K_t i_t - AC(K_t, i_t, p_t) \tag{1}$$

where Y_t is real output, K_t is the capital stock installed at the beginning of period t, L_t represents labor in physical units, i_t represents the investment rate, defined as $\frac{I_t}{K_t}$, where $I_t = K_t i_t$ represents new capital purchases, w_t and p_t are input prices relative to product price. We assume that labor can be adjusted costlessly, so the decision on employment is static. However, when the firm decides to adjust its capital stock it faces some adjustment costs represented by the function $AC(K_t, i_t, p_t)$. A well-known evidence that arises in any empirical study of firms' behavior is the large amount of heterogeneity in firms size, productivity and behaviour in general, even after controlling for location, industry or product characteristics. For this reason, we state the problem in terms of the investment rate i_t, instead of investment in physical units, I_t.

Output depends on labor and installed capital at the beginning of the period and a productivity shock a_t, according to the Cobb-Douglas production function:

$$Y_t = a_t K_t^{\alpha_K} L_t^{\alpha_L} \tag{2}$$

where $\alpha_K, \alpha_L \in (0, 1]$. We assume there is one period time-to-build, i.e, the new equipment is productive one period after its acquisition. The productivity shock is exogenous and follows a first order Markov process with transition density $\phi_a(a_{t+1} | a_t)$.

We assume that adjustment costs faced by the firm when it decides to invest can be variable or fixed costs:

$$AC(K_t, i_t, p_t) = VC(K_t, i_t, p_t) + FC(K_t) \tag{3}$$

Variable costs $VC(\cdot)$ include costs associated with the installation of the capital stock. We assume a convex structure for these costs, similar to the specification of adjustment costs in the traditional investment models. More specifically, we use the following quadratic function:

$$VC_t = VC(K_t, i_t, p_t) = \frac{\theta_Q}{2} p_t K_t i_t^2 \tag{4}$$

where θ_Q is a constant parameter.

Fixed adjustment costs $FC(\cdot)$ are internal costs related to the reorganization of the productive process and retraining of employees in the handling of the new equipment. We assume that these costs are proportional to the installed capital stock:

$$FC = FC(K_t) = 1(i_t > 0)\ \theta_F K_t \tag{5}$$

where $1(.)$ is the indicator function and θ_F is a constant parameter.

Since the firm operates in competitive markets, input prices are exogenous to the firm. We assume that capital price and wages follow a Markov process with transitional densities $\phi_p(p_{t+1}|p_t)$ and $\phi_w(w_{t+1}|w_t)$, respectively. Capital retirement and physical depreciation are exogenously given to the firm. The capital stock follows a transition rule given by

$$K_{t+1} = K_t \left((1-\delta_t) + i_t\right) \tag{6}$$

where $\delta_t \in (0,1)$ is the depreciation rate, which includes not only the economic depreciation of the capital stock but also the capital retirements due to obsolescense.

At the beginning of period t, the firm knows its level of capital stock and labor, the input prices in the industry where it operates and the value of productivity and cost shocks. Since the decision on labor is static, an optimal condition for labor can be obtained and the one-period profit function can be written as:

$$\Pi_t = Y_t(K_t, L_t^*, a_t) - w_t L_t^* - p_t K_t i_t - AC(K_t, i_t, p_t), \tag{7}$$

where L_t have been optimally chosen. The optimal condition for labor, under the assumption of a Cobb-Douglas production function with constant returns to scale, is given by

$$L_t^* = \left(\frac{a_t(1-\alpha_K)}{w_t}\right)^{\frac{1}{\alpha_K}} K_t \tag{8}$$

Thus, the profit function in terms of capital stock can be written as:[1]

$$\Pi_t = R_t K_t - p_t K_t i_t - AC(K_t, i_t, p_t) \tag{9}$$

where R_t is a profitability shock in terms of the productivity shock entering the production function, wages and technological parameters according to the following expression:

$$R_t = R(a_t, w_t, \alpha_K) = \left(\frac{a_t(1-\alpha_K)}{w_t^{1-\alpha_K}}\right)^{1/\alpha_K} \frac{\alpha_K}{1-\alpha_K} \tag{10}$$

[1]Note that under Cobb-Douglas production with constant returns to scale, and our specification of capital adjustment costs, the one-period profit function is linear in the capital stock.

We assume that the investment decision is completely irreversible, i.e, the firm decides purchases of capital stock and once a new equipment has been acquired, it cannot be sold.[2] Thus, the firm faces the decision of not to invest or to undertake a strictly positive investment, so the decision variable in this problem is $i_t \geq 0$. Let s_t be the vector of state variables, which can be observed by the firm and the econometrician, or only by the firm. The firm's decision problem can be written as:

$$\max_{\{i_t \geq 0\}} \sum_{t=0}^{\infty} \beta^t E\left[\Pi(i_t, s_t)\right] \tag{11}$$

where $\beta \in (0, 1)$ is the discount factor, related to the interest rate of the economy. The Bellman's equation for this problem is given by:

$$V(s_t) = \max_{\{i_t \geq 0\}} \Pi(i_t, s_t) + \beta EV(s_{t+1}|\, s_t, i_t) \tag{12}$$

where $EV(s_{t+1}|\, s_t, i_t)$ is the expected conditional value function

$$EV(s_{t+1}|\, s_t, i_t) = \int V(s_{t+1})\phi(ds_{t+1}|\, s_t, i_t) \tag{13}$$

and $\phi(ds_{t+1}|\, s_t, i_t)$ is the transition probability of the state variables.

3.2 Optimal decision rule

Firms in our model face a double decision: the discrete choice of not investing vs. investing, and, if they decide to invest, the continuous decision about the amount of investment. If firms decide not to invest, it can be due to two different reasons: on one hand, the imposibility of selling the purchased capital goods, i.e., the total irreversibility of the investment decision; on the other hand, the possible existence of high fixed adjustment costs, that penalize the small capital adjustments and can lead the firms to decide postponing the investment decision. So the decision variable in the intertemporal firm's problem, the investment rate, is censored at zero.

If the only source of censoring were the irreversibility of the investment decision, the value function would be continuous and concave, and thus, relatively easy to deal

[2] Alternatively, there are some papers assuming the existence of second-hand markets in which the selling price for capital is lower than the purchase price, so the decision on capital is partially irreversible.

with. However, the introduction of fixed adjustment costs brings a discountinuity in the one-period profit function, that makes the value function to be nonconcave. The decision rules for these kind of problems have been characterized by Bertsekas (1976), using properties of K-concave functions. We can find examples of these type of decision rules in Scarf (1959), Slade (1998) or Aguirregabiria (1999) in the context of inventories and price adjustment models.

The optimal decision rule for our investment decision problem is given by:

$$i(s_t, \theta) = \begin{cases} i^*(s_t, \theta) & \text{if} \quad i^*(s_t, \theta) > 0 \text{ and } \gamma(s_t, \theta) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where $i^*(s_t, \theta)$ is the optimal interior solution characterized by

$$\tilde{\pi}_i(s, i^*(s, \theta), \theta) + \beta EV_i(s, i^*(s, \theta), \theta) = 0, \tag{15}$$

with $\tilde{\pi}_i \equiv \partial\tilde{\pi}/\partial i$ and $EV_i = \partial EV/\partial i$ and the function $\gamma(s_t, \theta)$ is given by

$$\tilde{\pi}(s_{rt}, i^*(s_t, \theta), \theta) - FC(s_t, \theta) - \tilde{\pi}(s_t, 0, \theta) + \beta \left[EV(s_t, i^*(s_t, \theta), \theta) - EV(s_t, 0, \theta) \right]. \tag{16}$$

That is, there is a first order condition of optimality for the interior solution, given by (15), and there are two conditions for the discrete choice between interior and corner solution. The first one, $i^*(s_t, \theta) > 0$, is related to the non-negativity constraint, i.e. the irreversibility of the decision: the interior solution will be optimal only if it is positive. If condition (15) holds for a negative value $i(s_t, \theta) < 0$, the firm will choose $i(s_t, \theta) = 0$, due to total irreversibility. The second condition, $\gamma(s_t, \theta) > 0$, is related to the existence of fixed adjustment costs. If $\gamma(s_t, \theta) > 0$, it means that the fixed costs are not high enough to lead the firm to decide not to invest.

Our model is a dynamic choice model in which the decision variable is censored at zero as a consequence of inaction. As it is explained above, there are two sources of censoring, irreversibility and fixed adjustment costs, which are indistinguishable for the econometrician. When the intertemporal profit, gross of fixed adjustment costs, is maximized for a negative value of investment, the optimal decision is inaction due to irreversibility. When it is maximized for a positive level of investment, but the value obtained with this level is lower than the value obtained with zero investment, the optimal decision is inaction due to the presence of fixed adjustment costs.

Although the optimal decision rule (14) involves marginal conditions of optimality and optimal discrete choices, in this paper we obtain estimates of the structural parameters which only exploit conditions associated to the optimal discrete choice between interior and corner solution.[3]

4 Estimation method

We have a panel of firms with information on output, capital, labor, investment and input prices.

$$\{Y_{nt}, K_{nt}, I_{nt}, p_{nt}, w_{nt}; \quad n = 1, ..., N; \quad t = 1, ..., T_n\}$$

We are interested in exploiting this sample to estimate the structural parameters. According to the estimation procedure that we describe here, we can classify the structural parameters in four groups: a) the parameters entering the production function; b) the parameters that describe the transition probabilities of input prices and profitability shock; c) the adjustment costs parameters: θ_Q and θ_F; and d) the parameters of the distribution of the state variables which are unobservable for the econometrician.

For estimation purposes, we proceed in two stages. In a first stage we estimate the parameters of the production function and the transition probabilities of the state variables. Once we have estimates of the parameters entering the production function, we can obtain estimates of the productivity shock a_{nt}, and construct the profitability shocks R_{nt} as in (10). In a second stage we estimate the rest of the structural parameters. In order to do this, we exploit the optimal discrete choice "to invest vs. not to invest" to obtain estimates of the adjustment costs parameters θ_Q and θ_F and the parameters in the distribution of the unobservable state variables.

With respect to the estimation of the production function, we summarize below the main features of the modelization and estimation method used in Alonso-Borrego

[3] Since corner solutions are very frequent in our dataset, the subsample of observations that we can use to exploit moment conditions associated to marginal conditions of optimality (i.e, Euler equations) is relatively small. Besides, parameters associated with fixed costs can only be identified by exploiting the discrete decision between interior and corner solution.

and Sánchez-Mangas (2001), where we estimated a production function using the dataset we use in this paper.[4] We begin by considering a Cobb-Couglas production function without imposing constant returns to scale:

$$y_{nt} = \alpha_K \; k_{nt} + \alpha_L \; l_{nt} + u_{nt} \tag{17}$$

where $y_{nt} = \ln(Y_{nt})$, $k_{nt} = \ln(K_{nt})$, $l_{nt} = \ln(L_{nt})$ and $u_{nt} = \ln(a_{nt})$. We allow the following structure for the productivity shock:

$$\begin{aligned} u_{nt} &= A_t + \eta_n + v_{nt} \\ v_{nt} &= \rho \; v_{n,t-1} + \xi_{nt} \end{aligned} \tag{18}$$

where A_t is an aggregate effect, η_n is a time invariant firm-specific effect, v_{nt} is an $AR(1)$ idiosyncratic shock and ξ_{nt} is $iid \; N(0, \sigma_\xi^2)$.

In order to estimate the parameters $(\alpha_K, \alpha_L, \rho)$, we formulate the dynamic representation of (17):

$$y_{nt} = \alpha_K \; k_{nt} - \alpha_K \; \rho \; k_{n,t-1} + \alpha_L \; l_{nt} - \alpha_L \; \rho \; l_{n,t-1} + \rho \; y_{i,t-1} + (A_t - \rho \; A_{t-1}) + (1-\rho) \; \eta_n + \xi_{nt}$$

or

$$y_{nt} = \pi_1 \; k_{nt} + \pi_2 \; k_{n,t-1} + \pi_3 \; l_{nt} + \pi_4 \; l_{n,t-1} + \pi_5 \; y_{i,t-1} + A_t^* + \eta_n^* + \xi_{nt}$$

subject to two non-linear restrictions: $\pi_2 = -\pi_1 \pi_5$ and $\pi_4 = -\pi_2 \pi_5$, and where $A_t^* = A_t - \rho A_{t-1}$ and $\eta_n^* = (1-\rho)\eta_n$.

Given consistent estimates of the unrestricted parameter vector $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5)'$ and its variance-covariance matrix, the restrictions can be tested and imposed by minimum distance to obtain estimates for the restricted parameter vector $(\alpha_K, \alpha_L, \rho)'$.

In the estimation of the unrestricted parameter vector, we apply the "extended GMM" estimation method proposed by Arellano and Bover (1995). It is based on a system including not only differenced equations with lagged levels as instruments, but also level equations with lagged differences as instruments. In a context of highly persistent variables, such as sales, capital or employment, the application

[4] See more details on the estimation of the production function in Alonso-Borrego and Sánchez-Mangas (2001).

of standard GMM estimators which take first differences to eliminate unobserved firm-specific effects and use as instruments lagged levels has produced unsatisfactory results (Mairesse and Hall, 1996). More specifically, it yields a low and statistical insignificant capital coefficient and suggest decreasing returns to scale. These problems, due to the weakness of the instruments considered, are dramatically reduced when applying the extended GMM estimation method.

Since our specification of the profit function as a linear function of the capital stock is based on the constant returns to scale hypothesis, we tested the validity of this hypothesis and obtained estimates imposing constant returns to scale. The estimation results are shown in the Appendix.

Estimation of the adjustment costs parameters

Once we have estimated the tecnological parameters, the productivity shock a_{nt} can be recovered and we can obtain the profitability shock R_{nt} according to (10). This profitability shock will be treated as an observable state variable in the estimation of the adjustment costs parameters.

Let s_{nt} the vector of state variables, which can be decomposed as $(x_{nt}, \varepsilon_{nt})$, where x_{nt} stands for state variables observed by the firm and the econometrician and ε_{nt} stands for state variables which are unobservable for the econometrician. In the firm's decision problem, once the production function has been estimated, the vector of observable state variables is given by $x_{nt} = (p_{nt}, K_{nt}, R_{nt})'$.

Let $d = \{0, 1\}$ be the index for the optimal discrete choice, where $d = 0$ means that the optimal decision for the firm n at period t is not to invest, i.e, $i(s_{nt}) = 0$, and $d = 1$ means that the optimal decision is to undertake an investment project, i.e, $i(s_{nt}) > 0$.

Under the *Additive Separability (AS)* assumption *(*Rust, 1987), the vector of unobservable state variables is given by $\varepsilon_{nt} = (\varepsilon^0_{nt}, \varepsilon^1_{nt})$, where ε^0_{nt} is associated with the decision $d = 0$ and ε^1_{nt} with the decision $d = 1$, and these unobservable state variables enter the one-period profit function in an additive fashion. The additive separability assumption allows us to write:

$$\pi^d(s_{nt},\theta) = \pi^d(x_{nt},\theta) + \varepsilon^d_{nt} \qquad \text{for } d = 0,1 \tag{19}$$

where

$$\begin{aligned}\pi^1(s_{nt},\theta) &= R_{nt}K_{nt} - p_{nt}K_{nt}\ i_{nt} - \frac{\theta_Q}{2}p_{nt}K_{nt}i^2_{nt} - \ \theta_F K_{nt} + \varepsilon^1_{nt} \\ \pi^0(s_{nt},\theta) &= R_{nt}K_{nt} + \varepsilon^0_{nt}\end{aligned}$$

The unobservable state variables represent the uncertainty of the researcher about the actual expected profit that is observable to the firm. We assume they ε^d_{nt}, for $d = \{0,1\}$ are independent and identically distributed with zero mean and variance σ^2_ε.

Let us consider the following multiplicative decomposition of the expected current profits:

$$E\left[\pi^d(x_{nt},\theta)\right] = \pi^d(x_{nt})'\ \mu(\theta) \qquad \text{for } d = 0,1 \tag{20}$$

Since the adjustment costs parameters enter this function linearly, the decomposition (20) is given by:

$$\pi^0(x_{nt}) = \begin{pmatrix} R_{nt}K_{nt} \\ 0 \\ 0 \end{pmatrix} \qquad \pi^1(x_{nt}) = \begin{pmatrix} R_{nt}K_{nt} - p_{nt}K_{nt}E\left[i_{nt}|\, x_{nt}, d_{nt} = 1\right] \\ -\frac{1}{2}p_{nt}K_{nt}E\left[(i_{nt})^2\right| x_{nt}, d_{nt} = 1\right] \\ -K_{nt} \end{pmatrix}$$

$$\mu(\theta) = \begin{pmatrix} 1 \\ \theta_Q \\ \theta_F \end{pmatrix}$$

The first component of $\pi^1(x_t)$ is related to the revenues realized by the firm net of the acquistion price of the new capital stock. The second and third components are related, respectively, to the quadratic and fixed adjustment costs.

Let us consider the *Conditional Independence (CI)* assumption (Rust, 1987), which establishes that the conditional transition probability of the state variables can be factorized as:

$$pdf(x_{t+1}, \varepsilon_{t+1}|\, x_t, \varepsilon_t, d_t) = pdf(\varepsilon_{t+1}|\, x_{t+1})\ pdf(x_{t+1}|\, x_t, d_t) \tag{21}$$

This assumption implies, on one hand, that conditional on the discrete choice and the current value of the observable state variables, the future observable state variables do not depend on unobservables. On the other hand, this assumption rules out the existence of autocorrelated unobservable state variables that difficult extremely the estimation of the decision problem.

Under assumptions AS and CI and the multiplicative decomposition given by (20), the optimal discrete choice can be written as:

$$d_{nt}^{*}=d \Longleftrightarrow d=\arg\max_{j=0,1}\left\{\pi^{j}(x_{nt})'\mu(\theta)+\varepsilon_{nt}^{j}+\beta\ EV^{j}(x_{nt};\theta)\right\}$$

The log-likelihood function for this problem is

$$\ln L=\sum_{n=1}^{N}\sum_{t=1}^{T_n}\sum_{d=0,1}\mathbf{1}\left(d_{nt}^{*}=d\right)\ln\left(\Pr(d_{nt}^{*}=d|\, x_{nt})\right) \tag{22}$$

where , for $d=\{0,1\}$,

$$P^{d}(x_{nt})=\Pr(d_{nt}^{*}=d|\,x_{nt})=$$

$$=\Pr\left\{d=\arg\max_{j=0,1}\left\{\pi^{j}(x_{nt})'\mu(\theta)+\varepsilon_{nt}^{j}+\beta\ EV^{j}(x_{nt};\theta)\right\}\middle|\, x_{nt}\right\}=$$

$$=\int\mathbf{1}\left\{d=\arg\max_{j=0,1}\left\{\pi^{j}(x_{nt})'\mu(\theta)+\varepsilon_{nt}^{j}+\beta\ EV^{j}(x_{nt};\theta)\right\}\right\}q\left(d\varepsilon|\, x\right)$$

These conditional choice probabilities entering the log-likelihood function are expressed in terms of unknown conditional value functions $EV^{d}(x_{nt};\theta)$. An obvious approach to estimate the structural parameters is a solution method consisting in some nested algorithm in the spirit of Rust's Nested Fixed Point (1987). This technique consists in an outer algorithm that maximizes the likelihood function and an inner algorithm which solves the dynamic programming problem, i.e., which computes the functions $EV^{d}(x_{nt};\theta)$, at each iteration in the search for the parameter estimates. The main drawback of this kind of techniques that solve the dynamic programming problem is its high computational cost.

In order to overcome this limitation, Hotz and Miller (1993) proposed an alternative estimation method, the Conditional Choice Probability (CCP) estimator, which allows to estimate the structural parameters without solving the dynamic programming problem. Is is based on the so-called Invertibility Proposition, which gives an alternative representation of the conditional value functions in terms of observable state variables, conditional choice probabilities, conditional transition probabilities and structural parameters. This estimation method has been applied in Aguirregabiria (1999) and Slade (1998) for the estimation of models of inventories and price change decisions. In Sánchez-Mangas (2002), we estimated a dynamic structural investment model similar to the model in this paper applying this method.

Although the CCP estimator has clear advantages over the NFXP algorithm in terms of computational cost, since it avoids the solution of the dynamic programming problem, it has a clear disadvantage in terms of the efficiency of the estimation. Thus, there is a trade-off between these two techniques in terms of computational cost and precision.

In a recent work, Aguirregabiria and Mira (2002) proposed the Nested Pseudo-Likelihood estimator (NPL), which has the computational advantages of the Hotz and Miller's CCP estimator, but allows to reach the effciency of the Rust's NFXP algorithm. As it is occurs with the CCP estimator, the NPL is based on the representation of conditional value functions in terms of observable state variables, conditional choice and transition probabilities and structural parameters. The keypoint of this estimation method is the so-called Policy Iteration operator. It is an operator in the space of the conditional choice probabilities:

$$P = \Psi\left(P\right) \equiv \Lambda\left(\varphi\left(P\right)\right)$$

where $\varphi(\cdot)$ is an operator which maps a vector of conditional choice probabilities into a vector of conditional value functions using Hotz and Miller's Invertibility Proposition . The operator $\Lambda(\cdot)$ maps a vector in the value function space into a vector of conditional choice probabities. Aguirregabiria and Mira (2002) show that the set of optimal choice probabilities P a is fixed point of $\Psi(\cdot)$. Thus, the NPL algorithm is, as the NFXP algorithm, a nested algorithm in which a fixed point problem must be solved. But this fixed point problem is not defined in the value function space, but in the probability space. In the NPL algorithm, unlike the NFXP algorithm, is the outer algorithm which computes the fixed point, while the inner algorithm iterates in a pseudo-likelihood function using Hotz and Miller's representation.

This representation of conditional value functions in terms of observable state variables, conditional choice and transition probabilities and structural parameters was reformulated by Aguirregabiria (1999), who showed that these value functions could be expressed as:

$$EV^d(x_{nt};\theta) = W^d(x_{nt})' \ \lambda(\theta)$$

where

$$W^d(x_{nt}) = \hat{F}^d(x_{nt}) \left(I - \beta\hat{F}(x_{nt})\right)^{-1} \left(\sum_{d=0,1} \hat{P}^d(x_{nt}) * \pi^d(x_{nt}) \quad \sum_{d=0,1} \hat{P}^d(x_{nt}) * g^d(x_{nt}) \right) \tag{23}$$

$\lambda(\theta) = (\mu(\theta)' \quad 1)'$, $*$ denotes the element-by-element product, the functions $g^d(x_t)$ are given byr:

$$g^d(x_{nt}) = E\left[\varepsilon^d_{nt} \middle| x_{nt}, d^*_{nt} = d \right]$$

and $\hat{P}^d(x_t)$, $\hat{F}^d(x_t)$ and $\hat{F}(x_t)$ are nonparametric estimators of the conditional choice probabilities, and the conditional and unconditional transition probabilities respectively.

The vector $W^d(x_t)$ is related to the expected and discounted stream of the future components associated with the corresponding components of the one period profit function $\pi^d(x_t)$. The conditional expectation of the unobservable state variables, $g^d(x_{nt})$, can be written if terms of conditonal choice probabilities. If we assume, for example, an extreme value distribution for ε^d_{nt}, this function is given by $E\left[\varepsilon^d_t \middle| x_t, d^*_t = d \right] = \gamma - \ln\left[P^d(x_t)\right]$, where γ is the Euler's constant. With this distributional assumption, it is straightforward from (23) to obtain a closed expression for the conditional value functions $EV^d(x_{nt};\theta)$.

For an arbitrary vector of choice probabilities P, the pseudo-likelihood function is defined as:

$$\tilde{l} = \sum_{n=1}^{N} \sum_{t=1}^{T_n} \sum_{d=0,1} \mathbf{1}\left(d^*_{nt} = d\right) \ln \Psi^d_\theta(x_{nt}, P) \tag{24}$$

where

$$\Psi_\theta^d(x_{nt}, P) = \frac{\exp\left\{\pi^d(x_t)'\mu\left(\theta\right) + \beta\ W^d(x_t)'\ \lambda(\theta)\right\}}{\sum\limits_{j=0,1} \exp\left\{\pi^j(x_t)'\mu\left(\theta\right) + \beta\ W^j(x_t)'\ \lambda(\theta)\right\}} \tag{25}$$

Once the pseudo-likelihood function is formulated, how does the NPL algorithm go? Let us assume that we have obtain nonparametric estimates of the conditional transition probabilities, $\hat{F}^d$, for $d = 0, 1$. Let $\hat{\theta}^{(0)}$ be an initial vector of parameters and $\hat{P}^{(0)}$ an initial vector of conditional choice probabilites (e.g, a nonparametric consistent estimator). For the iteration $R \geq 1$, the NPL algorithm consists in the following steps:

Step 1: To obtain the representation of the conditional choice value functions in terms of the conditional choice probabilities, using the Hotz and Miller's representation as in (23).

Step 2: To obtain a new pseudo-likelihood estimator $\hat{\theta}^{(R)}$:

$$\hat{\theta}^{(R)} = \arg\max_{\theta\in\Theta} \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{d=0,1} 1\left(d_{nt} = d\right| x_{nt}) \ \ \ln \Psi^d_{\hat{\theta}^{(R-1)}}\left(\hat{P}^{(R-1)}\right| x_{nt}\right)$$

where

$$\Psi^d_{\hat{\theta}^{(R-1)}}\left(\hat{P}^{(R-1)}\right| x_{nt}\right) = \Pr\left(d_{nt} = d\right| x_{nt}) \equiv \Psi^d\left(\hat{\theta}^{(R-1)}, \hat{P}^{(R-1)}, \hat{F}^d\right| x_{nt}\right)$$

Step 3: To update the vector of conditional choice probabilities using the estimator $\hat{\theta}^{(R)}$ obtained in step 2.

$$\hat{P}^{(R)} = \Psi^d\left(\hat{\theta}^{(R)}, \hat{P}^{(R-1)}, \hat{F}^d\right| x_{nt}\right)$$

Iterate in R until convergence in $\hat{P}$ and $\hat{\theta}$.

As it is showed in Aguirregabiria and Mira (2002), when the NPL is initialized with consistent estimators of the vector of conditional choice probabilities, successive iterations return a sequence of estimators, the R-stage Policy Iteration estimators, that includeas extreme cases the Hotz and Miller's CCP estimator (for $R = 1$) and the Rust's NFXP estimator (when $R \to \infty$). The gains in efficiency form the first to the second iteration is important, but the gains in succesive iterations is much lower. Furthermore, the asymptotic distribution of all the estimators in the sequence is the same and equal to that of maximum likelihood estimator.

5 Estimation results

Once we have estimated the profitability shock R_{nt} from the estimates of the technological parameters, we decompose it in an aggregate and an idiosyncratic shock $\tilde{R}_{nt}$, such that $R_{nt} = R_t \tilde{R}_{nt}$. Following Cooper and Haltiwanger (2000), the aggregate shock is simply the yearly mean of the profitability shock R_{nt}, and the idiosyncratic shock $\tilde{R}_{nt}$ is the deviation from that mean. Both components have been taken in logarithms, so

$$r_{nt} = r_t + \tilde{r}_{nt}$$

where $r_{nt} = \ln(R_{nt})$, $r_t = \ln(R_t)$ and $\tilde{r}_{nt} = \ln(\tilde{R}_{nt})$.

With respect to the capital stock K_{nt}, since its range of variability is very different for the different firms, we have considered the logarithm of the capital stock in deviation with respect to its firm mean. That is, we have considered $\tilde{k}_{nt} = \ln(K_{nt}) - \ln(\bar{K}_n)$, where $\bar{K}_n$ is the mean capital of firm n in the whole period in which this firm is observed. This means to consider some kind of heteroskedasticity in the model. The one-period profit function conditional on the decision, in terms of the mean capital stock, is given by:

$$\tilde{\pi}^1(s_{nt}) = R_{nt}\tilde{K}_{nt} - p_t\tilde{K}_{nt}i_{nt} - \frac{\theta_Q}{2}p_t\tilde{K}_{nt}\left(i_{nt}\right)^2 - \theta_F\tilde{K}_{nt} + \tilde{\varepsilon}^1_{nt}$$

$$\tilde{\pi}^0(s_{nt}) = R_{nt}\tilde{K}_{nt} + \tilde{\varepsilon}^0_{nt}$$

where $\tilde{K}_{nt} = K_{nt} / \bar{K}_n$ and $\tilde{\varepsilon}^d_{nt} = \varepsilon^d_{nt} / \bar{K}_n$, for $d = 0$ and $d = 1$.

Thus, the vector of observable state variables we use in the estimation is given by $x_{nt} = (p_t, r_t, \tilde{r}_{nt}, \tilde{k}_{nt})$. The main descriptive statistics for these variables are shown in Table 4.

	Mean	Std. deviation	Minimum	Maximum
r_t	1.6301	0.1061	1.4149	1.7394
$\tilde{r}_{nt}$	0	1.1749	-4.6440	5.2184
$\tilde{k}_{nt}$	1	0.1877	0.1881	2.5297
p_t	0.9917	0.0474	0.8668	1.1905

Table 4: Descriptive statistics of the observable state variables

The NPL estimation method, as in Rust's NFXP or Hotz and Miller's CCP estimators, requires a discretization of the observable state variables. The details on this discretization and on the initial estimates of the conditional choice probabilities and the conditional transition probabilities are shown in the Appendix.

In the NPL algorithm, the inner algorithm maximizes the pseudo-likelihood function. The conditional choice probabilities entering this function takes the expression of the probabilities in a logit model, in which the explanatory variables are the components of the vectors $\pi^d(x_{nt})$ and $W^d(x_t)$. In general, in this type of models it is not possible to identify the variance of the error term. However, in this case, since one of the explanatory variables, the one corresponding to the revenue function, appears with parameter restricted to be 1, it is possible to identify the variance of the error term.

The structural estimation results using the NPL algorithm are shown in Table 5. The discount factor β has been fixed at 0.975. We have estimated the model with different values of β (from 0.95 to 0.99) obtaining similar results.

Structural parameter estimates (NPL algorithm)						
	1 stage	2 stages	3 stages	4 stages	5 stages	6 stages
θ_Q	124.20	157.30	140.2	141.5	141.4	141.4
	(30.76)	(37.87)	(34.44)	(34.67)	(34.65)	(34.65)
θ_F	5.458	8.002	8.394	8.371	8.377	8.377
	(0.928)	(1.124)	(1.130)	(1.126)	(1.127)	(1.127)
σ_ε	8.807	9.550	9.305	9.317	9.318	9.318
	(1.061)	(1.218)	(1.140)	(1.142)	(1.142)	(1.142)
LogL	-3862	-3693	-3696	-3696	-3696	-3696
Pseudo-R^2	0.8129	0.8187	0.8186	0.8186	0.8186	0.8186

Table 5: Structural parameter estimates. Standard errors in parenthesis

We have obtained very precise estimates of all the parameters. Since the capital has been considered in terms of the firm's mean capital, the estimator of the fixed cost parameter implies that a firm with a capital stock of one million euro that decides to undertake an investment project must face a fixed cost which is equivalent to a percentage of its installed capital between 0.83% and 1.05% (95% confidence interval). Table 6 shows the median proportion that fixed and variable adjustment costs implied by our estimates represents on average over the installed capital stock and the sales in the considered categories: small, medium and large firms, according to the clasiffication criterion established by the European Commission.

Type of firm	VC/Cap. stock	FC/Cap. stock	VC/Sales	FC/Sales
Small firms	0.7319	3.3975	0.1348	0.6010
Medium firms	0.0343	0.1838	0.0088	0.0469
Large firms	0.0109	0.0431	0.0023	0.0111

Table 6: Proportion of adjustment costs over installed capital stock and sales.
VC: Variable adustment costs. FC: Fixed adjustment costs.

On average, fixed adjustment costs are much more relevant than convex costs in each category. It is worthwhile to emphasize the importance of fixed adjustment costs in small firms, representing around 3.34% of the installed capital stock and around 0.60% of the total sales. These proportion decreases a lot in the other categories: medium and large firms. In the group of medium firms, fixed adjustment costs represent 0.18% of the installed capital and 0.04% of the firm sales. In the group of large firms, fixed costs represent 0.04% of the installed capital and only 0.01% of the sales. The very different magnitude of fixed adjustment costs for the categories considered can explain the very different importance of inaction found in each of them. As we saw in Section 2, the percentage of observations accounting for zero investment was very high in small firms, 29.24%, while in medium and large firms this percentage was much smaller, 4.51% and 0.93% respectively.

There exist very few papers on structural estimation of a dynamic investment model with fixed adjustment costs. We can cite Cooper and Haltiwanger (2000), who estimate an investment model for american firms. Their estimations imply that for the firms analyzed, fixed adjustment costs represent approximately 0.04% of total profits. In our opinion, given the enormous heterogeneity which is present among firms, it is more informative to analyze the implications of the estimators distinguishing more homogeneous groups of frms.

Up to our knowledge, the work in this paper is the first one using the NPL estimator, which presents clear advantages over other previous estimation methods.

6 Conclusions

In this paper we have estimated a dynamic structural model of irreversible investment for Spanish manufacturing firms. The dataset we use exhibit some of the characteristics reported in the recent microeconomic investment literature. More specifically, we have found strong evidence of inaction and lumpy investment. Based on these empirical features, we have proposed a dynamic structural investment model in which irreversibilities and nonconvex adjustment costs have been included. The adjustment cost function we consider includes quadratic and fixed components.

We have stated the model through a dynamic programming problem of discrete choice, in which firms decide between buying some capital goods or postponing the purchase decision. The estimation method used in this paper has been the nested pseudo-likelihood (NPL) algorithm recently proposed by Aguirregabiria and Mira (2002). This method presents clear advantages over other estimation methods in this context. It is a technique based on Hotz and Miller's CCP estimator, i.e., it is based on a representation of the conditional value functions in terms of observable state variables, conditional choice and transition probabilities and structural parameters. Unlike Rust's NFXP, this estimation method does not require to solve the dynamic programming problem to obtain structural parameter estimates, so it has a clear computational advantage over NFXP estimator. Furthermore, successive iterations in the NPL algorithm return a sequence of estimators with the asymptotic distribution of the maximum likelihood estimator. Thus, from the point of view of the efficiency, it has clear advantages over CCP estimator. Up to our knowledge, the estimation exercise presented in this paper is the first one applying this new estimation technique.

Our estimation results reflect the importance of fixed adjustment costs, which can represent a considerable proportion of installed capital and sales. The magnitude of these costs varies a lot depending on the firm size. For small firms they can represent around 3.4% of the installed capital and 0.60% of the firm sales, while for medium and large firms these proportions are much smaller. This can explain the observed investment behavior in our dataset. In the group of small firms, approximately 29% of the observations accounted for zero investment, while this percentage was much smaller for medium and large firms.

Appendix

A1. CONSTRUCTION OF VARIABLES

Employment: Number of employees at december 31th, is the sum of permanent workers and the average number of temporary workers. The weights to calculate the average number of temporary workers is: 1/4 if the average time in the firm is less than 6 months, 3/4 if it is more than 6 months and less than one year and 1 if it is more than one year.

Output: Gross output at retail prices is calculated as total sales.

Capital stock: The dataset contains information on the book value and the average age of the stock of fixed capital and the year of the last regulation. It also includes data on gross nominal investment during the year. Following Alonso-Borrego and Collado (1999), taking period t as reference year, the market value of the stock of fixed capital in period t is calculated as:

$$K_{nt} = (1 - \delta_n)^{age_{nt}} KB_{nt} \frac{q_t}{q_{m_n}}$$

where age_{nt} is the average age of the capital stock of firm n at period t, δ_n is the depreciation rate of the sector in which firm n operates, KB_{nt} is the book value of the stock of fixed capital, q_t is the price deflator of the stock of fixed capital and m_n is the year of the last regulation in firm n. The price index is the GDP implicit deflator of investment goods, which is constant over time. The depreciation rate varies across sectors.

Taking t as the reference year, the market value of the stock of fixed capital for any year $s \neq t$ is calculated using a perpetual inventory method:

$$K_{ns} = (1 - \delta_n) K_{n,s-1} \frac{q_s}{q_{s-1}} + I_{ns} \quad \text{if} \quad s > t$$

$$K_{ns} = \frac{(K_{n,s+1} - I_{n,s+1})}{(1 - \delta_n)} \frac{q_s}{q_{s+1}} \qquad \text{if} \quad s < t$$

where I_{ns} is the investment accounted by the firm n in period t. Using this approach it is possible to obtain negative values of K_{ns} for $s < t$. In that case the market value of the capital stock is set to missing. In an attempt to reduce this problem, the market value of the capital stock for any firm has been calculated using different years as reference. Finally, the reference year was chosen to minimize the number of missing values in the capital stock.

A2. PRODUCTION FUNCTION ESTIMATES

The following table shows the standard GMM and system GMM estimates for the production function with constant retuns to scale. Standard errors are in parenthesis. We have performed some specifcation tests: $m1$ and $m2$ are tests for first and second order correlation in the first differenced residuals. *Sargan* is the statistic for the Sargan test of overidentifying restrictions. *Dif Sargan* is the statistic for the test of the validity of the additional instruments used in the system estimation with respect to the standard estimation. MD stands for minimum distance. Year dummies have been included in all models.

GMM estimates		
	First differences $t-2$	System $t-2$
k_{nt}	0.289 (0.154)	0.436 (0.113)
$k_{n,t-1}$	-0.301 (0.117)	-0.368 (0.092)
$y_{n,t-1}$	0.864 (0.090)	0.844 (0.061)
$m1$ p-value	-7.446 0.000	-10.659 0.000
$m2$ p-value	0.036 0.971	-0.245 0.807
Sargan p-value	25.811 0.529	39.539 0.357
Dif. Sargan p-value	— —	13.728 0.800
Minimun distance estimates		
α_k	0.357 (0.133)	0.436 (0.109)
α_L	0.643 —	0.564 —
ρ	0.878 (0.088)	0.844 (0.058)
p-value MD test	0.382	0.993

A3. DISCRETIZATION OF THE STATE VARIABLES

The aggregate shock r_t has been discretized in only two cells corresponding to low and high shock. The idiosyncratic shock and the capital stock have been discretized in 7 cells using a unifom grid on the empirical distribution of these variables. Due to the very low variability of the capital price in the dataset, it has been taken as constant. Besides, preliminary analysis on the relevance of this variable on the firms' investment pattern in the dataset yield to consider it nonsignificant at the usual levels. The discretization we have carried out yields 98 cells in the space of the state variables.

El shock agregado r_t ha sido discretizado en dos únicas celdas, indicando valores bajo y alto de dicho shock. En cuanto al shock idiosincrático y al capital, se han discretizado en 7 celdas utilizando una parrilla uniforme en la función de distribución empírica de ambas variables. En cuanto al precio del capital, éste presenta poca variabilidad en los datos, y por tanto, se ha tomado como constante. Además, análisis preliminares acerca de la relevancia de esta variable sobre las pautas de inversión de las empresas de nuestra base de datos llevaron a considerarla no significativa a los niveles habituales. La discretización que hemos llevado a cabo da lugar a un total de 98 celdas en el espacio de las variables de estado.

A4. NONPARAMETRIC ESTIMATION OF CONDITIONAL CHOICE PROBABILITIES AND CONDITIONAL TRANSITION PROBABILITIES

We have obtained nonparametric estimates of the probability that a high (low) value of the aggregate shock is followed by a high (low) value, obtaining the following transition probability matrix for the aggregate shock:

$\Pr(r_{t+1} \mid r_t)$	low r_{t+1}	high r_{t+1}
low r_t	0.682	0.318
high r_t	0.318	0.682

Let us denote by M_1, M_2 and M_3 the number of cells in the discretizaton of the variables $r_t, \tilde{r}_{nt}$ and $\tilde{k}_{nt}$ respectively. In this case, $M_1 = 2$ and $M_2 = M_3 = 7$. Let $m = 1, ..., M$ be the index for the cells of tridimensional state variable $x_{nt} = \left(r_t, \tilde{r}_{nt}, \tilde{k}_{nt}\right)$, where $M = M_1 \times M_2 \times M_3 = 2 \times 7 \times 7 = 98$. Let $r^c, \tilde{r}^c$ and $\tilde{k}^c$ be the values of the discretized state variables and let $r^m, \tilde{r}^m$ and $\tilde{k}^m$ be the values of discretized state variables correspondig to the m-th cell, that is, $x^m = (r^m, \tilde{r}^m, \tilde{k}^m)$.

The initial estimates of the conditional choice probabilities and the conditional transition probabilities of the capital stock and the idiosyncratic shock have been obtained using trivariate kernel estimators.

The conditional choice probability $\Pr(d = 1|\, x^m)$ has been estimated as:

$$\overset{\wedge}{\Pr}(d = 1|\, x^m) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T_n} \mathbf{1}(d_{nt} = 1)\, K_3(x_{nt}, x^m)}{\sum_{n=1}^{N}\sum_{t=1}^{T_n} K_3(x_{nt}, x^m)}, \qquad \text{for } m = 1, ..., M$$

where K_3 is the trivariate gaussian kernel:

$$K_3(x_{nt}, x^m) = \frac{1}{(2\pi)^{3/2}} \exp\left\{-\frac{1}{2}\left[\left(\frac{r_t - r^m}{h_1}\right)^2 + \left(\frac{\tilde{r}_{nt} - \tilde{r}^m}{h_2}\right)^2 + \left(\frac{\tilde{k}_{nt} - \tilde{k}^m}{h_3}\right)^2\right]\right\}$$

where h_1, h_2 and h_3 are bandwith parameters chosen using the Silverman's rule.

Since $\tilde{r}_{nt}$ is an exogenous variable, its conditional transition probability is estimated as:

$$\overset{\wedge}{\Pr}\left(\tilde{r}^c_{t+1} = r^m \middle|\, \tilde{r}^c_t = r^l\right) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T_n} 1\left(\tilde{r}^c_{n,t+1} = r^m\right) K_1\left(\tilde{r}_{nt}, r^l\right)}{\sum_{n=1}^{N}\sum_{t=1}^{T_n} K_1\left(\tilde{r}_{nt}, r^l\right)}$$

for $m, l = 1, ..., M_2$, where K_1 is a univariate gaussian kernel:

$$K_1\left(\tilde{r}_{nt}, r^l\right) = \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}\left(\frac{\tilde{r}_{nt} - r^l}{h_1}\right)^2\right\}.$$

The capital stock is an endogenuous variable and we must estimate the conditional transition probability conditional on $d = 0$ and conditional on $d = 1$. We have obtained nonparametric estimates of these probabilities:

$$\overset{\wedge}{\Pr}\left(\tilde{k}^c_{t+1} = \tilde{k}^l \middle|\, x^m, d\right) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T_n} \mathbf{1}\left(\tilde{k}^c_{n,t+1} = \tilde{k}^l\right) \mathbf{1}(d_{nt} = d)\, K_3(x_{nt}, x^m)}{\sum_{n=1}^{N}\sum_{t=1}^{T_n} \mathbf{1}(d_{nt} = d)\, K_3(x_{nt}, x^m)}$$

for $d = 0, 1$, $l = 1, ..., M_3$ and $m = 1, ...M$.

From these estimates we obtain the $M \times 1$ vector $P^1(x) = \Pr(d = 1|\, x)$ of estimated conditional choice probabilities and the $M \times M$ matrices $F^1(x)$ and $F^0(x)$ of estimated transition probabilities of the state variables, conditional on $d = 1$ and $d = 0$ respectively.

A5. ESTIMATION OF THE AMOUNT OF INVESTMENT IF $d = 1$

The functions $E\left[i_{nt} \middle| x_{nt}, d_{nt} = 1 \right]$ and $E\left[(i_{nt})^2 \middle| x_{nt}, d_{nt} = 1 \right]$ appear in the one-period profit function conditional on the decision d, $\tilde{\pi}^d(s_{nt})$. Following a methodology similiar to Slade (1998), we have obtained nonparametric estimates of these expectations. First, we have discretized the variable $\{i_{nt};\ d_{nt} = 1\}$, that is, considering the observations such that $i_{nt} > 0$, using a uniform grid on the empirical distribution function of this variable. Let H be the number of cells in this discretization. We have considered $H = 7$. Let i^c be the value of the discretized investment rate and i^h the value of the discretized investment rate in the cell $h = 1, ..., H$. The function $E\left[i \middle| x^m, d = 1 \right]$, for $m = 1, ..., M$, has been estimated as:

$$\sum_{h=1}^{H} i^h \Pr\left(i^h \middle| x^m, d = 1 \right)$$

where the probability $\Pr\left(i^h \middle| x^m, d = 1 \right)$ has been estimated nonparametrically as:

$$\Pr\left(i^h \middle| x^m, d = 1 \right) = \frac{\sum_{n=1}^{N} \sum_{t=1}^{T_n} \mathbf{1}\left(i^c_{nt,} = i^h \right) \mathbf{1}\left(d_{nt} = 1 \right) K_3\left(x_{nt}, x^m \right)}{\sum_{n=1}^{N} \sum_{t=1}^{T_n} \mathbf{1}\left(d_{nt} = 1 \right) K_3\left(x_{nt}, x^m \right)}$$

for $h = 1, ..., H$ y $m = 1, ..., M$. So we have obtained the $M \times 1$ vector of estimated values of $E\left[i \middle| x, d = 1 \right]$. The $M \times 1$ vector of estimated values of $E\left[i^2 \middle| x, d = 1 \right]$ has been estimated in a similar way.

REFERENCES

Aguirregabiria, V. (1999): "The dynamics of markups and inventories in retailing firms", *Review of Economic Studies*, 66, 275-308.

Aguirregabiria, V. and Mira, P. (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models", *Econometrica*, 70, 1519-1543.

Alonso-Borrego, C. and L. Collado (1999): "Factor demands and innovation activity: Evidence from Spanish manufacturing firms", Documento de Trabajo 99-02, Fundación Empresa Pública.

Alonso-Borrego, C. and R. Sánchez-Mangas (2001): "GMM estimation of a production function with panel data: an application to Spanish manufacturing firms", Working Paper 01-55 (27), Universidad Carlos III de Madrid.

Arellano, M. and O. Bover (1995): "Another look at the instrumental variable estimation of error-component models", *Journal of Econometrics*, 68, 29-52.

Barnett, S. and P. Sakellaris (1995): "Non-linear response of firm investment to Q: Testing a model of convex and non-convex adjustment costs", *Journal of Monetary Economics*, 42, 261-288.

Bertsekas (1976): "Dynamic programming and stochastic control", New York, Academic Press.

Cooper, R. and J. Haltiwanger (2000): "On the nature of capital adjustment costs", *NBER* , Working Paper Series No. 7925

Cooper, R., J. Haltiwanger and L. Power (1999): "Machine replacement and the business cycle: lumps and bumps", *American Economic Review*, 89, 921-946.

Doms, M and T. Dunne (1998): "Capital adjustment patterns in manufacturing plants", *Review of Economic Dynamics*, 1, 409-429

Hotz, J. and R. A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models", *Review of Economic Studies*, 60, 497-529.

Nilsen, O. and F. Schiantarelli (1998): "Zeroes and lumps in investment: Empirical evidence on irreversibilities and non-convexities", Mimeo, Boston College.

Rust, J. (1987): “Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher”, *Econometrica*, 55, 999-1034.

Sánchez-Mangas, R. (2002): ”Estimation of a dynamic discrete choice model of irreversible investment”, Working Paper 01-56 (28), Universidad Carlos III de Madrid.

Scarf, H. (1959): “The optimality of (S,s) policies in the dynamic inventory problem”, in K.J. Arrow, S. Karlin and P. Suppess (eds), *Mathematical Methods in the Social Sciences*, Stanford University Press.

Slade, M. (1998): “Optimal pricing with costly adjustment: Evidence form retail-grocery prices”, *Review of Economic Studies*, 65, 87-107.