Fondazione Eni Enrico Mattei

Dynamic Labour Demand with Lumpy and Kinked Adjustment Costs

Paola Rota*

NOTA DI LAVORO 20.2001

APRIL 2001

ETA – Economic Theory and Applications

*University of Modena and Reggio Emilia, Italy

Fondazione Eni Enrico Mattei
Corso Magenta, 63, 20123 Milano, tel. +39/02/52036934 – fax +39/02/52036946
E-mail: letter@feem.it
C.F. 97080600154

Dynamic Labour Demand with Lumpy and Kinked Adjustment Costs

Paola Rota

Universita' di Modena e Reggio Emilia

January 1997
Revised September 1998
Second Revision July 2000

Abstract

We analyze the dynamics of firms' employment decisions which underlie lumpy and kinked adjustment costs. We consider a dynamic structural model in which, in each period, firms face a choice of whether to vary the labour input or to postpone the adjustment to the future. By exploiting the first order condition for optimality, we derive a semi-reduced form in which firms' intertemporal employment are defined by a standard static marginal productivity condition augmented by a forward-looking term. In this way we obtain a marginal productivity equilibrium relation which takes into account the future alternatives of adjustment or non-adjustment that firms face as the result of the presence of fixed and linear adjustment costs. Linear costs amount to 35% of average labour costs and fixed costs are estimated to be about 3.65 times average unit labour costs.

Keywords: adjustment costs, dynamic labour demand, discrete decision process.
JEL: J32 C23

I would like to thank Orazio Attanasio, Richard Blundell, Alessandra Del Boca, Christopher Gilbert, Daniel Hamermesh and Costas Meghir. I gratefully acknowledge financial support from the Economic & Social Research Council; ESRC award ref. H53627501795. All errors, however, remain my responsibility.

Address for correspondence:
Corso Canalgrande 14
41100 Modena - Italy
email: p.rota@pianeta.it

1. Introduction

Despite a period of real wage moderation in Europe, the response of employment has been modest. Adjustment costs have been claimed to be responsible for the great persistence in unemployment: a phenomenon defined by some economists as "Eurosclerosis" (see Bentolila and Bertola, 1990). Firms face costs if they choose to adjust their staffing levels. During recessions firms hoard labour and during expansions they may operate to some extent understaffed in order to avoid the cost of adapting employment to the level they would attain in the absence of these costs. Both economists and employers acknowledge that the non-wage component of the costs of varying the labour input has played an increasingly important role in determining the temporal pattern of labour demand in response to exogenous shocks. Labour adjustment costs affect the timing and the extent of the employment variation in response to shocks. The form of the constraint that they impose on optimal employment decisions is crucial in understanding firms' behaviour over time. In particular, the presence of fixed and linear costs in hiring and firing can give very different adjustment paths from those resulting from the standard assumption of strictly convex (generally quadratic) costs.

A rapidly growing literature has considered forms of adjustment costs alternative to the standard quadratic case and has examined the possibility of non-smooth or asymmetric employment adjustment paths - see Nickell (1986), Hamermesh (1993) and Hamermesh and Pfann (1994) for surveys. These developments have involved positing either linear costs and, typically, asymmetries in hiring and firing, or fixed adjustment costs and lumpiness. See, for example, Bentolila and Bertola (1990) and Aguirregabiria and Alonso-Borrego (1999) for the former case; and Hamermesh (1989) and Caballero, Engel and Haltiwanger (1997) for the latter case. These last two studies show how reference to the plant or firm-level structure of adjustment is important if aggregate dynamics are to be correctly understood. In particular, the effects of heterogeneity, idiosyncratic uncertainty and lack of coordination across firms or plants may result in highly nonlinear behaviour in aggregate time series (see also Bertola and Caballero, 1990 and Caballero and Engel, 1993).

In this paper we analyze firms' intertemporal employment decisions in the presence of both lumpy and kinked adjustment costs and emphasize the importance of non-convexities for modelling and estimating dynamic labour demand. Since the optimizing response in the presence of non strictly-convex adjustment costs may be inaction, there is the possibility of corner solutions in the demand for labour. In this case, the standard marginal conditions for optimality given by the Euler equation fail to hold. In our model we explicitly take into account the existence of corner solutions by considering a discrete-time-discrete-choice dynamic structural model. While discrete time is obviously a characteristic of firm level data, discrete choice also represents the natural framework within which to analyze firms' intermittent adjustment. The use of the information contained in the optimal discrete choice faced by the firm allows us to characterize the adjustment cost function. Our focus is empirical. Previous studies on labour demand which have adopted a continuous-time and continuous-state-spaces framework have obtained important analytical results which prove the optimality of (S,s) policies, but, due to the difficulty of obtaining closed-form solutions, direct estimation of the structural parameters has not been viable (see for instance Bentolila and Bertola, 1990). We utilize a two stage procedure within a more general framework. At the first stage, we estimate a first order condition augmented by forward-looking terms. These estimates allow us to test for the presence of both linear and fixed adjustment costs, and to quantify the linear costs. We are then in a position to use these estimates to develop a structural model of the employment adjustment decision which in principle allows identification of the fixed cost component of the adjustment costs.

Our framework makes Hamermesh's (1989) fixed cost model dynamic. In that model, the instantaneous jump to the desired level of employment and the implied discontinuity in the demand for labour result in the optimal decision problem being perceived as static rather than dynamic. This is because we are left with a (static) marginal productivity type of relation since the term which represents adjustment costs drops from the first order conditions, in contrast to the case of strictly convex costs which generate the familiar Euler equation. Labour demand is governed by a threshold rule: either the firm decides not to vary the labour input or it chooses to completely adjust the input to the long run level corresponding to the (static) marginal productivity relationship and the dynamics implied by the presence of fixed adjustment costs seem to disappear.

We estimate the structural parameters using a panel dataset of more than 3000 Italian firms. Italy has often been regarded as an extreme case in which the regulatory burden of hiring and firing has imposed a major constraint on employers' decisions with regard to their desired level of employment (Emerson, 1988; Del Boca and Rota, 1998). Our emphasis is on adjustment costs relating to the number of jobs and on the effects of adjustment costs on firms' decisions as to whether to expand, contract or maintain their size unaltered.

We proceed in two stages:

1. We estimate the parameters of the production function and linear cost function, by exploiting the first order conditions for optimality, we derive a semi-reduced form and define a (dynamic) marginal productivity equilibrium relation which takes into account the future alternatives that firms face of adjustment or non-adjustment, due to the presence of fixed and linear costs. The important new result is that, conditional on having adjusted employment between *t*-1 and *t*, the marginal revenue product of labour will differ from the wage the firm pays by a forward looking term which embodies the discrete choices implied by the cost structure. We show that this term may be estimated non-parametrically. This procedure allows us to exploit standard orthogonality conditions in estimation and it may also provide a way of properly estimating intertemporal relations for a variety of other problems (eg investment demand and inventory adjustment).
2. We exploit the structural parameters obtained in part 1 in the estimation of the discrete choice structural model. Here, we use a method similar to Hotz and Miller (1993) and Aguirregabiria (1999). Our procedure is simplified by using the parameters estimated in the dynamic marginal productivity equation.

Our estimates support the view that adjustment costs in Italy are high and non convex, and in particular, linear costs amount to 35% of average unit labour costs. We find also clear evidence for fixed costs which amount to 3.65 times average unit labour costs. Simulation analysis indicates that linear costs are responsible for the largest proportion of non-adjusting firms, but these results are qualified by the presence of a substantial residual component.

The paper is organized as the following. In Section 2 we illustrate the dataset. Section 3 we present the dynamic stochastic discrete choice model of labour demand. In Section 4 we obtain an empirical representation of the optimal decision rule. In Section 5 we describe the empirical model, derive the first order condition and obtain a dynamic marginal productivity condition which allows us to take into account the choice of non adjusting the firm face in the presence of non-convex adjustment costs. In Section 6 we estimate the production function parameter and the linear cost function. In Section 7 we estimate the dynamic discrete choice structural model and discuss identification issues. Conclusions follow in Section 8.

2. The dataset

The dataset, which we use to estimate the structural model, was drawn from the Centrale dei Bilanci databank of company accounts. It contains information on companies located in the highly industrialized district of Lombardia and which in 1982 had less than 500 employees. The data cover 3247 manufacturing companies for the period 1982-1989. Relatively few of these firms are publicly quoted, from 20 to 25 throughout the period, reflecting their predominant small-size. The configuration of firms by number of employees in the starting year is reported in Table I.

Table I Configuration of firms by number of employees in 1982

No. of employees	1-19	20-49	50-99	100-199	200-500
Frequency	494	1125	827	519	288

Small-sized firms (1-49 employees) account for nearly 50 per cent of the sample. To obtain a first approximation to the process of employment adjustment we consider the distribution of the rate of change in employment each year, starting from 1982, in relation to the rate of change in sales and wages. A number of features suggest the existence of some non-convexity in the cost of adjusting labour. First, changes in employment show a recurring pattern throughout the sample period. In Fig. 1 we report the percentage rate of change in 1982-1983; the distributions are similar in subsequent years.

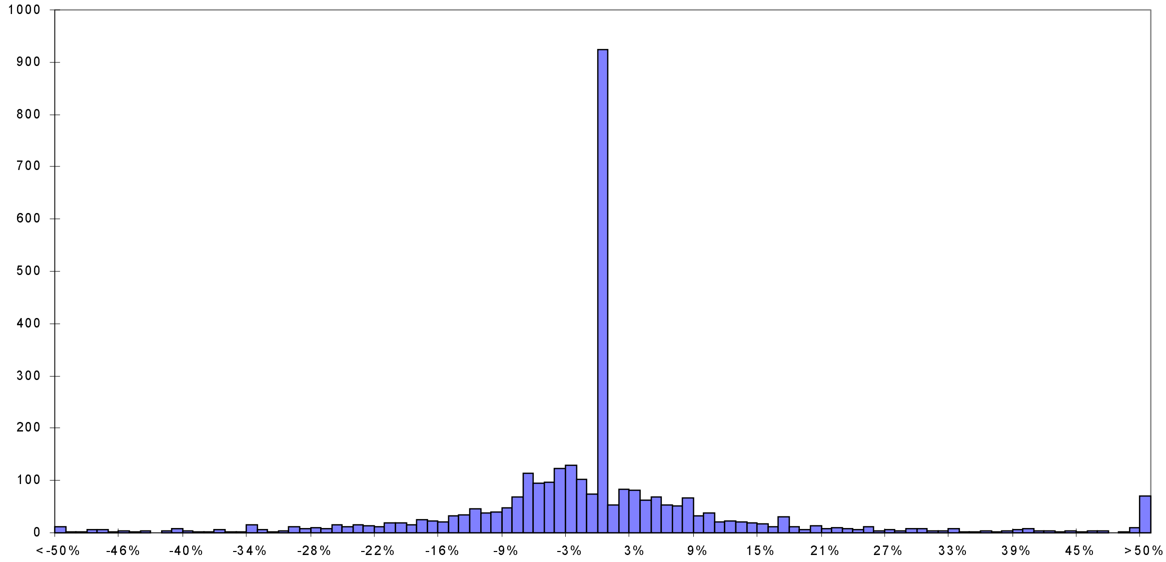

Figura 1 Rate of change in employment: 1982-1983

The very high spike at zero shown, in the graphs and recorded reveals a considerable stickiness in employment. In Table II, the first column shows the number of firms which did not adjust each year and the second column the sample size.

Table II Number of firms with no change in employment

year	Firms with ΔL_t=0	Total no of firms
1982-83	913	3253
1983-84	640	3258
1985-85	717	3258
1985-86	735	3260
1986-87	662	3263
1987-88	648	3263
1988-89	587	3266

Throughout the period, the rate of change in employment was zero on average for more than 20 per cent of the firms. This is confirmed by Gavosto and Sestito (1994) who aggregate monthly employment data and find a very similar proportion of firms not changing the employment from one year to another. There is some evidence of a higher frequency of small decreases in employment relative to small increases. By contrast the rate of change in sales and in wages is characterized by a standard bell-shaped distribution. The distribution of the percentage change in output and in the wage rate, in year 1982-1983 are shown in Figures 2 and 3 respectively where

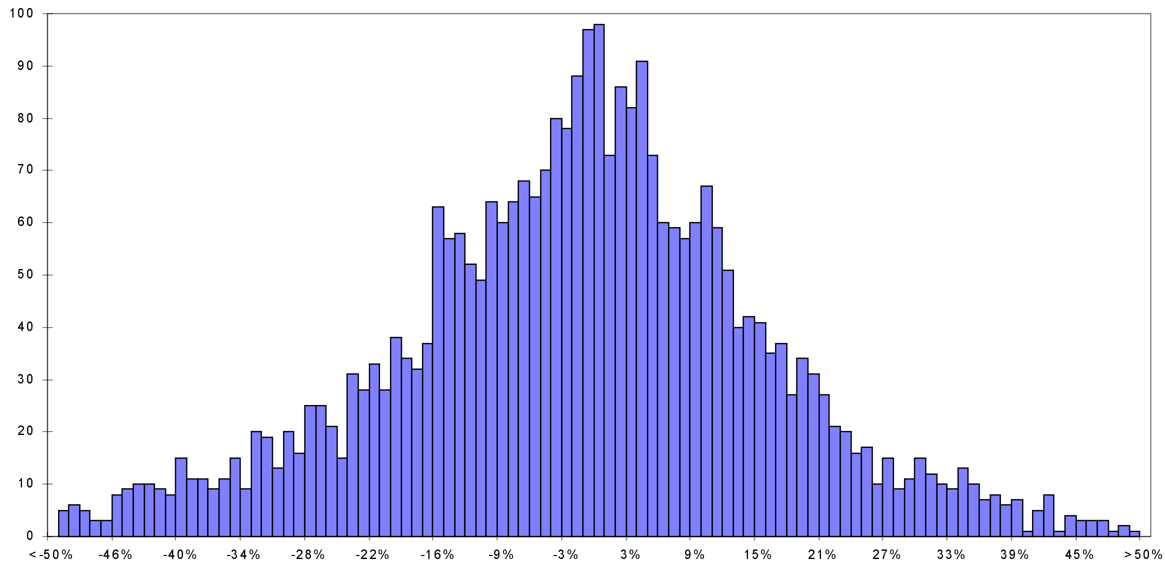

Figure 2 Percentage rate of change in real output

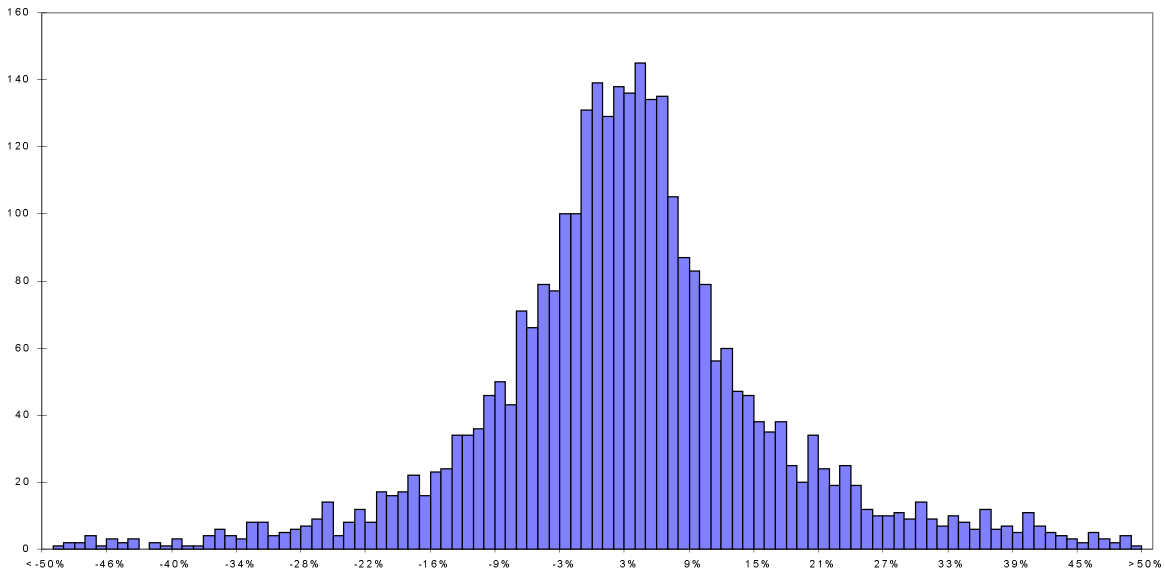

Figure 3 Percentage rate of change in real wage

the histograms omit changes of less than -50% and greater than 50%. A similar pattern is evident throughout the sample period. Our focus is on firms' decisions of whether to expand or to contract, or to maintain the number of employees unaltered. This is particularly important in a dynamic context where adjustment costs may affect a firm's optimal size. However we are aware that hiring and firing costs may also affect job and worker reallocation. In this context, zero changes in employment may represent the net sum of people who have quit, together with their replacements. (Note that in Figure 1, negative changes include dismissals as well as voluntary quits and retirements.) In some cases, these may coincidentally cancel leaving an unintended zero net-change, while in other they may cancel to give a planned net zero adjustment. We consider the former case as reflecting measurement errors, which we will model explicitly as imperfect knowledge on behalf of the econometrician. By contrast, we view the latter possibility, as outcome of the firm's optimal decision problem.

It is also possible that apparent non-adjustment arises from reporting deficiencies and does not represent actual employment changes. Without independent measures of employment we cannot be completely confident that this is not the case. We checked whether there is an association between round numbers of numbers (for instance, multiples of five) and zero changes. We did not find any recurrent pattern; indeed zero adjustment occurs randomly throughout the sample.

3. The Dynamic Model

In a discrete time, infinite horizon framework, the firm's intertemporal problem is

$$max_{c,d} E\left\{\sum_t \beta^t \Pi\left[d_t, c_t, s_t\right]\right\} \tag{1}$$

where $\beta \in (0,1)$ is a discount factor; $d_t \equiv d(s_t)$ is a two-component vector of mutually exclusive discrete choices: A if the firm decides to adjust employment and NA, if the firm decides not to vary the stock of employees at time t; $c(s_t)$ is an M-dimensional vector of continuous choices made at time t; $s_t=[L_{t-1}, W_t, \omega_t]$ is the vector of the state variables, respectively beginning of period stock of workers and current wages, W_t, and ω_t is an idiosyncratic productivity shock.[1] In order to keep notation simple we consider a single firm and hence omit the subscript i, i=1,...,N. The firm knows the stock of workers at the beginning of the period and current period wage rates, but it is uncertain about the future wage. The exogenous state variables, the wage and the productivity shock, evolve stochastically according to a stationary first order Markov process with transition

[1] In principle we could also add an aggregate shock, but the length of the time series in our dataset does non allow us to consider this type of shock.

probability $p_W(W_{t+1}=W'|W_t=W)$ and $p_\omega(\omega_{t+1}=\omega'|\omega_t=\omega)$ respectively. We assume that W_t varies across firms because of workers' differing skill (human capital) levels. Π(.) indicates a time-additive, separable, one-period profit function which may be specified as

$$\Pi_t = \left\{ A L_t^{\alpha} - b\,|\,\Delta L_t\,| - K\,1\,(L_t \neq L_{t-1}) - W_t L_t \right\} \qquad \textbf{(2)}$$

where A>0, 0<α<1, b>0 is the cost per unit of labour, the absolute value indicates that linear costs have a kink at $\Delta L_t=0$, K represents the fixed cost and $1(L_t \neq L_{t-1})$ is an indicator function. When the firm adjusts employment it pays both kinked and lump-sum costs. We assume here that adjustment costs are symmetric[2].

The assumptions of infinite time horizon, time-separability of the profit function, $\beta \in (0,1)$ and stationary Markov transition probabilities characterizing the motion of wages guarantee the stationarity of the decision rule and of the optimal value function (Blackwell, 1965). We may thus omit the time subscript in the following Bellman equation:

$$V(s) = \max_{d,c} V^d(c,s) = \max_{d,c} \left\{ \left[\Pi^d(c,s) + \beta \int V(s')p(ds'\,|\,c,s) \right] \right\} \qquad \textbf{(3)}$$

where d=(A, NA) is the discrete choice relating to the decision of adjusting and not-adjusting, respectively; $c(s_t)$ indicates the continuous decision ie, conditionally on $d(s_t)$=A, the firm chooses how many workers to hire/fire (otherwise it leaves employment at the previous period level); the prime indicates variables that are unknown to the firm at the time of the decision and p(.) is the Markov transition probability which stochastically determines next period's state s' as a function of the current state vector and the current choices (d,c).

The presence of both a continuous and a discrete control makes this problem non-standard. We follow Pakes (1994) and Miller (1997) and show that the problem may be reduced to a standard discrete choice problem. Conditional on the discrete choice, the continuous control implies that a maximum is reached. For each discrete decision d=(A, NA), the firm chooses c to maximise

---

[2] The role of asymmetries has been discussed in various studies. However, those studies assume only linear costs (or strictly convex in some case) whereas we adopt a more complicated adjustment cost structure. Since our focus is the identification of linear and fixed adjustment costs, we use symmetry as simplifying assumption. Asymmetric adjustment costs are considered in Rota (1997).

$$\left[\Pi^{d}(c,s)+\beta\int V(s')p^{d}(ds'|c,s)\right] \quad \mathbf{(4)}$$

Let c_d be the solution to maximizing (4) and successively substitute it to the d-subproblems into current profits and transition probability. Then we obtain reduced form profits $\Pi^d(s_t)\equiv\Pi^d(c^d,s)$ and transition probabilities $p^d(s'|s)\equiv p^d(s'|c^d,s)$.The optimal discrete choice d^* satisfies: [3]

$$max_{d}\left\{\Pi^{d}(s)+\beta\int V(s')p^{d}(ds'|s)\right\} \quad \mathbf{(5)}$$

over d=A,NA and we obtain a discrete choice problem where the control variable is restricted to a countable set of alternatives (see Eckstein and Wolpin, 1989 and Miller, 1997, for a survey on the stochastic discrete choice models). The firm chooses to adjust its labour stock if its value given the optimal adjustment exceeds its value with no adjustment. Equation (5) represents the starting point of our representation of firms' intertemporal employment decisions.

4. Optimal Decision Rule

At the beginning of period t, the firm chooses the profit maximizing level of employment. It observes the number of workers, the current wage and the productivity shock, but it is uncertain about future wages and productivity shocks.

Following the literature on discrete decision processes, we introduce an error term which

---

[3] The proof follows Miller (1997, p. 256). Consider profit as a function of the two state variables, L_{t-1} and W_t and write the transition probability as $h^d(s_{t+1}|s_t)$, assumed to be differentiable almost everywhere. Set $c_t = L_t$ to obtain the following first order conditions:

$$\frac{\partial V(L_t,s_t)}{\partial L_t}=\frac{\partial \Pi_d(L_t,s_t)}{\partial L_t}+\beta\int\frac{\partial V(L_t,s_t)}{\partial L_t}h_d(s_{t+1}|s_t)ds_{t+1}+$$
$$\beta\int V(s_{t+1})\frac{\partial h_d(s_{t+1}|s_t)}{\partial L_t}dW_{t+1}=0$$

where the optimal choice is to adjust. The continuous choice has two dynamic effects: it directly affects future profits, since outcomes of the past endogenous choices form part of the state variable and it may affect the transition probability. Substituting the optimal decision rule into the intertemporal profit function we obtain the mapping (5).

captures the lack of perfect information, from the viewpoint of the econometrician, on the state variables relevant for the decision process. Lack of perfect knowledge on the part of the econometrician may be the consequence of difficulties in obtaining exact information in large datasets: prohibitive costs, inaccessibility and imperfect monitoring - no dataset is sufficiently rich to fully measure all the characteristics of a firm.[4] In our context, wages are imperfectly observed by the econometrician due to lack of information on the costs of monitoring workers, and heterogeneity with respect to human capital. We thus allow the state vector, s, to comprise an unobservable component, ϵ, which captures the heterogeneity in both the outcome paths and the employment choices observed in the data and define it as:

$$s \equiv [L_{t-1}, W_t, \omega_t, \epsilon_t] \quad \textbf{(6)}$$

where L_{t-1} is previous period's employment.

We denote the vector of observables as

$$x \equiv [L_{t-1}, W_t, \omega_t] \quad \textbf{(7)}$$

We make the following assumption:

Assumption 1.

The profit function is additively separable and may be written as

$$\Pi(s,d) = E[\Pi(x,d)] + \epsilon(d) \quad \textbf{(8)}$$

where the unobserved state variable is a vector with at least as many components as the number of alternative choices (McFadden, 1981; Rust, 1987) [5]

$$\epsilon(d) = \left[\epsilon^{A}, \epsilon^{NA}\right] \quad \textbf{(9)}$$

---

[4] Miller (1997, p. 259) explains very well the role of unobserved state variables to avoid obtaining a deterministic decision rule.

[5] As in static discrete choice, a sufficient condition for obtaining a saturated decision rule is that:

$$\delta(x,\epsilon) = argmax_{d \in D(x)}\left[V(x,d) + \epsilon(d)\right]$$

where $\delta(.)$ is the optimal decision rule and D(x) is the choice set. (See McFadden, 1981 for the proof and Rust, 1991 for a discussion)

Set $E[\Pi^d(x)] \equiv \pi^d(x)$, the one-period profit, conditional on the observables x and the choice d, may be partitioned into two components: an observable component π^A or π^{NA} and a non observable component, ϵ^A or ϵ^{NA}:

$$\Pi^d(s) = \begin{cases} \pi^A(x) + \epsilon^A & if \quad d = A \\ \pi^{NA}(x) + \epsilon^{NA} & if \quad d = NA \end{cases} \qquad \textbf{(10)}$$

where ϵ^d and x are orthogonal. Since our problem is dynamic we need to introduce the following assumption about serial correlation of the ϵ's:

Assumption 2.

Problem (5) satisfies "conditional independence" (Rust, 1987,1991,1992) according to which, conditional on x, the unobservables ϵ are independent of their past values:

$$Pr\{dx', d\epsilon' \mid x, \epsilon, d\} = Pr\{\epsilon' \mid x'\} Pr\{x' \mid x, d\} \qquad \textbf{(11)}$$

where Pr{x' /x,d) is the transition probability of state x' given state x and decision d.

`

Assumption 2 restricts the role of unobservables, implying that any serial dependence between ϵ and is transmitted entirely through *x'*, which becomes a sufficient statistic for ϵ'

$$Pr\{\epsilon' | x', x, \epsilon, d\} = Pr(\epsilon' | x') \qquad \textbf{(12)}$$

The second term on the right hand side of (11) restricts the law of motion of the observables. Conditional on the discrete choice, d, and on the observables, the next period stock of workers does not depend on ϵ'. The only way the unobservable can affect the stock of worker is through the decision the firm makes:

$$Pr\{x' | x, \epsilon, d\} = p(x' | x, d) \qquad \textbf{(13)}$$

Assumption 2 makes the problem tractable avoiding an increasingly expanding state space, in particular together with assumption 1 allows us to integrate out the unobservables from the next period value function.

Thus, Bellman's equation takes the following form:

$$V(s) = \max_d \left\{ \pi^d(x) + \epsilon^d + \beta \int [V(s') \mid x,d] pr(d\epsilon' \mid x') p(dx' \mid x,d) \right\} \quad \textbf{(14)}$$

Under Assumption 2, we may express the value function and the conditional value functions as expectations which only depend on the observable state variables:

$$\begin{aligned} &v^A(x) \equiv E[V(s) \mid x,d^*(s)=A] \\ and \quad &v^{NA}(x) \equiv E[V(s) \mid x,d^*(s)=NA] \end{aligned} \quad \textbf{(15)}$$

This allows us to define the conditional value function, ie the value function after having optimally decided to adjust as:

$$\begin{aligned} v^A(x) &= \pi^A(x) + \epsilon^A + \beta E\overline{v^A}(x') \\ &= \pi^A(x) + \epsilon^A + \beta \int [P^A v^A(x') + P^{NA} v^{NA}(x')] p(x' | x,d^*=A) \\ where \quad \overline{v^A} &= max_j \{ v^j [(x')|x,d^*=A] \} \end{aligned} \quad \textbf{(16)}$$

and where j=A,NA, d^*=A, indicates the optimal choice and $P^A(x')$ and its complement $P^{NA}(x')$ are the conditional choice probabilities. The conditional value function for decision NA is defined symmetrically. We have now expressed the value function in terms of the observables, as the weighted sum of the expected value functions, conditional on having optimally taken one of the two mutually exclusive choices; weights are given by the conditional choice probabilities.

The optimal decision takes the following form:

$$d^* = \begin{cases} NA \;\; if \;\; [\epsilon^{NA} - \epsilon^A] > \left[\pi^A(x) - \pi^{NA}(x) + \beta E\left\{\overline{v^A}(x') - \overline{v^{NA}}(x')\right\}\right] \\ A \;\; if \;\; [\epsilon^{NA} - \epsilon^A] \leq \left[\pi^A(x) - \pi^{NA}(x) + \beta E\left\{\overline{v^A}(x') - \overline{v^{NA}}(x')\right\}\right] \end{cases} \quad \textbf{(17)}$$

The optimal decision depends on the differences in the expected profits associated with the two alternative choices. For instance, the firm will decide to adjust if the difference between the choice specific stochastic terms (ϵ^A-ϵ^{NA}) is less than the difference between profits (current and expected), and it will decide not to adjust if the reverse inequality holds.

5. The Empirical Model

Rather than directly solve the Bellman equation, we reduce this complex intertemporal model to a directly estimable form following the method suggested by Hotz and Miller (1993). We may now express the second term on the right hand side of equation (16) as the following:

$$\begin{aligned} E\overline{v^{A}}(x') &= P^{A}(x')v^{A}(x') + \left(1 - P^{A}(x')\right)v^{NA}(x) \\ &= v^{A}(x') + \left(1 - P^{A}(x')\right)\left(v^{NA}(x') - v^{A}(x')\right) \end{aligned} \tag{18}$$

Then the following applies:

Proposition *(Hotz and Miller, 1993)*
The difference in the conditional value functions may be expressed as a mapping from the conditional choice probabilities.

The conditional probability is given as

$$\begin{aligned} P^{A} &= \int_{\epsilon^{A'}=-\infty}^{\infty} \int_{\epsilon^{NA'}=-\infty}^{\epsilon^{A'}+\Delta v} dg\left(\epsilon^{A'}, \epsilon^{NA'}\right) = Q(\Delta v) \\ & \textit{where} \quad \Delta(v) = \left(v^{NA}(x') - v^{A}(x')\right) \end{aligned} \tag{19}$$

where g is the choice specific error joint density and Q(.) is a real valued, invertible function such that $\Delta v = Q^{-1}(P^{A})$. In the special case that ϵ^{A} and ϵ^{NA} follow a type 1 extreme value distribution, the difference in the conditional valuation functions, Δv, reduces to a log-odds transformation which we denote as q:

$$\Delta v(x') = Q^{-1}\left[P^{A}(x')\right] \equiv \gamma_t q(x') \equiv \gamma_t log\left[\frac{1 - P^{A}(x')}{P^{A}(x')}\right] \tag{20}$$

where γ_t is a constant of proportionality.[6] Conditional on the optimal decision to adjust, $d^{*}=A$, the value function may be expressed in terms of the continuous control L_t as:

[6] We assume γ_t is constant across firms, but that it may vary over time as the result of heteroscedasticity. We have not attempted a conditional representation for the γ process.

$$v^A(L_t,x_t) = \Pi^A(L_t,x_t) + \epsilon_t^A + \beta E\left[\left(v^A(x_{t+1}) + P^{NA}(x_{t+1})\gamma_{t+1} q(x_{t+1})\right) \mid x_t, d_t^* = A\right] \qquad \textbf{(21)}$$

where we have reintroduced the time subscript to focus on the arguments of each term in the value function and x_t=[L_{t-1}, W_t, ω_t]. The term $P^{NA}(.)\gamma_{t+1}$ q(.) embodies all the alternatives the firm faces in the future between adjusting and not. We may take the first order condition with respect to L_t and, given consistent estimates of the future conditional choice probabilities, we may directly estimate the marginal productivity condition. The first order condition is therefore:

$$\frac{\partial v_t^A}{\partial L_t} = \frac{\partial \Pi_t^A}{\partial L_t} + \beta E(Z_{t+1}) + \beta\gamma_{t+1} E\left[\frac{\partial P_{t+1}^{NA}}{\partial L_t} q_{t+1} + P_{t+1}^{NA}\frac{\partial q_{t+1}}{\partial L_t}\right] = 0 \qquad \textbf{(22)}$$

where Z_t=sign(ΔL_t). Here, the first term is standard and relates to increased profits from the current period adjustment, but this is augmented by the square-bracketed term which relates to the consequences of maintaining this new level of employment in subsequent periods, as the result of lumpy and linear adjustment costs. The second component on the right hand side is an indicator variable which takes the sign + (-) if hiring (hiring) is expected to occur in period t+1.

Using the profit function specified in (2), we obtain the following marginal productivity condition:

$$\left(\frac{Q}{L}\right)_t = \psi_1 W_t + \psi_2\left[Z_t - \beta E(Z_{t+1})\right] + \psi_3 (pq)_{t+1} + \nu_{t+1} \qquad \textbf{(23)}$$

where

$$(pq)_{t+1} = \left[\frac{\partial p_{t+1}^{NA}}{\partial L_t} q_{t+1} + p_{t+1}^{NA}\frac{\partial q_{t+1}}{\partial L_t}\right] \qquad \textbf{(24)}$$

and ν_{t+1} is a realization error. The structural parameters are related to the coefficients of (23) by

$$\psi_1 = \frac{1}{\alpha} \;; \qquad \psi_2 = \frac{b}{\alpha} \;; \qquad \psi_3 = -\frac{\beta\gamma}{\alpha} \qquad \textbf{(25)}$$

Equation (23) represents a simple marginal productivity relation augmented by the forward-looking term $(pq)_{t+1}$ which captures the alternatives of future adjustment or non-adjustment arising from the presence of fixed and kinked adjustment costs. The indicator variable Z_t arises from the presence of linear costs.

Suppose that the firm increases (decreases) employment in period t but, because of high adjustment costs, expects not to adjust in the subsequent period ($P_{t+1}^{NA} > P_{t+1}^{A}$). Thus P_{t+1}^{NA} is high and q_{t+1} positive. Suppose that at the beginning of period t+1, L_t is above the desired level, then the derivatives of both P^{NA}_{t+1} and q_{t+1} with respect to L_t will typically be positive. This implies that the term $(pq)_{t+1}$ will be positive. The firm expects to reduce employment in the future but, at present, L_t is higher than it would have obtained in the absence of adjustment costs, and marginal productivity is low relative to the wage rate. In the opposite case, the firm intends to hire new employees in the future, but at present L_t is below the desired level and marginal productivity is high with respect to the wage rate (the term $(pq)_{t+1}$ is negative). The term $(pq)_{t+1}$, therefore allows us to take into account the future choices the company faces subsequent to adjustment at t.

Our approach extends the Euler equation methodology to a world in which adjustment costs are linear and/or lumpy. When adjustment costs introduce the possibility of corner solutions, the standard marginal conditions for optimality given by the Euler equation fail to hold. [7] Despite this, Pakes (1994) showed that the presence of both discrete and continuous control variables does not prevent derivation of stochastic Euler equations. His method involves selection of only those observations associated with adjustment - ie he considered the set of changes irrespective of the time between adjustment. This allowed him to obtain a random sample of interior solutions and exploit the moment conditions to obtain the Hansen and Singleton-type estimator (Hansen and Singleton, 1982). The mixed discrete and continuous control approach to Euler equation has several limitations - see Aguirregabiria (1997). In our case, the censored observations represent a large part of the sample and the duration between two adjustment periods may be quite long with respect to the sample time dimension. Thus, selecting only the adjustment periods causes a considerable loss of information and introduces a potentially serious sample selection bias. Moreover, the fixed cost parameter does not enter the first order condition directly and it is

[7] In Rota (1994) we assumed quadratic adjustment costs and analysed the Euler equations characterizing firms' adjustment processes. Estimation of this model demonstrated that the type of dynamics implied by strictly convex adjustment costs, i.e. smooth continuous adjustment, is strongly rejected by the data. Moreover, the near coincidence of the long run and short run elasticities of employment to wages suggested that the dynamics of labour market adjustment cannot be satisfactorily explained within the quadratic costs model. This supports the hypothesis that non-convex costs may be present.

therefore not possible to identify the effects of fixed costs. In order to properly take lumpy adjustment costs into account, we need to model the information contained in the discontinuous adjustment process and consider the discrete decision explicitly.

6. Estimation of the First Order Condition

We estimate equation (23) using a two stage procedure; we first estimate the conditional choice probabilities P_{t+1}^{A} and P_{t+1}^{NA} and the slopes $\partial P_{t+1}^{A}/\partial L_t$ non-parametrically, using a kernel method, and construct the term $(pq)_{t+1}$. The estimation method is described in Appendix A.1.

In the second stage we use GMM to estimate the augmented marginal productivity equation (23), by conditioning the model on adjustment occurring at time t. We thus consider the sub-sample of interior solutions, which comprises the firms which have optimally adjusted at time t. For this purpose, we adopt the following selection window: we keep or drop the observation on L_t according to whether it is different from or equal to L_{t-1}. This is consistent with the theoretical model which requires conditioning on the sample and it differs from conditioning estimation on the dynamic adjustment followed by companies, which would give rise to potential sample selection bias. In this case, selecting firms which have adjusted will not affect the error term in the levels equation (23), since the theoretical model is conditional on the decision to adjust. In other words, on the null hypothesis of a correctly specified model, the selection process is not endogenous and we do not need to add any correction for selectivity. Nevertheless, we will test for non-random selectivity in estimation. If we restrict our attention to a model without linear costs, errors, due to imperfect knowledge on the state variables on the part of the econometrician, would enter the model only through the term $(pq)_{t+1}$. Hence, this would rule out, at least in principle, the possibility of fixed effects in the empirical equation. The absence of fixed effects, in that formulation, is a direct consequence of the conditional independence assumption. In the actual estimated equation, see equation (26) below, there is a second source of error which arises from replacement of an expectational term by its realization. If firms form expectations rationally, the expected value of this error term will be zero for each firm. Again there should be no fixed effects. We test for the presence of fixed effects.

The theory characterizing the discrete decision process defines the sources of stochastic disturbance in the model. First, the error term ν_{t+1} reflects a realization error due to the use of actual values of variables t+1 as an approximation to their expected values in the calculation of $(pq)_{t+1}$. The treatment of this disturbance term as purely innovational implies that this error is only correlated with the variables dated t+1 or later. If this were the only source of disturbance in the model, we would be able to use as all the available history relating to period t and earlier in the levels equation instruments. However, with linear costs, the indicator variable Z_t is jointly determined with productivity and we are restricted to using instruments dated t-1 or earlier. Rust's assumption that any serial dependence between ϵ_t and ϵ_{t+1} is transmitted entirely through the observed state L_{t+1}, thereby limiting the pattern of dependence in the $\{L_t,\epsilon_t\}$ process, implies that the probability density function of L_{t+1} depends only on last period's stock L_t and not on the error ϵ_t. If this assumption is valid it follows that the error term is serially independent in the conditional representation (23) and that instruments dated t and earlier remain valid. We use as instruments: employment, output, wages, wage-bill, new investment in machinery, gross profits, firm size, the absolute value of the change in output and the absolute value of the change in employment. [8] We estimate the following empirical specification [9]:

$$\left(\frac{Q}{L}\right)_{it} = \psi_1 W_{it} + \psi_2\left(Z_{it} - \beta Z_{i,t+1}\right) + \psi_3 (pq)_{it+1} + \nu_{it+1} \qquad \textbf{(26)}$$

where β is set equal to 0.95. Equation (26), estimated in levels, supposes the absence of both fixed effects and endogenous selection, using instruments dated (t-1), (t-2) and (t-3) as instruments, which allows for the possibility of errors dated t as in the case of disturbances to the adjustment function, or in the case of technological shocks. Results are reported in Table III.

In column (a) both the linear cost and the forward-looking term are estimated with coefficients which conform to the theory and are significantly different from zero. This model conforms well to the theoretical specification although the Sargan instrument validity test suggests the presence a bias arising from the presence of fixed effects. As noted, the theoretical specification implies that selection should be random and should not contain fixed effects.

---

[8] Size is measured as the reciprocal of employment and gross profits as value added minus wages, salaries and other personnel costs, among which the mandatory fund for severance payments).

[9] Estimation is based on the Arellano and Bond (1988) DPD procedure modified to take into account non-adjustment. The program is available from the author.

Table III
Dependent variable: $(Q/L)_t$

Estimation in:	**LEVELS**				**DIFFERENCES**
Model:	**(a)**	**(b)**	**(c)**	**(d)**	**(e)**
Constant			-0.301 (2.260)	-0201 (1.783)	
W_t	7.923 (67.366)	9.092 (23.465)	9.384 (13.559)	7.081 (4.853)	7.763 (5.485)
Z_t-βZ_{t+1}	0.799 (2.781)	0.180 (2.957)	0.076 (1.404)	0.331 (1.328)	-0.152 (3.221)
$(pq)_{t+1}$	-3.932 (4.413)	-3.652 (4.276)	-1.881 (2.036)	-2.095 (2.226)	-12.586 (1.778)
λ_t		-0.215 (3.381)	-0.099 (1.515)		
size82			9.718 (5.622)	9.418 (5.886)	
empl82			-0.002 (7.235)	-0.002 (5.752)	
nfi82			-0.603 (1.851)	-0.685 (2.019)	
prof82			2.469 (9.347)	2.504 (9.538)	
wages82			-0.507 (0.818)	0.769 (0.740)	
No. obs.	12794	12974	12794	12794	10195
Sargan tests for overident. restrictions	$\chi^2(152)$ 3047.7	$\chi^2(156)$ 4387.3	$\chi^2(175)$ 3752.0	$\chi^2(171)$ 3402.9	$\chi^2(111)$ 178.23
First order residual autocorr.	0.320	0.845	0.887	0.715	-0.243
Hausman-Wu test	$\chi^2(2)$=1.78		$\chi^2(2)$=5.92		

The dependent variable is productivity: $(Q/L)_t$. The regressors are: the real wage W_t, the sign L_t variable, the forward-looking term $(pq)_{t+1}$, the inverse Mills ratio λ_t, and size, employment, new fixed investment, gross profits and wage in the pre-sample year: size82, empl82, nfi82, prof82, wage82. The firm subscript is omitted from the table. Absolute t-statistics in parentheses.

We test for selectivity bias this by inclusion of an estimated inverse Mills' ratio λ_{it} calculated from a prior ordered probit using as regressors the two state variables.[10] The results reported in column (b) show the inverse Mills' ratio to be significant, providing apparent evidence for selectivity bias. In column (c), we test for possible fixed effects by including a vector of pre-sample firm characteristics.[11] In columns (c) and (d) we have allowed for an intercept in order to obtain residuals with zero mean. A likelihood ratio test of the estimates in column (b) against those in column (c) gives a value of $\chi^2(5) = 970.28$ confirming the presence of fixed effects. In these estimates, we can accept the hypothesis of random selection indicating that the evidence for non-randomness in the previous estimates may have arisen from omission of fixed effects. The estimates in column (d) omit the inverse Mills' ratio but retain the vector of firm characteristics. The estimated coefficients are broadly comparable with those in column (a), which omitted the firm characteristics, except that the coefficient of Z_t-βZ_{t+1}, which measures the size of linear costs, is substantially lower.

The standard alternative procedure for accounting for fixed effects is to estimate the equation in differences. Here, the error terms become $\Delta\nu_{it+1}=\nu_{it+1}-\nu_{it}$ and $\Delta u_{it}=u_{it}-u_{it-1}$. Instruments dated t-1 are no longer valid and we therefore use as instruments all the available history starting from t-2 backwards for the difference equation. Moreover, the method for selecting only the firms which have adjusted requires that we now use only observations corresponding to two successive periods of employment changes. Results from the difference equation are less encouraging in that the linear costs coefficient ψ_2 is estimated with a theoretically uninterpretable negative sign. The coefficient ψ_3 of the forward looking term is on the margin of conventional significance. In order to compare the specification in levels and differences we consider two Hausman-Wu tests over the reduced sample (see Appendix A.3). The values of the χ^2 statistics are reported in the final row of Table III . Both tests fail to reject the hypothesis of equality between the column (a) and column (e) estimates and between the column (d) and column (e) estimates, although in the latter case the

[10] See Appendix A.2 for a description of the procedure used to obtain the inverse Mill's ratio.The negative estimated coefficient implies a negative covariance between the equation error ν_t and the error ζ_t on the selection (adjustment) equation - see Appendix A.2. This suggest that the term $(pq)_{t+1}$ may not fully account for forward looking behaviour.

[11] Firm size, employment, new fixed investment, profitability and wage (respectively, size82, empl82, nfi82, prof82, wage82).

failure is marginal. This allows us to infer that the vector of firm characteristics incorporated in the column (d) estimates adequately account for fixed effects. Differencing involves loss of information and consequential inefficiency. The estimates in column (d) are therefore are preferred set.

The estimated coefficient of the forward-looking term($pq)_{t+1}$ implies that the discrete choice the firm faces in the future, imposed by lumpy and kinked adjustment costs, plays an important role in the marginal productivity condition and in the determination of optimal employment. In this dynamic framework, if, at any level of wages, the term $(pq)_{t+1}$ is positive, marginal productivity is lower than implied by the standard static marginal productivity condition and, conversely, is higher if pq_{t+1} is negative. If, for instance, current employment is higher than the level the firm would obtain in the absence of fixed and linear costs, then $(pq)_{t+1}$ is negative (the firm expects to adjust in the future) but current employment remains low relative to the level which would have been attained in the absence of this costs structure, and marginal productivity is correspondingly high relative to wages.

The value of α from the model is equal to 0.141. This is lower than estimates typically obtained from production functions, but not implausibly low given that we are holding capital stock fixed. It is not possible to separately identify β and γ. The estimated coefficients ψ_1 and ψ_2 are consistent with a level of linear costs around 35% of unit labour costs. [12] By contrast the fixed cost parameter K is not identified from these estimates. In order to estimate fixed costs we need to directly model the discrete choice.

7. Estimation of Fixed Costs

Hotz and Miller (1993), Slade (1998) and Aguirregabiria (1999) have developed a modelling methodology which directly exploits the information contained in the discrete decision which in principle might allow us to obtain an estimate of the fixed cost parameter K. We are able to simplify their procedure by exploiting the model reduction discussed in Section 7 and utilising the structural parameter estimates obtained from estimation of the first order condition. Rewrite the two conditional value functions as:

[12] Calculated as $\psi_2/(\psi_1*\bar{w})$ where $\bar{w}$ is the average wage.

$$v^{A} = \pi^{A} + \epsilon_{t}^{A} + \beta \int \left(P^{A} v^{A} + P^{NA} v^{NA}\right) p^{A}$$
$$v^{NA} = \pi^{NA} + \epsilon_{t}^{NA} + \beta \int \left(P^{A} v^{A} + P^{NA} v^{NA}\right) p^{NA} \tag{27}$$

where p^A and p^{NA} are the transition probabilities conditional on having optimally decided to adjust and non adjust at time t, respectively. The conditional value functions are known functions of the structural parameters, the observed state variables, the conditional choice probabilities and the transition probabilities. From the estimation of the first order condition in Section 7 we know the conditional choice probabilities, the linear cost, b, and the parameter of the production function, α. To estimate the size of fixed costs we need to estimate first the transition probabilities p^A and p^{NA}. To estimate the discrete decision process we require a discretization of the space of the observable exogenous state variables, W_t and ω_t. We consider a grid of $m=5^2=25$ cells with approximately equal numbers of observations. Write (27) as

$$\left(I_{2m} - \beta R\right) \begin{pmatrix} v^{A}(\tilde{x}) \\ v^{NA}(\tilde{x}) \end{pmatrix} = \begin{pmatrix} \pi^{A}(\tilde{x}) + \epsilon^{A} \\ \pi^{NA}(\tilde{x}) + \epsilon^{NA} \end{pmatrix} \tag{28}$$

where

$$R = \begin{pmatrix} T^{A} P^{A} & T^{A} P^{NA} \\ T^{NA} P^{A} & T^{NA} P^{NA} \end{pmatrix} \tag{29}$$

and where I is the identity matrix, T^A and T^{NA} are (m×m) matrices of the one-period transition probabilities and P^A and P^{NA} are diagonal matrices with the conditional choice probabilities on the diagonal. The value functions and profits are evaluated at the mean cell values of the state variable. The one-period transition probabilities are estimated non-parametrically. Then the future term is

$$\tilde{v}^{A}(\tilde{x}) - \tilde{v}^{NA}(\tilde{x}) = \left(I_{m} : -I_{m}\right) \left[\left(I_{2m} - \beta R\right)^{-1} - I_{2m}\right] \begin{pmatrix} \pi^{A}(\tilde{x}) + \epsilon^{A} \\ \pi^{NA}(\tilde{x}) + \epsilon^{NA} \end{pmatrix} \equiv \Lambda \begin{pmatrix} \pi^{A}(\tilde{x}) + \epsilon^{A} \\ \pi^{NA}(\tilde{x}) + \epsilon^{NA} \end{pmatrix} \tag{30}$$

where $\tilde{v}^A(\tilde{x})$ and $\tilde{v}^{NA}(\tilde{x})$ correspond to $v^A(\tilde{x})$ and $v^{NA}(\tilde{x})$ without the first term. This representation is equivalent to that in Aguirregabiria (1999). [13]

Define the indicator function

$$J_{it} = \left[A_{it}\left(\frac{\alpha A_{it}}{w_{it}}\right)^{\frac{1}{1-\alpha}} - W_{it}\left(\frac{\alpha A_{it}}{w_{it}}\right)\right] - \left[A_{it}L^{\alpha}_{i,t-1} - W_{it}L_{i,t-1}\right]$$
$$- b\left|\left(\frac{\alpha A_{it}}{w_{it}}\right)^{\frac{1}{1-\alpha}} - L_{i,t-1}\right| - K(g_i;\theta) + V\left(\tilde{W}_{j(W_{it})}\right) \qquad \textbf{(31)}$$

where g_i is a vector of firm-specific variables, $j(W_{it})$ is the discretization cell corresponding to W_{it} and V is the continuation value. The parameters α and b are obtained from the estimation of the first order condition and the firm-specific A_{it} are calculated from the residuals implied by the parameters estimates from the first order condition. The virtual wage w_{it} is implicitly defined by

$$w_{it} = W_{it} + b\left\{1\left[\left(\frac{\alpha A_{it}}{w_{it}}\right)^{\frac{1}{1-\alpha}} > L_{i,t-1}\right] - 1\left[\left(\frac{\alpha A_{it}}{w_{it}}\right)^{\frac{1}{1-\alpha}} \leq L_{i,t-1}\right]\right\}$$
$$- \beta b E\left\{1\left[\left(\frac{A_{i,t+1}}{w_{i,t+1}}\right) > \left(\frac{A_{i,t}}{w_{i,t}}\right)\right] - 1\left(\frac{A_{i,t+1}}{w_{i,t+1}}\right) \leq \left[\left(\frac{A_{i,t}}{w_{i,t}}\right)\right]\right\} \qquad \textbf{(32)}$$

According to the structural methodology, firm i should adjust in period t if $J_{it} > 0$ but otherwise it should remain with an unchanged labour force. This motivates definition of a set of a generalized residuals

$$e_{it} = d_{it} - \Phi\left(\frac{J_{it}}{\sigma_J}\right) - \delta \qquad \textbf{(33)}$$

where Φ is the standard normal distribution function, σ_J is the sample standard deviation of the J_{it} and δ is an intercept.[14] Following Slade (1998), we adopt a two stage procedure. In the first stage

[13] Note that the term relating to current period profits disappears from equation (30) as the result of subtraction of the identity matrix I_{2m} in the square-bracketed term.

[14] The alternative is to include the intercept within the normal distribution function $\Phi(.)$. That has the advantage of ensuring that the fitted values lie within (0,1) and are therefore interpretable as probabilities.

we choose θ to minimize

$$e'Z(Z'\Omega Z)^{-1}Z'e \qquad \textbf{(34)}$$

where Z is a matrix of instruments, Ω is a covariance-weighting matrix which corrects for heteroscedasticity (White, 1980). At the first stage Ω is replaced by the identity matrix and at the second stage Z'ΩZ is calculated using the first stage residuals.[15]

Note that if the discount factor beta were equal to zero, so that the forward-looking term disappears from (31), the fixed cost parameter K would simply be a constant term in the indicator function J_{it}. In that case, identification of a fixed cost parameter constant across firms implies the restriction that the function does not contain any independent intercept. It follows that, in the absence of a forward-looking term, the fixed cost parameter is identified by the functional specification of firms' technology and the choice of the normal distribution function Φ(.) in the definition of the generalized residuals - see Aguirregabiria (1999).

We specified fixed costs in three alternative ways, the last two are taking into account measures of firm size:

1. Fixed costs constant across firms.
2. Fixed costs proportional to the first year productivity level (a_{i1});
3. Fixed costs proportional to the inverse of the number of employees in the first year of the sample (size).

For each method we have both included and excluded the intercept δ. Only the third specification, with the intercept included, generated positive estimates of fixed costs. As in estimation of the first order condition (26) we set β=0.95. The estimated fixed cost coefficient, defined relative to the size variable, was 28.57 with a t statistic of 11.05.This implies that the fixed costs are

---

The specification in (33), which gives a marginally superior fit, is more directly compatible with the GMM criterion function (34).

[15] Our estimation procedure partitions the structural parameters between those estimates for the first order conditions (α and the linear cost parameter b) and those estimated from the adjustment equation (the fixed cost parameter K). The fixed cost parameter is not identified in the first order condition, but the other two parameters might in principle be estimated also form the adjustment equation. It turned out that b was very poorly determined by the adjustment equation, while the estimated value of α was implausibly high and associated with a very high estimated level of fixed costs (since firms' profits are rising in α). It is natural to consider joint estimation of the first order and adjustment equations, but this is complicated by the fact that the first order condition is only satisfied by adjusting firms. We therefore chose not to pursue this possibility.

disproportionately associated with firms with a relative small number of employees.

This result is consistent with the effects of the legislation on hiring and firing in Italy. While in general individual dismissals are very difficult except for rare cases of clearly unsatisfactory behaviour in the work place, medium and large firms may avoid unfair dismissal procedures by resorting to the legislation on redundancies which provides a relatively less expensive means of reorganizing personnel - a firm with more than fifteen employees must be willing to fire at least five employees to gain access to this legislation (*Mobilita'*).[16] The implication is that medium and large size firms have significantly greater flexibility in reducing their labour force, whereas for a smaller firm it is difficult to make five or more workers redundant without having to reschedule the whole production process. It also seems plausible that hiring and training costs are lower for large firms.

The estimation remains robust even when we discard firms with less than 15 employees which benefits from less strict hiring and firing legislation. In the case of firing, firms up to 15 employees are subject to lighter sanctions in case of unfair dismissal (up to 2-6 months pay) than larger firms. It has been argued that this has often provided incentives for firms not to expand their personnel beyond this thresholds, but to organize themselves as single-plant.

The median level of fixed costs corresponds to approximately 3.65 times average unit labour costs, while the level of linear costs, per worker, is 35% of the unit labour costs. For large firms, which will find it profitable to alter employment in almost all years, the linear cost terms is likely to dominate, while for smaller firms, which adjust less frequently and by smaller amounts,

[16] If it is to reduce its labour force, a firm with more than 15 employees must be willing to fire at least 5 of these. To do this, it has to prove that there are no alternatives to firing. The firm is then obliged to pay a month of supplementation for the redundant employees. Those workers laid-off receive a subsidy equal to 80% of the wage, up to a maximum of a monthly 1,250,000 lire ($778). The subsidy diminishes to 80% of the initial subsidy after 1 year (about 900,000 Lire; ($560)) and can be extended for 1 year if the employee is 40 or over and for 2 years if he/she is 50 or over. In the case of conflict, the firm is obliged to pay 6 months of subsidy if it does not reach an agreement with the union. This legislation is designed to discourage individual dismissals by making the cost of releasing workers very high. In practice, except for rare cases of clearly unsatisfactory behaviour in the work place, a dismissed employee will invariably initiate a claim for unfair dismissal. Bargaining between the firm and the employees will typically result either in a job buy-out, or in rehiring. Whenever an employee is fired without a fair reason she has the option either to be re-employed within 30 days, and compensated with an amount equal to at least five months and up to a maximum of 12 months wages or to accept a job loss indemnity up to 20 months wages, including 5 months wages as a fine on the company. (The legislation described was that current during the sample period and the lira/dollar exchange is the average over the period). See Del Boca and Rota, 1998.

the fixed cost term may be more important. Comparison of the practical importance of these magnitudes requires counterfactual simulation of the model.[17] The results are reported in Table IV. Linear costs account for nearly half of the sample non-adjustments, while fixed costs account for only one sixth. However, the implied non-adjustment bands are not additive, in a number of cases firms which decide not to adjust because of linear costs would also have done so because of fixed costs - if either linear or fixed costs become sufficiently high, all firms would remain with constant levels of employment. There is also a small proportion of non-adjustments which result from the fact that the smallest adjustment possible is that of one worker. The high residual arises from the failure of the model to fully capture the causes of non-adjustment, and requires that our conclusions be appropriately qualified.

Table IV
Simulated decomposition of non-adjustments

Fixed costs	3.32%
Linear costs	9.10%
Interaction	-1.37%
Integer adjustment	0.09%
Residual	9.23%
Total	20.37%

8. Conclusions

Adjustment costs affect firms' employment decisions in the face of shocks. In this paper we have focussed on the dynamics underlying lumpy and kinked costs of adjusting firm size. We have derived a simple dynamic marginal productivity equilibrium relation which takes into account the alternatives of adjustment and non-adjustment that firms face in the future, due to the presence of fixed and linear adjustment costs. The model requires addition of just one term to the static marginal productivity condition. This term is the result of a discrete decision process which takes explicitly into account the discontinuous character of the adjustment process and allows us to

[17] We performed an iterative fixed point simulation in which the adjustment probabilities P^A and P^{NA} in (29) were recalculated at each iteration until convergence.

exploit the information contained in the corner solution. We estimate the model using a two stage procedure: we first estimate the conditional choice probabilities non-parametrically and construct the forward-looking term; in the second stage we condition on firms which and exploit moment conditions to estimate the augmented first order condition.

The forward-looking term summarizes the departures from the static first order condition resulting from firms' rational anticipation of future adjustment. A marginal productivity regression which omits this term will only give asymptotically unbiased estimates if this term is uncorrelated with the wages paid by firms. This result is comparable to that obtained by Bertola (1992) in the case of linear adjustment costs, which also gives rise to "bang-bang" adjustment behaviour. We show that it is possible to characterize the forward looking aspects of adjustment in a particularly simple way in the presence of both fixed and linear costs.

Estimation underlines the importance and significance of non-convexities in the costs of adjusting employment. This is consistent with numerous prior studies of the Italian labour market which has been considered as highly regulated. We have found clear evidence for both linear and fixed costs by estimating the augmented first order condition, although this only permits quantification of the linear cost component. We estimate linear costs to be of the order of 35% of average unit labour costs. We also explored a structural model drawn from the literature of discrete decision processes, which allows us to estimate the level of fixed costs. We find that fixed costs amount to 3.65 times average unit labour costs. Simulation analysis indicates that linear costs are responsible for the largest proportion of non-adjusting firms, but these results are qualified by the presence of a substantial residual component.

References

AGUIRREGABIRIA, V. (1997), “Estimation of Dynamic Programming Models with Censored Dependent Variables”, *Investigationes Economicas*, **21**, 167-208.

AGUIRREGABIRIA, V. (1999), “The Dynamics of Markups and Inventories in Retailing Firms”, *Review of Economic Studies,* **66**,275-308.

AGUIRREGABIRIA, V. and ALONSO-BORREGO, C. (1999), “Labor Contracts and Flexibility: Evidence from a Labor Market Reform in Spain”, *manuscript*.

ARELLANO, M. and BOND, S. (1988), “Dynamic Panel Data Estimation Using DPD - A Guide for Users”, *Institute for Fiscal Studies*, Working Paper 88/15.

BENTOLILA, S. and BERTOLA, G. (1990), “Firing Costs and Labour Demand: How Bad is Eurosclerosis?”, *Review of Economic Studies*, **57**, 381-402.

BERTOLA ,G. and CABALLERO, R. (1990), “Kinky Adjustment Costs and Asggregate Dynamics”, *NBER Macroeconomic Annual* vol. **V**:237-287.

BERTOLA, G. (1992), “Labor Turnover Costs and Labor Demand”, *Journal of Labor Economic***,** **10**, 389-411.

BLACKWELL, D. (1965), “Discounted Dynamic Programming”, *Annals of Mathematical Statistics*, **36**, 226-235.

CABALLERO, R. and ENGEL, E. (1993), “Microeconomic Adjustment Hazards and Aggregate Dynamics”, *Quarterly Journal of Economics*, **108**, 359-364.

CABALLERO, R., ENGEL, E. and HALTIWANGER, J. (1997), “Aggregate Employment Dynamics: Building from Microeconomic Evidence”, *American Economic Review*, **87**, 115-137.

DEL BOCA, A. and ROTA P. (1998), "How Much Does Hiring and Firing Cost? A survey study on a sample of Italian firms," *Labour,* **12**:427-449.

ECKSTEIN, Z. and WOLPIN, K. (1989), "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models", *Journal of Human Resources*, **24**, 562-598.

EMERSON, M. (1988), "Regulation and Deregulation in the Labour Market", *European Economic Review*, **32**, 775-817.

GAVOSTO, A. and SESTITO, P. (1994), “Costi d’aggiustamento e flessibilita’ dell’occupazione: l’eterogeneita’ tra piccole e grandi imprese”, *Lavoro e relazioni industriali*.

HAMERMESH, D. (1989), "Labor Demand and the Structure of Adjustment Costs," *American Economic Review*, **79**, 675-689.

HAMERMESH, D. (1993), *Labor Demand*, (Princeton, N. J.: Princeton University Press).

HAMERMESH, D. and PFANN, G. (1994), "Adjustment Costs in Factor Demand," *Journal of Economic Literature*, **34**, 1264-1292.

HANSEN, L. and SINGLETON, K. (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models", Econometrica, **50**, 1269-1281.

HOTZ, J. and MILLER, R. (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, **60**, 497-529.

HOTZ, J., MILLER, R., SANDERS, S. and SMITH, J. (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, **61**, 265-289.

McFADDEN, D. (1981), "Econometric Models of Probabilistic Choice", in C. Manski and C. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, (Cambridge: MIT Press).

MILLER, R. (1997), "Estimating Models of Dynamic Optimization with Microeconomic Data", in H.Pesaran and P. Schmidt (eds.), *Handbook of Applied Econometrics*, vol. II (Oxford: Blackwell).

NADARAYA, A. (1964), "On Estimating Regression", *Theory Prob. Appl.*, 10, 186-90.

NICKELL, S. (1986), "Dynamic Models of Labor Demand", in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics* (Amsterdam: North Holland).

PAKES, A. (1994), "Dynamic Structural Models, Problems and Prospects: Mixed Continuous Discrete Controls and Market Interaction", in C. Sims (ed.), *Advances in Econometrics, Sixth World Congress*, vol. II.

ROTA, P. (1994), "Dynamic Labour Demand: Fixed and Quadratic Adjustment Costs" *unpublished Ph.D. thesis*.

ROTA, P. (1997), "Testing (S,s) Rules in Labour Demand" *manuscript*.

RUST, J. (1987), "Optimal Replacement of GMC Bus Engines: an Empirical Model of Harold Zurcher," *Econometrica*, **55**, 999-1033.

RUST, J. (1991), "Estimation of Dynamic Structural Models: Problems and Prospects Part I: Discrete Decision Processes," in C. Sims and J. Laffont (eds.), *Proceedings of the 6th World Congress of the Econometric Society Barcelona, Spain* (Cambridge, U.K: Cambridge University Press).

RUST, J. (1992), "Structural Estimation of Markov Decision Processes," in D. McFadden and R.

Engle (eds.), *Handbook of Econometrics*, vol. 4 (Amsterdam: North Holland).
SILVERMAN, B. (1985), *Density Estimation for Statistics and Data Analysis*, (London: Chapman and Hall).
SLADE, M. (1998), "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices", *Review of Economic Studies*, **65**, 87-107.
WATSON, G. (1964), "Smooth Regression Analysis", *Sankhyā*, Series A, **26**, 101-16.
WHITE, H. (1980), "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity", Econometrica, **48**:817-838.

Appendix

A.1. Non-parametric estimation of the conditional choice probabilities

In order to estimate (23) we require estimates of the term $(pq)_{t+1}$ and, hence, of the conditional choice probabilities P_{t+1}^{A}, and the slopes $\partial P_{t+1}^{A}/\partial L_t$, as may be seen by rewriting (23) as:

$$\left(\hat{pq}\right) = -\frac{\partial \hat{p}^{A}}{\partial L}\left[\hat{q} + \frac{1}{\hat{p}^{A}}\right] \qquad \textbf{(35)}$$

The conditional choice probabilities and the slopes, may be estimated from their sample frequencies. We use a multivariate kernel framework and adopt the Nadaraya-Watson estimator (Nadaraya, 1964; Watson, 1964). The regressors are given by the two state variables,ω_t and W_{t+1}. Given a sample of n observations on a binary variable y_i (i=1,...,n), the standard Nadaraya-Watson regression estimator of the probability $p(\xi) = \Pr(y_i=1|x_i=\xi)$ where ξ is a vector of regressors, is defined by

$$p(\xi) = \frac{\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{\xi - x_i}{h}\right) y_i}{\frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{\xi - x_i}{h}\right)} \qquad \textbf{(36)}$$

where K(.) is the kernel and h is the window. We use a normal kernel and, for two regressors, set the kernel width as

$$h = \frac{0.96}{n^{6}} \qquad \textbf{(37)}$$

following Silverman (1985, p.87). The estimates are not very sensitive to small variations in this choice. The slope $p_j(\xi)$ with respect to the jth regressor at x = ξ is given by

$$P_j(\xi) = \frac{\left(\sum_{i=1}^{n} K\left(\frac{\xi - x_i}{h}\right)\right)\left(\sum_{i=1}^{n} K_j\left(\frac{\xi - x_i}{h}\right) y_i\right) - \left(\sum_{i=1}^{n} K\left(\frac{\xi - x_i}{h}\right) y_i\right)\left(\sum_{i=1}^{n} K_j\left(\frac{\xi - x_i}{h}\right)\right)}{h\left[\sum_{i=1}^{n} K\left(\frac{\xi - x_i}{h}\right)\right]^2} \qquad \textbf{(38)}$$

where K_j(.) is the derivative of the kernel with respect to x_j.Our model requires us to estimate the probability of adjustment p^A, but it seems likely that this is non-monotonic in these variables. Although the nonparametric estimation procedure is capable, in principle, of accounting for this non-monotonicity, we found it preferable to adopt a conditional framework and to model

$$p^{A} = p^{U} + \left(1 - p^{U}\right) p^{D|\neg U} \qquad \textbf{(39)}$$

where p^U is the probability that the firm adjusts upwards and $p^{D|\neg U}$ is the conditional probability that it adjusts downwards, given that it did not adjust upwards. It is more plausible that the probabilities p^U and $p^{D|\neg U}$ are near monotonic. We found that this formulation resulted in more precise estimates.

We also investigated the converse formulation, setting

$$p^{A} = p^{D} + \left(1 - p^{D}\right) p^{U|\neg D}$$

The choice of conditioning did not make any appreciable difference to the estimates. The Gauss kernel regression program is available from the author on request.

A.2. Inverse Mill's ratio

We prefer ordered probit to standard bivariate probit in order to obtain monotonic responses. Write the ordered probit indicator function as $z_i = \delta' x_i + \zeta_i$ and let upward adjustment correspond to $z_i > a$ and downward adjustment to $z_i < b$. Noting that $E(z \mid z>c) = \phi(c)/[1-\Phi(c)]$ and $E(z \mid z<c) = -\phi(c)/[1-\Phi(c)]$, the inverse Mills' ratio is defined as

$$\begin{aligned} \lambda_i &= E(\zeta_i | z_i > a) + E(\zeta_i | z_i < b) = E(\zeta_i | \zeta_i > a - \delta' x_i) + E(\zeta_i | \zeta_i < b - \delta' x_i) \\ &= \phi(a - \delta' x_i)/[1 - \Phi(a - \delta' x_i)] - \phi(b - \delta' x_i)/[1 - \Phi(b - \delta' x_i)] \end{aligned} \qquad \textbf{(41)}$$

where $\phi(.)$ denotes the standard normal density function and $\Phi(.)$ denotes the standard normal distribution function.

A.3. Hausman-Wu test

Denote the estimate of the sub-vector of coefficients ψ in equation (23) from model (d or a) in Table III by b_0 with variance V_0, and that from model (e) by b_1 with variance V_1. Write $\Delta V = V_1 - V_0$. The Hausman test is:

$$H = (b_1 - b_0)' (\Delta V)^{-1} (b_1 - b_0) \sim \chi^2_k \qquad \textbf{(42)}$$

where k is the number of coefficients. We find V_0 to be very close to V_1 with the consequence that one of the three eigenvalues of ΔV is very small and negative. Write $\Delta V = P' \Lambda P$, where Λ is the diagonal matrix of eigenvalues of ΔV and P is the associated matrix of eigenvectors, and partition Λ as

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \qquad \textbf{(43)}$$

where λ_1 corresponds to the negative eigenvalue and Λ_2 corresponds to the k_2 positive eigenvalues.

This allows us to calculate a modified Hausman-Wu statistic

$$H^* = (b_1 - b_0)' P' \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_2 \end{pmatrix}^{-1} P(b_1 - b_0) \sim \chi^2_{k_2} \quad \textbf{(44)}$$

Note that the sample employed for the column (e) estimates in Table III is smaller than that used for the estimates reported in columns (a)-(d). The Hausman-Wu tests are calculated from estimates corresponding to those reported in columns (a) and (d) but over the smaller sample.