Corporate Investment, Irreversibilities and Lumpiness: An Empirical Model[*]

César Alonso-Borrego
Universidad Carlos III de Madrid

Rocío Sánchez-Mangas[†]
Universidad Autónoma de Madrid

June 2008

Abstract

We study the role of irreversibility and non convexities in firm investment decisions. For such purpose, we posit a dynamic structural investment model with irreversibility and nonconvex adjustment costs. We focus on the firm decision about whether to invest or not, which is characterized by means of a discrete choice dynamic programming problem. The adjustment cost parameters behind the investment decision are estimated with a longitudinal sample of Spanish manufacturing firms between 1990 and 2002. For these firms, we confirm that inaction and investment episodes account for a significant fraction of them. As estimation procedure, we apply the Nested Pseudo-Likelihood (NPL) algorithm by Aguirregabiria and Mira (2002).

JEL Codes: C25, C23, C13, D21
Keywords: Capital adjustment costs, Irreversible investment, Structural estimation, Discrete choice models, Dynamic programming, Nested algorithms

[*]We thank Victor Aguirregabiria, Pedro Mira, Alfonso R. Sánchez and seminar participants at Universidad Carlos III de Madrid, Universitat Autonoma de Barcelona, Universidad de Navarra, Universidad de Vigo and Universitat de les Illes Balears for helpful comments. We acknowledge research funding from Ministry of Education, Grants No. SEJ2006-05710/ECON and SEJ2006-04957/ECON, respectively. The second author also thanks funding form Comunidad de Madrid, Grant No. CCG07-UAM/HUM-1918.

[†]Contact address: Rocío Sánchez-Mangas, Dpto. de Análisis Económico: Economía Cuantitativa, Universidad Autónoma de Madrid. Campus de Cantoblanco, Avda. Francisco Tomás y Valiente 5,.28049 Madrid (Spain). Phone: + 34 914975522. e-mail: rocio.sanchez@uam.es

1 Introduction

The …rms' behaviour about investment in …xed capital has been widely studied in the last four decades because of its central role in economic growth and aggregate ‡uctuactions. The adjustment costs and irreversibility behind …rm investment decisions introduces interesting dynamic considerations. The baseline investment model arises from the neoclassical factor demand model in which the adjustment in the amount of physical capital that is productive is costly and takes time to complete (see Chirinko, 1993). Adjustment costs for investment have traditionally been assumed to be strictly convex, what usually provides a good description for the behaviour of investment in physical capital at the aggregate level. Nonetheless, at the …rm level, a signi…cant fraction of …rms show moderately long episodes of zero investment, followed by episodes of sharp positive investment which amounts a high percentage of installed capital. This evidence of infrequent and lumpy adjustments challenges the smooth adjustment pattern implied by the standard neoclassical investment model with convex adjustment costs. Since then, richer adjustment cost structures have been proposed (see Bertola and Caballero, 1994; Dixit and Pindyck, 1994; and Abel and Eberly, 1994 and 1997; Cooper and Haltiwanger (2006); among others). The structure of adjustment costs determines the timing of investment decisions and may entail rigidities in the process of capital accumulation, and in turn it may have macroeconomic consequences.

In this paper, we propose an empirical model of …xed capital investment at the …rm level, and estimate it using an unbalanced longitudinal sample of $1;428$ Spanish manufacturing companies. The preliminary analysis of our data con…rms the importance of infrequent and lumpy investment. In accordance with such evidence, we posit a dynamic model of investment with irreversibility and a general speci…cation of adjustment costs which includes convex and nonconvex components. We are very intrigued about the importance of di¤erent components of adjustment costs in the pattern of …rm investment.

Our methodological approach is built upon a dynamic programming framework. We concentrate on the …rm discrete decision on whether to invest or not, and derive the corresponding discrete choice dynamic programming problem. Unlike models for a continuous decision, the optimal decision rule of a dynamic discrete decision cannot be expressed as a set of differentiable first order conditions. Instead, optimal decision rules for dynamic discrete decisions are characterized by inequality conditions. Essentially, the econometric model consists on a discrete choice model that is nonstandard due to the fact that the critical thresholds depend on the comparison of the value functions evaluated at each alternative. This requires the resolution of a nested fixed point algorithm, which must solve the dynamic programming model at each iteration of the parameter estimation problem. In such spirit, Rust (1987) proposed a solution-estimation algorithm, which has been scarcely applied because of its high computational cost. Hotz and Miller (1993) exploited the existence of a one-to-one mapping between the normalized value functions and the conditional choice probabilities to propose the Conditional Choice Probability (CCP) estimator, which circumvents the need of solving the model at each iteration. The CCP estimator was shown much less computationally demanding, yet at the expense of efficiency. In Sánchez-Mangas (2002) the CCP is applied to the estimation of a dynamic structural model of irreversible investment. Recently, Aguirregabiria and Mira (2002) have proposed an estimation method, the Nested Pseudo-Likelihood (NPL) algorithm, based on a representation of the dynamic programming model solution in the space of conditional choice probabilities. The NPL bridges the gap between the two estimation strategies mentioned above, including the Hotz and Miller's CCP estimator and Rust's NFXP estimators as extreme cases. In this paper, we apply the NPL algorithm to estimate the adjustment cost parameters of our dynamic discrete choice model of investment. Up to our knowledge, this is the first exercise of application of this estimation method.

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

2 Data and preliminary evidence

The main data set is an unbalanced panel of Spanish manufacturing companies, recorded in the database Encuesta sobre Estrategias Empresariales (Survey on Business Strategies, ESEE hereinafter). The ESEE is produced by Fundación Empresa Pública, a public institute financed by the Spanish Ministry of Industry. The original data set was designed with the aim of providing a representative sample of Spanish manufacturing firms. For this purpose, all the companies with more than 200 employees were surveyed (and, accordingly with the information provided by those responsible for the data set, about 70% completed the survey), and smaller companies with more than 10 employees were selected on the basis of a stratified sampling. The data contain annual information at the firm level of the balance sheet and other economic variables. Our final sample contains $1,428$ companies between 1990 and 2002, whose nature has not been substantially altered in the sample period (so that we discard mergers or splits), with nonmissing information about the variables relevant for the study in at least four consecutive years. We also exclude observations with extreme changes in fixed capital, output, intermediate inputs and in the wage bill.[1]

We focus on grosss investment in fixed capital. The investment rate at year t is defined as the ratio between gross expenditure in that year and the capital stock at the beginning of that year. We analyze the investment rate for the whole dataset and by firm size. In Figure 1, we show the sample distribution of annual firm-level gross investment rates, that we have right-censored for investment rates above 100 percent. The distribution is strongly skewed to the right. A significant fractions of firms do not invest or invest very little. The investment rate is lower than 2.5% in 20 percent of cases, and 13 percent of observations show zero investment. On the other hand, we observe a significant fraction of observations on the right tail, pointing out that a large fraction of firms experience a large investment episode at some time. The evidence provided by Figure 1 resembles the findings in recent empirical studies

[1] The criterion used is based in the quantiles 1 and 99 for relative changes in these selected variables (defined symmetrically as the change in the variable with regard to the average value of the variable in t and $t-1$).

about investment behavior by …rms in di¤erent countries, like Barnett and Sakellaris (1995), Doms and Dunne (1998) or Nielsen and Schiantarelli (1998), among others. The pattern of smooth capital adjustment predicted by the baseline investment model is thus contradicted by the data. Instead, many …rms decide not to invest during a few years (inaction), and, when they do, its investment amounts a high proportion of installed capital (lumpiness).

[Figure 1]

In Table 1 we provide further evidence about …rm behaviour on investment. For each year, we report the percentage of observations with zero investment in the …rst column, and the percentage of cases with investment rates above 20 percent, in the second one. In any sample year, at least 10 percent of …rms do not invest at all. In addition, at least 24 percent of …rms show an investment spike larger than 20 percent of installed capital. Furthermore, 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 …xed capital underwent, respectively, a decrease of 0.68% and 11.72% with respect to 1992. On the contrary, investment spikes are a cyclical phenomena.

[Table 1]

We also analyze the incidence of infrequency and lumpiness distinguishing by …rm size: small, medium and large …rms.[2] In Table 2, we show the distribution of …rms in the sample in the …rst 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. In accordance with the size de…nition, about half of the …rms in our sample are small (up to 50 employees). The most striking fact is that the frequency of inaction

[2] We follow the classi…cation criterion established by the European Commission. According to this criterion, small …rms are those with no more than 50 employees and no more than 7 million euro of annual turnover. Medium …rms 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 …rms are those with more than 250 employees and an anual turnover greater than 40 million euro.

di¤ers very much by …rm size. Almost a quarter of small …rm observations experience zero investment, in contrast to medium and large …rms, for which inaction frequencies amount to 4:7 and 1:5 percent, respectively. Hence, the probability inaction appears to tend towards zero as …rm size increases, so that indivisibilities of capital goods are likely to be behind inaction in smaller …rms. Nonetheless, the incidence of lumpiness, does not di¤er across …rms sizes, being greater than 30 percent in all categories.

[Table 2]

In order to provide a better description of the investment behaviour by …rm, we have ranked, following Doms and Dunne (1998) each …rm investment rate in descending order, showing the mean and the median investment rate in each rank in Figure 2. We can see that the average investment rate of the two highest investment realizations by …rm is above 30 percent, exceeding 50 percent in the highest episode, and drops signi…cantly in the subsequent occurrences. The average is below 20 percent for the third highest realization, and drops to 12 percent after that. The median reproduces a similar pattern, although with lower values, in accordance with the right-skewed distribution of the investment rate.

[Figure 2]

In order to have a clearer insight of the relevance of investment spikes, we have computed, for each rank of investment constructed earlier, its weight in total investment of the …rm in the sample period. We present this information on Figure 3, from which we can see that, on average, about 44 percent of investment has been carried out in an only year, and about two thirds in just two years.

[Figure 3]

In addition, we look at the importance of large investment spikes by time. In Figure 4, the solid line shows the fraction of observations in the year whose investment rate is above 20%, and the dotted line shows the percentage of total investment in that year attributed to these lumpy observations. As it was seen in Table 1, observations with investment spikes in any year amount between 24 and 37 percent of observations, and account for a fraction of total investment in the year between 42 and 74 percent. On average, lumpiness accounts approximately for half of the total gross investment, and smoother adjustments account for another half. Our evidence resembles the empirical …ndings in Cooper, Haltiwanger and Power (1999) for a large set of US manufacturing …rms.

[Figure 4]

This descriptive analysis of investment behavior in Spanish manufacturing …rms 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. An appropriate empirical model at the micro level ought be able to capture this sort of behaviour, since the empirical predictions of the baseline neoclassical model are in contradiction with the data. We then take into account irreversibilities and non convex components in our speci…cation of adjustment costs.

3 A dynamic structural model of …xed capital investment

3.1 Theoretical framework

Our theoretical setup follows closely Sánchez-Mangas (2002) and Cooper and Haltiwanger (2006). Consider a risk neutral …rm that produces an homogeneous good using as inputs labor and capital equipment with some …rm-speci…c characteristics. At each year, the …rm decides its level of employment and purchases of new capital in order to maximize the expected discounted stream of current and future pro…ts over an in…nite time horizon. We assume that the …rm behaves as a price taker with respect to input prices, and thus they are exogenous to the …rm. However, we allow for imperfect competition in the product market, asssuming monopolistic competition, by which the …rm faces a constant elasticity demand function. Its current gross pro…t function at period t (gross of investment decisions), in output units, is given by:

$$\Pi_t = Y_t(K_t; L_t; A_t) - w_t L_t \tag{1}$$

where Y_t is real output, which depends on K_t, the capital stock installed at the beginning of the period; L_t, physical units of labor, and a productivity shock A_t; and w_t is the wage rate relative to output price. We assume that labor can be adjusted costlessly, so the decision on employment is static, and then the optimal condition for labor can be obtained and substituted into current gross profits to obtain

$$\Pi_t = Y_t(K_t; L_t^*; A_t) - w_t L_t^* \tag{2}$$

where L_t^* is the optimally chosen employment level.[3] Hence, the profit function in terms of capital stock can be written as:

$$\Pi_t = R_t K_t^{\theta_m} \tag{3}$$

where R_t is a profitability shock, which depends on the productivity shock, the relative wage and the technological parameters. Assuming a Cobb-Douglas production function

$$Y_t = A_t K_t^{\alpha_K} L_t^{\alpha_L}$$

and constant returns to scale, and recalling our assumption for imperfect competition in the product market, the expression for the markup parameter θ_m would be

$$\theta_m = \frac{(1-\alpha_L)(1-\varepsilon)}{1-\alpha_L(1-\varepsilon)} \tag{4}$$

where α_L is the labor technological parameter and ε is the inverse of the product demand elasticity. Under perfect competition, $\varepsilon = 1$.

The net current profits (net of investment) are given by

$$\pi_i = \Pi_t - p_t K_t i_t - AC(K_t; i_t; p_t),$$

[3] Assuming a Cobb-Douglas production function $Y_t = a_t K_t^{\alpha_K} L_t^{\alpha_L}$ with constant returns to scale, optimal choice of labor is given by

$$L_t^* = \left[\frac{A_t^{1-\varepsilon}\alpha_L\left(1+\frac{1}{\varepsilon}\right)}{w_t}\right]^{\frac{1}{1-\alpha_L(1-\varepsilon)}} K_t^{\frac{(1-\alpha_L)(1-\varepsilon)}{1-\alpha_L(1-\varepsilon)}},$$

where ε is the inverse of demand elasticity.

where p_t is the unit price of capital relative to product price, and i_t denotes the investment rate, de…ned as I_t / K_t, where I_t denotes gross investment (new capital purchases). It is worth noticing that the investment problem is de…ned in terms of the investment rate i_t; instead of the investment in physical units, I_t.[4] Besides, when the …rm acquires new capital equipment, it faces some adjustment costs, which we represent through the function $AC(K_t; i_t; p_t)$. We assume there is one period time-to-build, i.e, the new equipment is productive one period after its acquisition. Capital retirement and physical depreciation are exogenously given to the …rm. The capital stock follows a transition rule given by

$$K_{t+1} = K_t \left((1 ¡ \pm_t) + i_t \right), \tag{5}$$

where $\pm_t$ 2 $(0; 1)$ is the depreciation rate, which includes not only the economic depreciation of the capital stock but also the capital scrapping due to obsolescence.

The speci…cation for adjustment costs includes both variable and …xed costs components:

$$AC(K_t; i_t; p_t) = VC(K_t; i_t; p_t) + FC(K_t) \tag{6}$$

Variable costs $VC(¢)$, which include those costs associated with the installation of new capital equipment, are assumed to be convex, following the standard speci…cation used in the baseline neoclassical investment model. Our speci…cation for variable costs considers the following quadratic function:

$$VC_t = VC(K_t; i_t; p_t) = \frac{1}{2} µ_Q p_t K_t i_t^2, \tag{7}$$

$µ_Q$ being a constant parameter. The …xed adjustment cost component $FC(¢)$ captures internal costs entailed by …rm reorganization needed to make the new equipment fully productive, such as reorganization of the productive process and retraining of employees. We assume these costs to be proportional to the installed capital stock:

$$FC_t = FC(K_t) = 1(i_t > 0)\, µ_F K_t, \tag{8}$$

[4] There is widespread evidence in empirical work on …rms' behavior about the large amount of …rm heterogeneity even after controlling for observed characteristics. For this reason, we have chosen as decision a measure of investment normalized by the installed capital stock.

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

We assume that the investment decision is completely irreversible, i.e., the firm decides its purchases of new equipment that, after being acquired, cannot be sold.[5] The firm then faces the decision of inaction (not invest) or to undertake a strictly positive investment, so the decision variable in this problem is $i_t \geq 0$. Let $s_t = (K_t, L_t, A_t)$ be the vector of state variables at period t. The firm's decision problem can be written as:

$$\max_{\{i_t \geq 0\}} \sum_{t=0}^{\infty} \beta^t E\left[\pi(i_t, s_t)\right] \tag{9}$$

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:

$$V(s_t) = \max_{\{i_t \geq 0\}} \pi(i_t, s_t) + \beta EV(s_{t+1} | s_t, i_t) \tag{10}$$

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

$$EV(s_{t+1} | s_t, i_t) = \int V(s_{t+1}) \phi(ds_{t+1} | s_t, i_t) \tag{11}$$

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

3.2 Optimal decision rule

In our model, each firm faces a double decision: the discrete choice on whether to invest or not, and, the continuous decision about the amount of positive investment if it decided to invest. There are two potential reasons behind the zero investment decision. On the one hand, the irreversibility of investment because of the imposibility of re-selling purchased capital once after acquired by the firm. On the other hand, the existence of sizeable fixed adjustment costs, which hamper small capital adjustments, and may eventually force the firms to postpone worthy investment. These two sources of inaction make that the investment rate is censored at zero.

If irreversibility were the only source of censoring in investment, then the value function would keep being continuous and concave. However, the existence of fixed

[5] In contrast with models assuming the existence of imperfect second-hand markets in which the selling price for capital is lower than its true value, here we take the extrem case in which the second-hand market does not exist because the re-selling price is zero.

adjustment costs brings a discontinuity in the one-period profit function, making the value function to be nonconcave. The decision rules for these class of problems have been characterized by Bertsekas (1976), using properties of κ-concave functions. Scarf (1959), Slade (1998) or Aguirregabiria (1999) are examples of these type of decision rules in the context of inventories and price adjustment models.

The optimal decision rule for investment is given by:

$$i(s_t;\theta) = \begin{cases} i^*(s_t;\theta) & \text{if} \quad i^*(s_t;\theta) > 0 \text{ and } \gamma(s_t;\theta) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

where θ is the vector of structural parameters and $i^*(s_t;\theta)$ is the optimal interior solution characterized by

$$\pi_i(s; i^*(s;\theta);\theta) + \beta EV_i(s; i^*(s;\theta);\theta) = 0; \tag{13}$$

with $\pi_i \equiv \partial\pi/\partial i$ and $EV_i = \partial EV/\partial i$ and the function $\gamma(s_t;\theta)$ is given by

$$\pi(s_t; i^*(s_t;\theta);\theta) - FC(s_t;\theta) - \pi(s_t;0;\theta) + \beta\left[EV(s_t; i^*(s_t;\theta);\theta) - EV(s_t;0;\theta)\right]: \tag{14}$$

Hence, the set of optimality conditions are given by a first order condition for the interior solution, given by (13), and the two inequalities determining the discrete choice between an interior and a corner solution. The first inequality, $i^*(s_t;\theta) > 0$, features the non-negativity constraint due to the irreversibility of the investment decision, so that the interior solution will be optimal only if optimal investment is positive. If condition (13) holds for a negative value $i(s_t;\theta) < 0$, then the firm will choose $i(s_t;\theta) = 0$; due to irreversibility. The second inequality, $\gamma(s_t;\theta) > 0$, is related to the existence of fixed adjustment costs: $\gamma(s_t;\theta) > 0$ requires intertemporal profits to be high enough so as to overcome fixed costs of investment.

Our model is a dynamic choice model in which the decision variable is left-censored at zero. The two sources of censoring, irreversibility and fixed adjustment costs, discussed earlier, 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 does not suffice to overcome the fixed costs of adjustment, the optimal decision is also inaction.

We have seen that the full set of first order conditions for the optimal investment rule (12) consists on a marginal condition of optimality for interior solutions and threshold conditions for the optimal discrete choice on whether to invest or not. This set of optimality conditions depends on the structural parameters. Our estimation strategy will disregard the first order condition for strictly positive investment, and exploit only those conditions determining the optimal discrete choice between zero and nonzero investment.[6]

4 Model estimation

We have an unbalanced 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. Our econometric model consists on the profit function, the stochastic process for the profitability shock, and the dynamic model for investment. The set of structural parameters includes the markup parameter in the gross profit function, μ_m; the parameters that describe the transition probabilities of input prices and profitability shock; the adjustment costs parameters: μ_Q and μ_F; and the parameters of the distribution of the profitability shock.

For estimation purposes, we proceed in two stages. In a first stage, we measure gross profits ¦$_{nt}$ and capital K_{nt} and exploit the gross profit function (3) so as to estimate the markup parameter and the transition probabilities of the state variables. Afterwards, we can recover profitability shock R_{nt} through (3). This strategy has been also used by Cooper and Haltiwanger (2006). . In a second stage, we estimate the remaining structural parameters. For that purpose, we exploit the optimal discrete choice on whether to invest or not, to obtain estimates of the adjustment costs parameters μ_Q and μ_F.

[6] 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.

4.1 Estimation of the gross profit function

With respect to the estimation of the gross profit function, taking natural logarithms in (3), we have

$$\pi_{nt} = \theta_m k_{nt} + r_{nt} \qquad (15)$$

where $\pi_{nt} = \ln(\Pi_{nt})$, $k_{nt} = \ln(K_{nt})$; $r_{nt} = \ln(R_{nt})$. We allow the following structure for the (unobserved) profitability shock:

$$\begin{aligned} r_{nt} &= A_t + \eta_n + v_{nt} \\ v_{nt} &= \rho v_{n,t-1} + u_{nt} \\ u_{nt} &\sim (0, \sigma_u^2) \end{aligned}$$

where A_t is an aggregate effect, η_n is a time invariant firm-specific effect, and v_{nt} is a first-order autoregressive idiosyncratic shock. Substituting in (15), we get the following dynamic representation,

$$\pi_{nt} = \gamma_1 \pi_{n,t-1} + \gamma_2 k_{nt} + \gamma_3 k_{n,t-1} + A_t^* + \eta_n^* + u_{nt} \qquad (16)$$

where $A_t^* = A_t - \rho A_{t-1}$, $\eta_n^* = (1-\rho)\eta_n$, and the parameters satisfy the following restrictions in terms of the parameters of interest as

$$\begin{aligned} \gamma_1 &= \rho, \\ \gamma_2 &= \theta_m, \\ \gamma_3 &= -\theta_m \rho. \end{aligned}$$

The model is thus overidentified in terms of the parameters of interest. Therefore, given consistent estimates of the unrestricted parameter vector $\gamma = (\gamma_1, \gamma_2, \gamma_3)'$ and its variance-covariance matrix, the restrictions can be tested and imposed by minimum distance to obtain estimates for the restricted parameter vector $(\theta_m, \rho)'$. To estimate the unrestricted parameter vector, we apply the "extended GMM" estimator proposed by Arellano and Bover (1995). In addition to moment conditions for the first-differences transformation of the model, by which the first difference of the idiosynchratic error is orthogonal to lagged values of the RHS variables, we also exploit moment conditions for the untransformed model, by which, under certain regularity conditions, fulfill no correlation between the composity error term r_{nt} and the first differences of the lagged RHS variables. Recent empirical work have noticed that in the presence of highly persistent variables, such as sales, capital or employment, GMM estimators based only on moment conditions for the first-differences transformation of the model provide imprecise results (see Mairesse and Hall, 1996, for an application to the estimation of production functions). In this setting, the further moment conditions for the untransformed model may improve parameter identification.

4.2 Estimation of the adjustment costs parameters

After estimating the gross profit function, we can recover the profitability shock R_{nt}, so that we can treat it as an observable state variable in the estimation of the adjustment costs parameters. Let $s_{nt} = (x'_{nt}, \varepsilon'_{nt})'$ be the vector of state variables, 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 gross profit 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 inaction is the optimal decision for the n-th firm at period t, 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:

$$\pi^d(s_{nt};\theta) = \pi^d(x_{nt};\theta) + \varepsilon^d_{nt} \qquad \text{for } d = 0, 1 \tag{17}$$

where

$$\pi^1(s_{nt};\theta) = R_{nt}K_{nt}^{\theta_m} - p_{nt}K_{nt}\, i_{nt} - \frac{\theta_Q}{2}p_{nt}K_{nt}i_{nt}^2 - \theta_F K_{nt} + \varepsilon^1_{nt}$$
$$\pi^0(s_{nt};\theta) = R_{nt}K_{nt}^{\theta_m} + \varepsilon^0_{nt}$$

The unobservable state variables account for the fact that the actual expected profit, though observable to the firm, is unobservable for the econometrician. We assume that $\varepsilon^d_{nt}; d = \{0;1\}$, to be independent and identically distributed with zero mean and variance σ^2_ε.

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

$$E\left[\tilde{\pi}^d(x_{nt};\theta)\right] = \pi^d(x_{nt})' \mu(\theta) \qquad \text{for } d = 0;1 \tag{18}$$

Since the adjustment costs parameters enter this function linearly, the decomposition (18) is given by:

$$\tilde{\pi}^0(x_{nt}) = \begin{pmatrix} R_{nt}K_{nt}^{\theta_m} \\ 0 \\ 0 \end{pmatrix} \quad \tilde{\pi}^1(x_{nt}) = \begin{pmatrix} R_{nt}K_{nt}^{\theta_m} - p_{nt}K_{nt}E[i_{nt}|x_{nt};d_{nt}=1] \\ -\frac{1}{2}p_{nt}K_{nt}E[i^2_{nt}|x_{nt};d_{nt}=1] \\ -K_{nt} \end{pmatrix}$$

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

The first component of $\tilde{\pi}^1(x_t)$ accounts for the firm gross profit minus the purchase cost of new capital stock. The second and third components are related, respectively, to the quadratic and fixed adjustment costs of new capital stock.

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:

$$pdf(x_{t+1};\varepsilon_{t+1}|x_t;\varepsilon_t;d_t) = pdf(\varepsilon_{t+1}|x_{t+1})\, pdf(x_{t+1}|x_t;d_t) \tag{19}$$

This assumption has two strong implications. First, conditional on the discrete choice and the current value of the observable state variables, the future observable state variables do not depend on unobservables. Second, the existence of autocorrelated unobservable state variables, which greatly complicate the estimation of the decision problem, is discarded.

Under assumptions AS and CI and the multiplicative decomposition given by (18), the optimal discrete choice can be written as:

$$d^*_{nt} = d \iff d = \arg\max_{j=0;1}\left\{\pi^j(x_{nt})'\mu(\theta) + \varepsilon^j_{nt} + \beta EV^j(x_{nt};\theta)\right\}$$

The log-likelihood function for this problem is

$$\ln L = \sum_{n=1}^{N} \sum_{t=1}^{T_n} \sum_{d=0,1} 1\left(d^{*}_{nt} = d\right) \ln\left(\Pr\left(d^{*}_{nt} = d \mid x_{nt}\right)\right) \tag{20}$$

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

$$P^{d}(x_{nt}) = \Pr(d^{*}_{nt} = d \mid x_{nt}) =$$

$$= \Pr\left\{ d = \arg\max_{j=0,1} \left[\pi^{j}(x_{nt})'\theta_1(\theta) + \varepsilon^{j}_{nt} + \beta\, EV^{j}(x_{nt};\theta) \right] \Big| \, x_{nt} \right\} =$$

$$= \int 1\left\{ d = \arg\max_{j=0,1} \left[\pi^{j}(x_{nt})'\theta_1(\theta) + \varepsilon^{j}_{nt} + \beta\, EV^{j}(x_{nt};\theta) \right] \right\} q(d\varepsilon \mid x)$$

The conditional choice probabilities entering the log-likelihood function are written in terms of the 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 these techniques that solve the dynamic programming problem is that they are computationally very demanding. Hotz and Miller (1993) proposed an alternative estimation method, the Conditional Choice Probability (CCP) estimator, which circumvents the need of solving the fixed point problem at each iteration of the outer algorithm. Is is based on a one-to-one mapping between the normalized value functions and the conditional choice probabilities. This estimation method has been applied in Aguirregabiria (1999) and Slade (1998) for the estimation of models of inventories and price change decisions. The much lower computational cost, however, is at the expense of lesser efficiency of the estimates. In a recent work, Aguirregabiria and Mira (2002) proposed the Nested Pseudo-Likelihood estimator (NPL), which enjoys the computational advantages of the Hotz and Miller's CCP estimator, but is able to reach the efficiency provided by the class of Rust's NFXP algorithms. As it 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 Policy Iteration operator, which maps the conditional value functions into the space of the conditional choice probabilities

$$P = \Psi(P) \equiv \Lambda(\varphi(P)),$$

where $\varphi(\cdot)$ is an operator which maps a vector of conditional choice probabilities into a vector of conditional value functions, and 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^* 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 conditional 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

$$W^d(x_{nt}) = \hat{F}^d(x_{nt}) \left(I - \beta \hat{F}(x_{nt})\right)^{-1} \left(\sum_{d=0,1} \hat{P}^d(x_{nt}) * \pi^d(x_{nt}) \quad \sum_{d=0,1} \hat{P}^d(x_{nt}) * g^d(x_{nt}) \right) \tag{21}$$

$\lambda(\theta) = (\mu(\theta)' \; 1)'$; $*$ 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 $E\left[\varepsilon^d_t \mid x_t, d^*_t = d\right] = \gamma - \ln\left[P^d(x_t)\right]$, where γ is the Euler's constant. With this distributional assumption, it is straightforward from (21) 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:

$$\ell = \sum_{n=1}^{N}\sum_{t=1}^{T_n}\sum_{d=0,1} 1(d^*_{nt} = d)\ln \Psi^d_\theta(x_{nt};P) \tag{22}$$

where

$$\Psi^d_\theta(x_{nt};P) = \frac{\exp\left\{\pi^d(x_t)'\alpha(\theta) + \beta W^d(x_t)'\lambda(\theta)\right\}}{\sum_{j=0,1}\exp\{\pi^j(x_t)'\alpha(\theta) + \beta W^j(x_t)'\lambda(\theta)\}} \tag{23}$$

Once the pseudo-likelihood function is formulated, the NPL algorithm is implemented as follows. Let $\hat{F}^d$ $(d = 0,1)$ be non parametric estimates of the conditional transition probabilities. 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 $B \geq 1$, the NPL algorithm consists in the following steps:

Step 1: Represent the conditional choice value functions in terms of the conditional choice probabilities, using the Hotz and Miller's representation as in (21).

Step 2: Obtain an update of the pseudo-likelihood estimator $\hat{\theta}^{(B)}$:

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

where

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

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

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

Iterate in B until convergence in $\hat{P}$ and $\hat{\theta}$:

Aguirregabiria and Mira (2002) showed that when the NPL is initialized with consistent estimators of the vector of conditional choice probabilities, successive iterations return a sequence of estimators that include the Hotz and Miller's CCP estimator (for $B = 1$) and the Rust's NFXP estimator (when $B \to \infty$) as extreme cases. The gains in efficiency from 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 the maximum likelihood estimator.

5 Estimation results

The estimation results for the gross profit function are shown in Table 3. We have provided the estimates for the full sample, and also for small and large firms, defined here using a threshold of 200 employees. In the upper panel, we provide the extended GMM unrestricted estimates. The coefficients show the expect signs in (16) for the whole sample and for the subsamples of small and large firms. The estimates are pretty precise in all cases. As specification tests, we provide the test for second order autocorrelation, which follows a standard normal under the null of no autocorrelation in the original idiosynchratic error term, and the Sargan test of overidentifying restrictions. We do not find evidence against the specification in none of the samples. We have used as set of instruments proper lags of capital and gross profits for the first differenced equations, and the corresponding proper differences of these variables for the level equations.

The minimum distance estimates are shown in the lower panel in Table 3. They provide the estimation of the structural parameters corresponding to the market power degree and the autocorrelation in the idionsyncratic component of the profitability shock. The estimates are very precise. For the whole sample we get an estimate of the pro…t function curvature of 0.715, being this estimate 0.877 and 0.659 for small and large …rms respectively. Thus, we …nd stronger evidence of market power in the case of large …rms. However, for the subsample of small …rms, given that the parameter is quite close to 1, we point that they are relatively close to perfect competition. The estimate of market power degree in this subsample of large …rms is in accordance with the one in Cooper and Haltiwanger (2006), who use a sample of US large plants. From these estimates and taking into account the labor input share in the value added, we have computed the implied demand elasticity through (4) and the markup. For the whole sample, we get an estimate of the markup of 12.04%. For the subsample of small and large …rms, the estimated markup has been 4.41% and 15.94% respectively.

[Table 3]

Once we have constructed the pro…tability shock R_{nt} from the previous, we decompose it into an aggregate and an idiosyncratic shock $\tilde{R}_{nt}$; such that $R_{nt} = R_t \tilde{R}_{nt}$: Following Cooper and Haltiwanger (2006), the aggregate shock is simply the annual mean of the pro…tability 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 $\tilde{r}_{nt} = \ln(\tilde{R}_{nt})$. Thus, the vector of observable state variables we use in the estimation is given by $x_{nt} = (p_t; r_t; \tilde{r}_{nt}; K_{nt})$:

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 ¼$^{d}(x_{nt})$ 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 that variance.

The structural estimation results using the NPL algorithm are shown in Table 4. The discount factor ¯ has been …xed at 0.95.

[To be written]

6 Conclusions

In this paper we have estimated a dynamic structural model of irreversible investment for Spanish manufacturing …rms. Our dataset exhibits some of the characteristics reported in the recent microeconomic investment literature, with sizeable frequencies of zero and lumpy investment. Based on these facts, we have proposed a dynamic structural investment model with irreversibilities in which nonconvex components have been included in the adjustment cost function. Our speci…cation for the adjustment cost function includes both quadratic and …xed components.

In line with Cooper and Haltiwanger (2006), we have derived a dynamic programming problem of discrete choice, in which …rms choose whether to make a positive investment or to postpone investment to the future. As estimation approach, we have proceed in two stages. In the …rst stage, we have estimated the gross pro…t function and use the entailed structural parameters to construct the pro…tability shock. We then have included such pro…tability shock as an observable state variable in the discrete choice model of investment in the second stage. To estimate the adjustment cost parameters, we implement the nested pseudo-likelihood (NPL) algorithm recently proposed by Aguirregabiria and Mira (2002).

[To be written]

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.

Profits: Gross profits are computed as value added net of labor costs. They are measured in thousand euros.

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 chosen reference year was tha one that minimizes the number of missing values in the capital stock. The unit measure is thousand euros.

A2. 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.

A3. 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; $\bar{r}_{nt}$ and $\bar{k}_{nt}$ respectively. In this case, $M_1 = 2$ and $M_2 = M_3 = 7$. Let $m = 1; \ldots; M$ be the index for the cells of tridimensional state variable $x_{nt} = (r_t; \bar{r}_{nt}; \bar{k}_{nt})$; where $M = M_1 \times M_2 \times M_3 = 2 \times 7 \times 7 = 98$. Let r^c; $\bar{r}^c$ and $\bar{k}^c$ be the values of the discretized state variables and let r^m; $\bar{r}^m$ and $\bar{k}^m$ be the values of discretized state variables correspondig to the m-th cell, that is, $x^m = (r^m; \bar{r}^m, \bar{k}^m)$.

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:

$$\widehat{\Pr}(d = 1|x^m) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T} 1(d_{nt} = 1)\, K_3(x_{nt}; x^m)}{\sum_{n=1}^{N}\sum_{t=1}^{T} K_3(x_{nt}; x^m)}; \qquad \text{for } m = 1; \ldots; M$$

where K_3 is the trivariate gaussian kernel:

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

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

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

$$\widehat{\Pr}\left(\acute{r}^c_{t+1} = r^m \,\middle|\, \acute{r}^c_t = r^l\right) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T} 1\left(\acute{r}^c_{n;t+1} = r^m\right) K_1\left(\acute{r}_{nt}; r^l\right)}{\sum_{n=1}^{N}\sum_{t=1}^{T} K_1(\acute{r}_{nt}; r^l)}$$

for $m; l = 1; \ldots; M_2$; where K_1 is a univariate gaussian kernel:

$$K_1\left(\acute{r}_{nt}; r^l\right) = \frac{1}{(2\pi)^{1/2}} \exp\left\{ -\frac{1}{2}\left(\frac{\acute{r}_{nt} - r^l}{h_1}\right)^2 \right\}:$$

The capital stock is an endogenous 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:

$$\widehat{\Pr}\left(\acute{k}^c_{t+1} = \acute{k}^l \,\middle|\, x^m; d\right) = \frac{\sum_{n=1}^{N}\sum_{t=1}^{T} 1\left(\acute{k}^c_{n;t+1} = \acute{k}^l\right) 1(d_{nt} = d)\, K_3(x_{nt}; x^m)}{\sum_{n=1}^{N}\sum_{t=1}^{T} 1(d_{nt} = d)\, K_3(x_{nt}; x^m)}$$

for $d = 0; 1$; $l = 1; \ldots; M_3$ and $m = 1; \ldots M$:

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.

A4. ESTIMATION OF THE AMOUNT OF INVESTMENT IF $d = 1$

The functions $E[i_{nt} | x_{nt}; d_{nt} = 1]$ 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; $\pi^d(s_{nt})$. 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 $E[i | x^m; d = 1]$; 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:

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

for $h = 1, ..., H$ y $m = 1, ..., M$. So we have obtained the $M \times 1$ vector of estimated values of $E[i | x; d = 1]$. The $M \times 1$ vector of estimated values of $E[i^2 | x; d = 1]$ has been estimated accordingly.

References

[1] Abel, A.B. and J.C. Eberly (1994), "A Uni...ed Model of Investment Under Uncertainty", The American Economic Review 84, 1369-1384.

[2] Abel, A.B. and J.C. Eberly (1997), "An exact solution for the investment and value of a ...rm facing uncertainty, adjustment costs", Journal of Economic Dynamics and Control 21, 831-852.

[3] Aguirregabiria, V. (1999), "The dynamics of markups and inventories in retailing ...rms", Review of Economic Studies 66, 275-308.

[4] 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.

[5] Alonso-Borrego, C. and M.D. Collado (1999), "Factor demands and innovation activity: Evidence from Spanish manufacturing ...rms", Documento de Trabajo 99-02, Fundación Empresa Pública.

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

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

[8] Bertola, G. and R.J. Caballero (1990), "Kinked Adjustment Costs and Aggregate Dynamics", NBER Macroeconomics Annual 5, 237-288.

[9] Bertsekas (1976), Dynamic programming and stochastic control. Academic Press, New York.

[10] Chirinko, R.S. (1993), "Business Fixed Spending: Modeling Strategies, Empirical Results, and Policy Implications", Journal of Economic Literature 31, 1875-1911.

[11] Cooper, R. and J. Haltiwanger (2006), "On the nature of capital adjustment costs", Review of Economic Studies 73, 611-633..

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

[13] Dixit, A.K., and R.S. Pindyck (1994), Investment under Uncertainty. Princeton University Press.

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

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

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

[17] Rust, J. (1987), "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", Econometrica 55, 999-1034.

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

[19] 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.

[20] Slade, M. (1998), "Optimal pricing with costly adjustment: Evidence form retail-grocery prices", Review of Economic Studies 65, 87-107.

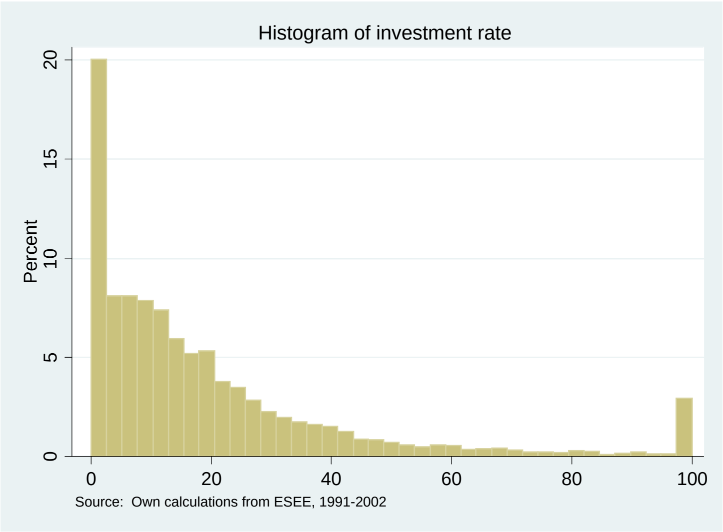

Figure 1: Empirical distribution of the investment rate

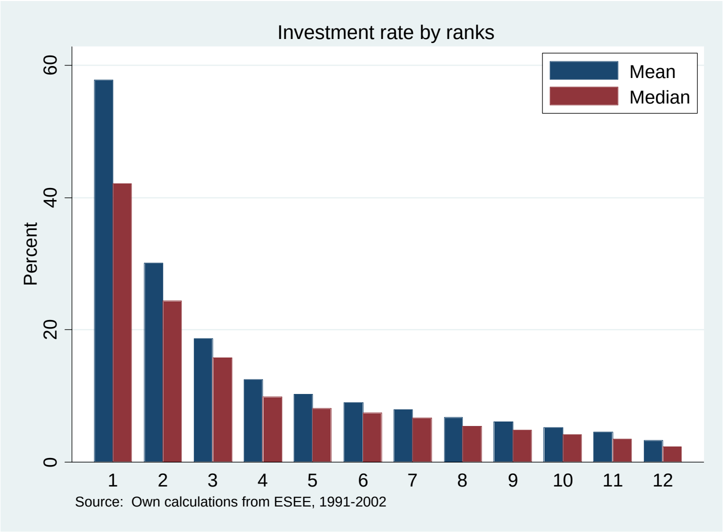

Figure 2: Investment rates, ordered by intensity

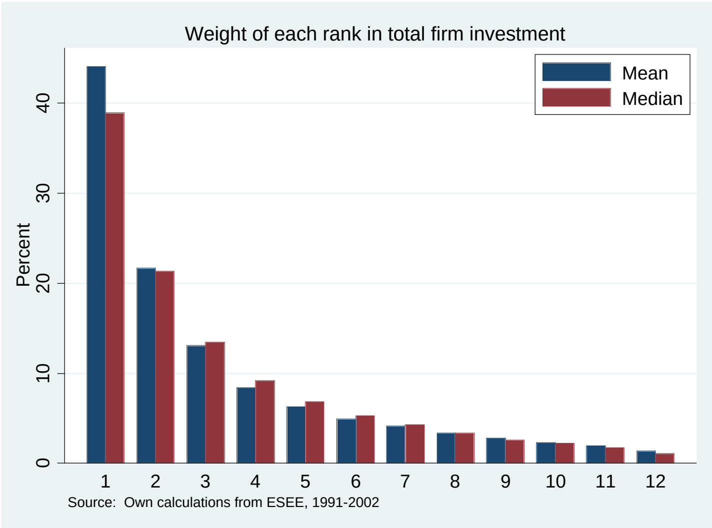

Figure 3: Relative size of each investment

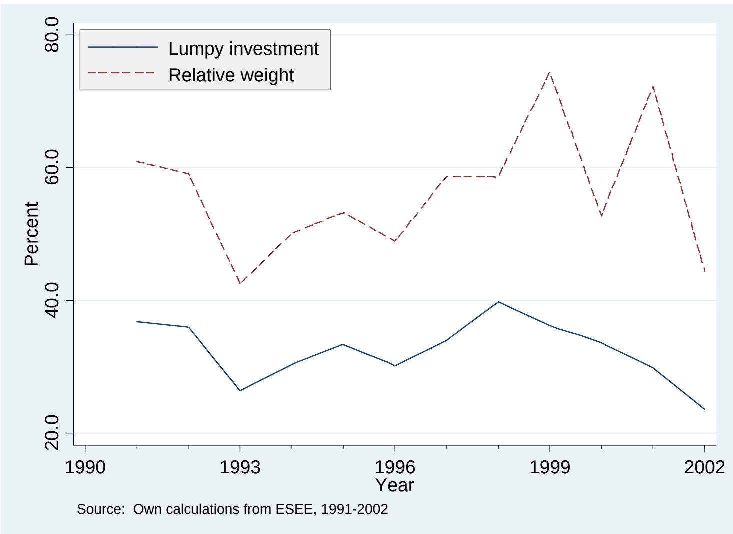

Figure 4: Relative importance of investment spikes

Year	Inaction¤	Lumpiness×
1991	14.71	36.76
1992	14.80	36.02
1993	17.41	26.39
1994	16.15	30.38
1995	15.15	33.40
1996	14.15	30.18
1997	12.61	34.08
1998	10.71	39.81
1999	10.99	36.24
2000	10.76	33.64
2001	10.00	29.86
2002	10.63	23.65
Total	13.06	32.81

¤Zero investment ×Inv. rate above 20%.

Table 1: Incidence of inaction and lumpiness by year.

Firm size	(%).	Inaction¤	Lumpiness×
Small	(49.89)	23.10	32.97
Medium	(24.41)	4.73	31.71
Large	(25.69)	1.50	33.53
Total	100	13.06	32.81

¤Zero investment ×Inv. rate above 20%.

Table 2: Incidence of inaction and lumpiness by …rm size

	Sample		
	All firms	Small firms	Large firms
GMM estimates			
k_{nt}	1.245	1.105	1.166
	(0.413)	(0.294)	(0.484)
$k_{n,t-1}$	-0.822	-0.451	-0.716
	(0.387)	(0.312)	(0.477)
$\pi_{n,t-1}$	0.386	0.220	0.287
	(0.106)	(0.024)	(0.038)
2nd order autocorr.	1.601	1.744	1.225
p-value	0.109	0.081	0.221
Sargan	38.821	35.700	38.440
p-value	0.224	0.530	0.404
Minimum distance estimates			
θ_m	0.715	0.878	0.659
	(0.094)	(0.091)	(0.145)
ρ	0.445	0.226	0.298
	(0.095)	(0.023)	(0.036)
p-value MD test	0.188	0.415	0.273
Markup	12.04%	4.41%	15.94%
Demand elasticity	-8.31	-22.66	-6.27
No. firms	1428	991	409
No. observ.	9867	6649	2782
Heteroskedasticity- robust standard errors in parentheses. Year binary dummies included to control for common shocks.			

Table 3: Estimates of the gross profit function

Structural parameter estimates (NPL algorithm)						
	1 stage	2 stages	3 stages	4 stages	5 stages	6 stages
μ_Q						
μ_F						
¾"						
LogL						
Pseudo-R^2						

Table 4: Structural parameter estimates. Standard errors in parenthesis