Estimating Job Destruction Costs with Heterogeneous Firms

Nicolas Roys *

CeMMAP and Paris School of Economics

December 14, 2008

PRELIMINARY AND INCOMPLETE

Abstract

This paper studies the dynamics of labor demand at the firm level and develops an estimation method to provide a quantitative evaluation of adjustment costs on the French labor market. The model endogenously accounts for different scale of operations that come from heterogeneous initial conditions and from permanent shocks. Structural parameters are allowed to be heterogeneous across firms by considering a finite number of types. The empirical strategy combines a simulated minimum distance estimator with the EM Algorithm to estimate the structural parameters. A general equilibrium analysis of job destruction costs is provided based on structural parameters estimate.

1 Introduction

Microeconomic flexibility is a major driving force in modern market economies. Its main obstacle is adjustment costs. Firms change their demand for inputs more slowly than the shocks to input demand warrant and that is because they incur adjustment costs. There are technological costs that design the costs of reduced efficiency during the period of adjustment and there are institutional costs that are mainly the effects of government policies. Chief among them is labor market regulation, especially in European countries where employment relations are highly regulated. These regulations seek to enhance job-security by reducing dismissals of workers and more generally fluctuations in employment.

Adjustment costs are used in many fields of economics but are intrinsically difficult to measure (Hamermesh and Pfann (1996)). Estimating adjustment costs

*nicolas_r@ifs.org.uk. The CeMMAP is a joint-venture between University College London and the Institute for Fiscal Studies, London.
may help the quantitative analysis in models that are using adjustment costs to explain a wide range of facts including the behavior of productivity and employment over the business cycle, the impact of job-securities policies, the process of creative-destruction and the micro-foundations of aggregate employment adjustment.

The simplest (and widely used) dynamic model of labor demand¹ is a so-called partial adjustment model in which the firm adjust the level of employment (say $n_t$) to a target (say $n_t^*$). Formally,

$$\log(n_t) - \log(n_{t-1}) = \lambda(\log(n_t^*) - \log(n_{t-1}))$$

(1)

The change in employment is proportional to the difference between the previous level of employment and a target where $\lambda$ parameterizes how quickly the gap is closed. This model implies a smooth and continuous adjustment of employment to shocks. Sargent (1978) shows that more elaborated version of this model may be derived as the solution to a firm’s dynamic profit maximization problem under the assumption that there are quadratic costs of adjusting the workforce. In part, the popularity of the quadratic adjustment cost structure reflects its tractability. The implications of this model conflict with evidence of inactivity and bursts at the firm level. In a seminal paper, Hamermesh (1989) examines monthly data on output and employment between 1983 and 1987 across seven manufacturing plants. For each plant, output fluctuates substantially over the sample. Employment exhibits long periods of constancy broken by infrequent and large jumps at times roughly coinciding with the largest output fluctuations. Hence, the plant data are not consistent with the smooth employment adjustment that would arise from convex adjustment costs. Caballero et al. (1997) show that employment change tends to be concentrated in a single-period, so that firms avoid paying adjustment costs too frequently. The adjustment process is lumpy and intermittent: in the face of a shock, a firm may decide that it is optimal to maintain the same number of employees and to postpone adjustment. The estimation on quarterly US data by Cooper et al. (2005) concludes that quadratic adjustment costs yields predictions that are very much at odds with the micro data while non-convex adjustment costs does a better job for explaining plant-level moments. Particularly, the quadratic adjustment cost model is unable to generate the negative correlation of hours and employment growth at the plant level.

To mimic these microeconomic facts, I consider a constant cost to destroy jobs which is simple but enough to reproduce observed patterns of employment at the firm level. Using a fully specified dynamic programming model, I develop an estimation strategy to quantify job destruction costs on the french labor market. The model is characterized analytically. Solving the model reduces to the resolution of a system of two non-linear equations. The usual discretizations and iterative algorithm are avoided. The interest is to considerably alleviate the computing costs and the numerical difficulties associated with discretization. Further the model endogenously account for different scale of operations that come from heterogeneous initial conditions and from permanent shocks. Related literature consider ex-ante identical firm

¹see Hamermesh (1993) and Bond and Reenen (2006) for exhaustive surveys of the literature on, respectively, labor demand and micro-economic models of investment and employment
but heterogeneous ex-post. This paper extends the literature by considering firm that are heterogeneous ex-ante and ex-post. Thus a second contribution is to allow structural parameters to be heterogeneous across rms by considering a finite number of types. The empirical strategy combines a simulated minimum distance estimator with the EM Algorithm. Unobserved heterogeneity is introduced in the auxiliary model via finite mixture modelling. Finally, the last contribution is to provide a general equilibrium evaluation based on structural parameters estimates. It gives a better understanding of the effect of adjustment costs compared. In particular it revisits the findings of the seminal paper of Hopenhayn and Rogerson (1993) that uses calibration to evaluate the effects of firing costs.

The remaining of the paper proceeds as follows. Section 2 presents the data and descriptive evidence on labor adjustment. Section 3 develops the theoretical framework. Section 4 tests the theoretical predictions and measure the importance of unobserved factors. Section 5 develops the methodology for the structural estimation of the parameters. Section 6 concludes.

2 Facts on Labor Adjustment

2.1 The Data and Descriptive Statistics

The data used is the BRN (Real Normal Profits). The BRN declarations are completed annually by firms with a turnover of more than 3.5 million francs (1992 threshold) liable for income tax in respect of BIC (Industrial and Commercial Profits). The BIC correspond to the profits declared by firms whose commercial, industrial or craft-work activity is carried out for lucrative purposes (60% of the firms, 94% of the turnover). The data cover the period 1994-2000. We focus on the manufacturing sector.

We drop firms with a level of employment below 100 employees. The reason is to avoid the effect of size-dependent policies (especially important when the size of 10 and 50 employees are crossed) which results in threshold effects in labor demand. Our empirical strategy can not accommodate these effects and we leave this for future research.

Figure 1 plots aggregate yearly rates of job turnover which are computed according to standard definitions (see Davis and Haltiwanger (1999)). As usually found in most industrial countries, aggregate job flows are large within the business cycle. Net employment growth is always much smaller than job creation or destruction.

Figures 2 and 3 show the range of variation of production and employment. Throughout the period, on average, the rate of change in employment was zero for almost 15% of the firms. Employment growth rates display high spikes around zero, compared to the smooth patterns of sales variations observed. It reveals a considerable stickiness in employment.

2 facts stand out from the distribution of job reallocation: 1. there is a significant amount of relatively small net employment adjustment and 2. these small adjustments are complemented by significant bursts of job creation and destruction: almost 30% of firms either contract or expand employment by more than 10% in a given year.
Figure 1: Job Turnover Statistics

Figure 2: Output Variation
2.2 Related Literature on the French Labor Market

There are two types of regular employment contracts in France: indefinite-term contracts (Contrats Dure Indetermine, CDI) and fixed-duration contracts (Contrats Dure Dtermine, CDD). Although their use is formally restricted, CDDs are the most common method of hiring: more than 2/3 of all hires are through CDD. Employment Protection Legislation heavily regulates the termination of CD: the employer must observe a mandatory waiting notice period and pay a severance payment.

Goux et al. (2001) use a panel of 915 French manufacturing firms for which they can measure the number of hirings and firings for indefinite and fixed term contract workers separately. They estimate the costs of employment adjustment for these two types of workers using a dynamic labor demand model with quadratic adjustment costs. For indefinite contract workers they find firing cost to be much higher than hiring costs (around 40 times higher). Their estimates only allow to make these comparative statements. They do not allow to measure the absolute amount of adjustment costs. They also find that it is practically costless to adjust workers on fixed time contracts. Using survey data on actual severance payments and actual costs (such as training hours, expenses on job advertising, etc.) upon hiring, Abowd and Kramarz (2003) provide an analysis of hiring and firing costs for French firms. They conclude that for permanent contracts, the cost of hiring are much lower than the cost of firing. Based on two cross-sections, Kramarz and Michaud (2004) revisit the findings of Abowd and Kramarz (2003). Again they find that collective terminations have the highest costs. They also find hiring cost to be small. For instance they mention that in two surveys of French firms that hire workers respectively 49 % (for 1992) and 62 % (for 1996) of the firms declare a zero hiring cost. It is hard to tell in advance the size of hiring or firing cost as most of these costs are hard to observe directly. In principle, in the case of hiring they include job advertising, interview cost and training cost of the employee and the cost of reorganizing work. Likewise, the cost of reducing employment includes

Figure 3: Employment Variation
besides legal severance payments, the cost of helping employee finding new job, and other costs. \( ? \) examines French data on the entry and exit of workers from firms and attempted to use information on the size of costs associated with movements of workers. They found that they were very high fixed costs associated with firing workers and most adjustment was through varying the hiring rate. Overall, the findings of the French literature suggest that hiring costs are likely to be low and firing cost are mostly high for collective terminations and for employees on indefinite term contracts. However they are low for fixed term contracts.

3 A Model of Labor Demand

3.1 Framework

The explanation chosen for the described facts is the existence of adjustment costs that generate inaction. With adjustment costs, the simple condition which states that the marginal productivity and the marginal costs of labor are equated in every period, is no longer efficient. The cost of destroying jobs requires a firm to adopt a forward-looking employment policy.

Time is discrete and indexed by \( t \). Consider a risk neutral firm that produces an homogeneous good.

Each firm has a production function \( Bn^a \) in productivity \( B \) and employment \( n \) where capital and others inputs are assumed to be maximized out. The firm faces an iso-elastic demand curve with elasticity \( \eta \): \( CP^{-\eta} \). The firm’s monopoly power decreases when \( \eta \) rises.

These can be combined into a revenue function: \( BC^{1-\frac{1}{\eta}}n^{a(1-\frac{1}{\eta})} \). To alleviate notation, I define \( A^{1-a} = BC^{1-\frac{1}{\eta}} \) and \( \alpha = a(1-\frac{1}{\eta}) \). Combined with a constant wage rate \( (w) \), it gives the following profit function as a function of \( A \) and \( n \):

\[
\pi(A, n) = A^{1-a}n^a - wn
\]

(2)

\( \alpha < 1 \) reflects decreasing returns-to-scale and/or market power. Variations in profitability \( A \) could reflect variations in product demands or variations in the productivity of inputs. Thus a firm may have a high productivity because it has market-power and/or because of the higher quality if its production. I do not have information on output prices so that I can not disentangle these two effects.

The manager problem is to choose a state contingent sequence of employment to maximize the present discount value of current and future profits. The time horizon of the decision problem is infinite. The parameter \( \beta \) represents the rate at which the agent discounts utility at future periods, and it belongs to the interval \( ]0, 1[ \). The decision problem of a firm at time \( t \) is to maximize the present discounted value of current and future profits, given the previous level of employment and the current state of the profitability. At the beginning of the period, the manager knows his past employment level \( (n_t) \) and current level of profitability \( (A_t) \).

Every period, a new value of productivity is drawn from a Markov transition \( Z(A_{t+1}, A_t) \). Shocks to the Given \( (n_t, A_t) \) the manager hires or fires employees \( (d_t) \) who become immediately productive. The level of employment at the start of the next period, \( t + 1 \), is \( n_{t+1} = n_t + d_t \).
There is a constant cost \( c \) that is paid for every job destroyed. We consider net adjustment costs not brute: adjustment costs are associated with job destruction but not with workers flows. Thus adjustment costs here are unrelated to the identity of the workers who fill these positions. This choice is entirely determined by data availability.

I do not consider fixed adjustment costs since it creates many technical difficulties (see Roys (2007)) and it requires informations on plants (and not firms). I do not consider convex adjustment costs since they generates smoothing which do not appear in the data. The model also assumes no entry or exit. This is for tractability and because I do not have data on entry/exit.

Define the value function at period \( t \), \( V(A_t, n_t) \) as the discounted expected value of current and future cash-flows:

\[
V(A_t, n_t) = \max_{d_j, j \geq t} E \left\{ \sum_{j=t}^{\infty} \beta^{j-t} \left[ A_j^{1-\alpha} (n_j + d_j)^\alpha - w(n_j + d_j) - c \min(0, d_j) \right] | n_t, A_t \right\}
\]

Under standard regularity conditions (see Stokey et al. (1989)), we can write the following recursive functional equation where the value function \( V(A, n) \) is given by the solution to Bellman’s equation:

\[
V(A, n) = \max_d \left\{ A^{1-\alpha} (n + d)^\alpha - cd - \beta \int V(A', n + d) Z(dA', A) \right\}
\]

\( V \) satisfies standard properties. Nevertheless it does not satisfy the differentiability conditions in Benveniste and Scheinkman (1979) because of the kinked in the adjustment costs function when employment stays constant. Fortunately Rincon-Zapatero and Santos (2007) provides conditions under which the value function is continuously differentiable without interiority assumptions. A formal derivation of the optimal decision rule is given in the Appendix.

Given \( (A, n) \) the optimal choice \( d \) must satisfy the first order conditions:

\[
\alpha A^{1-\alpha} (n + d)^{\alpha-1} - w + \beta E (V_n(A', n + d)) \leq 0
\]

with equality if \( d < 0 \), and

\[
\alpha A^{1-\alpha} (n + d)^{\alpha-1} - w + c + \beta E (V_n(A', n + d)) \leq 0
\]

with equality if \( d > 0 \).

The optimal decision rule is then as follows:

\[
n'(A, n) = \begin{cases} \bar{n}(A) & \text{if } n > \bar{n}(A) \\ n & \text{if } \underline{n}(A) < n < \bar{n}(A) \\ \underline{n}(A) & \text{if } n < \underline{n}(A) \end{cases}
\]
\[ \alpha A^{1-\alpha} \pi(A)^{\alpha-1} + \beta E(V_n(A', \pi(A))) = w \]
\[ \alpha A^{1-\alpha} \mu(A)^{\alpha-1} + \beta E(V_n(A', \mu(A))) = w - c \]

Optimality requires the firm to create and destroy jobs as needed to keep the marginal value of labor in the closed interval \([w - c, w]\). The optimal decision rule consists of two targets that verify an Euler equation. If the level of employment at the beginning period lies between the two target, it is not worth hiring/firing and the firm stays put until the next period. The optimization problem has a sequential nature: 1. choosing a target for employment and 2. whether to hire/ fire or stay put. And there is no smoothing: if the manager decides to adjust, he directly jumps to the target without smoothing and independently of lag employment value. Figure 4 plots the optimal employment for a given level of profitability.

![Figure 4: Optimal Decision rule for a given level of Profitability](image)

### 3.2 Solving the Model

In the Appendix, the model is characterized analytically. Solving the model reduces to the resolution of a system of two non-linear equations. The usual discretizations and iterative algorithm are avoided. The interest is to considerably alleviate the computing costs and the numerical difficulties associated with discretization. The model is solved without numerical approximations.

In the model, the value function is homogeneous of degree 1:

\[ V(A, n) = AV(1, \frac{n}{A}). \]

A virtue of the homogeneity of the value function is that what matters for the dynamic programming problem is \(y = \frac{n}{A}\). Using the Cobb-Douglas revenue function
it follows: \( y = p^{\frac{1}{\alpha}} \) where \( p \) is the observed average productivity of labor. Log-profitability follows a gaussian random-walk: \( \ln(A') = \ln(A) + \epsilon \) where \( \epsilon \) is i.i.d and gaussian with drift \( \mu \) and standard deviation \( \sigma \).

The optimal policy is summarized by firm’s choice for labor productivity:

\[
p(p_{-1}u) = \begin{cases} 
  p & \text{if } p_{-1}u < p \\
  p_{-1}u & \text{if } p < p_{-1}u < \bar{p} \\
  \bar{p} & \text{if } p_{-1}u > \bar{p}
\end{cases}
\]

where \( p \) is the optimal choice, \( p_{-1} \) is past period choice of labor productivity, \( u = e^{\epsilon(\alpha-1)} \) is a measure of the shock to profitability. \( p \) and \( \bar{p} \) are two labor productivity target. Their expression is given in the appendix.

Note that \( p \geq \bar{p} \). Labor productivity is higher in expansion than in contraction. This is because the firm is not destroying (creating) as much jobs as it would in a frictionless labor market. This is here not coming from lagging effects of employment.

Turning to the effect of adjustment costs on labor productivity. The effect is ambiguous. Following a positive(negative) shocks, the firm hires (fires) relatively less (more) which lead to a higher (lower) productivity of labor in expansion (recession) period. Overall the effect is ambiguous.

### 3.3 Identifying The Structural Parameters with a Panel of Firms

As will be described below, I use simulated minimum distance to estimate the structural parameters. This estimator needs observed moments that are a well-behaved function of the structural parameters. I show in this section how the parameters of the model affect features of employment and output at the firm level leaving the exact description of the estimator to the empirical section.

The overall link between the structural parameters and the observed moments is summarized in table 1. A sign \(+\) \((-\) indicates a positive (negative) relationship. No sign does not mean that the effect is exactly null but that it is likely to be small.

<table>
<thead>
<tr>
<th></th>
<th>( \bar{p} )</th>
<th>RR</th>
<th>NET</th>
<th>INA</th>
<th>( \frac{\sigma}{\bar{p}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>( \epsilon )</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>(</td>
<td>\mu</td>
<td>)</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>( \sigma_\epsilon )</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Structural Parameters and Observed Moments

The time-series average of labor productivity (say \( \bar{p} \)) identifies the parameter of the production function \( \alpha \). To understand the mechanism at work, consider a firm facing no job destruction costs. Its decision consists in equalizing at every period the marginal productivity of labor with wage \( w \). Simple calculations show that labor productivity is then:

\[
p = \frac{O}{n} = \frac{A^{1-\alpha}n^\alpha}{n} = \frac{w}{\alpha}
\]

(12)
Obviously, a high wage rate (α) discourages (encourages) hirings. In a dynamic framework, this is still true but labor productivity follows a non-trivial dynamics due to adjustment costs.

Identifying adjustment costs will rely on three moments. First, adjustment costs increase the within-firm variance in labor productivity. Second it is not always optimal to create and destroy jobs. Then adjustment costs increase inaction. Third, adjustment costs decrease the propensity to create and destroy jobs: reallocation rate should be low when adjustment costs are high.

The standard deviation of shock σ increases the within firm variance in labor productivity, the reallocation but reduces the inaction rate. A high within-firm variance may mean high adjustment costs and/or highly volatile environment. To separately identify c and σ I use their opposite effect on the inaction rate and the reallocation rate.

Obviously the trend in profitability μ increases the growth rate of employment.

More moments will be added later on notably to account for measurement errors (see below).

Finally, note that the identification strategy does not depend on functional forms assumptions.

4 Preliminary Evidence from a Double Selection Model

While I could directly turn to the structural estimation, estimating a reduced-form model is an insightful step for a least two reasons. First, it is an easy way to check the empirical relevance of the model. Two easily testable properties of the model are the following: 1. the probability that employment increase is an increasing function of current profitability and 2. the probability that employment increase is a decreasing function of past employment. Second, the chosen structural method is motivated by the existence of permanent unobserved differences between firms. With reduced-form estimation, the importance of these unobserved factors can be estimated.

Define the employment target for positive and negative net variation s = +, − :

\[ n_{it}^* = a^s A_{it} + c_i^s + u_i \] (13)

The firm might not wish to hire or fire. For this reason there is selection equation that determines in which regime the firm is:

\[ d_{it}^* = c A_{it} + d n_{it-1} + \omega_{it} + v_i \] (14)

Then, we have a panel of firms and we observe the following:

\[ d_{it} = I(d_{it}^* > s^+) - I(d_{it}^* \leq s^-) \] (15)

\[ n_{it} = I(d_{it}^* > s^+)n_{it-1}^+ + I(d_{it}^* \leq s^-)n_{it-1}^- + I(s^- \leq d_{it}^* \leq s^+)n_{it-1} \] (16)

\(^2\)all variables considered in this section are in logarithm
More structure has to be imposed to estimate this problem. I choose a parametric approach. As $T$ is small (even if $N$ is large), the fixed effects model suffers from the incidental parameters problem. Indeed, as there does not exist a sufficient statistic allowing the fixed effects to be conditioned out of the likelihood, the fixed-effects estimator is biased. So I used a random effects estimator even if its properties are more dependent on the parametric hypothesis we make. More precisely, I use correlated random-effects a la Chamberlain. Regarding the initial conditions problem, I use Wooldridge (2005) solution. It follows that individual effects are assumed to be:

\begin{align}
    u_i &= \lambda_u + \lambda_u' z_i + \lambda_u n_{i0} + u_i^* \\
    v_i &= \lambda_v + \lambda_v' z_i + \lambda_v n_{i0} + v_i^*
\end{align}

($u_i^*, v_i^*$) are assumed i.i.d across individuals and normally distributed. The estimation is performed by Maximum Likelihood using Double Gauss-Hermite Quadrature. To incorporate aggregate-effects, temporal dummies are added.

### 4.1 Production Function Estimation

I infer unobserved productivity form the estimation of a production function. The estimation of the production function is relatively standard. The main issue is the fact that inputs are endogenous: the productivity shock observed by the firm changes the factor input decision leading to a classical simultaneity problem. A standard approach is to assume a Cobb-Douglas production function with an additive, time-constant firm effect, and to solve the unobserved heterogeneity problem by using fixed effects estimation. The main forthcoming of this approach is that it requires the strong assumption of strict exogeneity of the inputs, conditional on firm heterogeneity. Instrumental variables methods can be used to relax the strict exogeneity assumption of the inputs. In particular, after differencing or quasi-differencing, lagged inputs can be used as instruments for changes in the inputs. But, the application of standard GMM estimators which take first differences and use as instruments lagged levels has produced unsatisfactory results (?). These problems are due to weak instruments as the series are highly persistent. Blundell and Bond (2000) apply a system GMM (which adds to the preceding moment conditions, the regression of the levels using first differences as instruments) to production function. They find that the system GMM greatly improves the result.

I pursue a more structural approach that is consistent with the labor demand econometric and theoretical model. Rather than allowing for time-constant firm heterogeneity, Olley and Pakes (1996) show that, under certain assumptions, investment can be used as a proxy variable for unobserved, time-varying productivity. Specifically, productivity can be expressed as an unknown function of capital and investment (when investment is strictly positive). They present a two-step estimation method where, in the first stage, semi-parametric methods are used to estimate the coefficients on the variable inputs along with the nonparametric function linking productivity to capital and investment. In a second step, the parameters on capital inputs can be identified under assumptions on the dynamics of the productivity process. More recently, Levinsohn and Petrin (2003) propose a modification of the
former approach to address the problem of lumpy investment. They suggest using intermediate inputs to proxy for unobserved productivity and propose a two-step estimation method to consistently estimate the coefficients on the variable inputs and the capital inputs. Wooldridge (2005) shows how the moment conditions can be implemented in a GMM framework.

For firm $i$ at time $t$, the production can be written as:

$$y_{it} = c + l_{it} \alpha + a_{it} + e_{it}$$  \hspace{1cm} (19)

where $\{e_{it} : t = 1, \ldots, T\}$ is measurement error in output. Some variables inputs are a strictly monotonic function of $z$ (for example materials): $m = g(a, l)$. This function can be inverted and $g^{-1}$ can be approximated with polynomials.

The results of the structural method are reported in Table 4.1 along with standard OLS and within estimation.

Consistently with the Literature, OLS estimate is bias upward while Within estimator is bias downward. The structural method provides a plausible and precisely estimated coefficient for $\alpha$. It is then possible to obtain measure of profitability using the residuals of the estimation. The profitability level will be treated as an observable state variable in the next sub-section.

### 4.2 Estimation Results for the Selection Model

The results are reported in Table 4.2. The theoretical predictions are verified: 1. employment is increasing in profitability, 2. the probability of an adjustment is increasing in profitability and 3. conditional on profitability, the probability of an adjustment is decreasing in past employment level.

Unobserved heterogeneity variance decreases significantly when is taken into account initial conditions and the correlation between the regressors and firm-specific
effects: its contribution to the total variance decreases from 31% to 11.5%.

5 Structural Estimation: Framework

Estimating the parameters using Euler equation techniques is not adequate here because corner solutions are relatively important: around 15% of the observations correspond to zero adjustment. This means that selection bias is likely to be severe. Pakes (1994) proposed to estimate the structural parameters using some modified Euler equations which take into account the number of periods between two consecutive interior solutions. But there remain several limitations of the Euler equations. For example in the consumption literature, the estimation using Euler equation techniques has been very disappointing (cf. Browning and Alan (2008) among others). The problems identified are manifold but the most important seem to be the paucity of long panels and the substantial measurement error in observed variables. The data-set used here is relatively short $T = 7$. Measurement errors issues are discussed below.

Exploiting the discrete decision (whether to adjust employment level of not) would be another possibility. For example, Rota (2004) estimate a labor demand model with fixed costs using Hotz and Miller (1993)’s estimator. The former is suited for discrete decision processes (the decision to adjust or not). The continuous decision (how much to invest) is estimated non-parametrically. Three important limitations appear here. First, it heavily relies on the conditional independence assumption (Rust (1987)) which essentially states that once we condition on observables, there are no serially correlated unobservables. But the dataset used contains a very large number of firms with a limited number of variables. There is considerable heterogeneity among the observed variables. Second, identification of the structural parameters would entirely rely on the discrete decision. Important measurement errors can be expected in the observed employment. It is an average of the number of employees in the firm that ignore the flows during the year. Lastly, we have information on firms and not plants.

As an alternative, I use a simulated minimum distance estimator as recently described by Hall and Rust (2003).

5.1 Firm Heterogeneity

In the current literature, the shock is supposed to follow a stationary autoregressive process $\ln(A') = \rho \ln(A) + \epsilon$. Why assuming a random walk rather than an AR(1) ? First of all, there is little agreement on the degree of persistence. In similar contexts, persistence has been estimated to be between .4 and .9 (Cooper and Haltiwanger (2007), Cooper and Ejarque (2003), Gomes (2001), Hennessy and Whited (2005)) ! More importantly, under the assumption that profitability follows a random walk, the model endogenously account for different scale of operations that come from heterogenous initial conditions and from permanent shocks. The former are relevant from Abbring and Campbell (2005) who show that heterogeneity in firms’ pre-entry scale decisions accounts for most of firms’ heterogeneity. The latter are relevant from Franco and Philippon (2007) and Gourio (2008) who document the importance of
permanent shocks at the firm level. Essentially assuming an AR(1) is assuming that firms oscillate around a fixed size. But the aforementioned paper shows that firms drift one way or another due to permanent shocks. Some firms are growing permanently while some firms are shrinking permanently.

In reality, firms operate in heterogeneous environment with different level of volatility and different trends. They have different degree of returns to scale, face level of competition and pay different job destruction costs. For these reasons, I allow for firm specific coefficients. Since I do not have access to a long panel, this is doable only through the use of a finite number of types.

5.2 Measurement Errors

The variable employment in the dataset is defined as full-time equivalent employment that is total hours worked divided by average annual hours worked in full-time jobs. The observed variable includes some part-time and temporary workers. Further, heavy rounding can be expected. I explicitly introduce employment measurement error into the simulated moments to mimic the bias these impute into the actual data moments. I follow Bloom (2007). Assume firm log wages can be decomposed into: \( \log(w_{it}) = \log(n_{it}) + \eta_t + u_i + \omega_{it} \) where \( \eta_t \) is a time dummy, \( u_i \) is a firm specific wage rate, \( n_{it} \) is observed employment and \( \omega_{it} \) is a residual. If this decomposition is correct, the coefficient on \( \log(n_{it}) \) is \( \frac{\sigma_n^2}{\sigma_n^2 + \sigma_{\omega n}^2} \) where \( \sigma_n^2 \) is the variance in log employment and \( \sigma_{\omega n}^2 \) is the variance of the measurement error in log employment. I find a coefficient (standard error) on \( \log(n_{it}) \) of .882(0.003). I incorporate these into the simulation estimation by multiplying each firm annual employment by \( mn_{it} \) that is i.i.d over firm and time and follows a log-normal distribution with mean 0 and standard deviation 0.0247.

Measurement error in employment severely bias our estimates of the inaction rate. To handle this, we consider as inaction any observation with a job creation or destruction of less than a threshold of 1%. This arbitrary value is not a source of bias as long as observed and simulated moments are treated identically.

Output is measured by value-added which is defined in the dataset as the difference between production and intermediate consumptions net of all variations in stocks. From the literature on structural estimates of productivity (see for example the recent survey of Ackerberg et al. (2007)), conflating true variation in productivity and measurement error in output would overestimate the degree of profitability dispersion over both time and firm. I incorporate a measurement error in log output into the simulation estimation by multiplying each firm annual value-added by \( mo_{it} \) that is i.i.d over firm and time and follows a log-normal distribution with mean 0 and standard deviation \( \sigma_o \). How to separately identify \( \sigma_o \) and true variation in profitability \( \sigma_f \)? It is identify by looking at variation of output that are correlated with variation in employment as distinct form variation in output that are uncorrelated with variation in employment. An informative moment is then the covariance between first-difference employment and first-difference output. Measurement error attenuates this correlation while firm volatility has a small positive impact on it.
5.3 The Auxiliary Model

The auxiliary model is a finite mixture model that account for a finite number of firms' type whose respective proportions are estimated. The basic idea is that data come from a population with several subpopulations. The overall population is a mixture of subpopulations each having its own model. Let $Z_i$ denote a random vector whose components are the time series average for firm $i$ of moments related to employment and production dynamics. $f(z)$ is its density function. It is a finite mixture of gaussian that can be written in the form:

$$f(z_i; \Psi) = \sum_{k=1}^{K} \lambda_k \phi_k(z_i; \mu_k, \Sigma_k)$$

where the $\phi_k(z_i; \mu_k, \Sigma_k)$ are multivariate gaussian densities. For each $k \in 1, ..., K$, $\lambda_k$ is the proportion of the population in the group $k$. The $\lambda$s are nonnegative quantities that sum to one. $\Psi = (\mu_1, ..., \mu_K, \Sigma_1, ..., \Sigma_K, \lambda_1, ..., \lambda_{K-1})$ is the vector of auxiliary parameters. Under the assumtion that $z_i$ are iid, the likelihood of the model writes:

$$L(\Psi) = \prod_{i=1}^{N} \sum_{k=1}^{K} \lambda_k \phi_k(z_i; \mu_k, \Sigma_k)$$

The model is estimated using the EM Algorithm. The EM algorithm (see Dempster et al. (1977) and McLachlan and Peel (2000)) is a general approach to maximum likelihood estimation for problems in which the data can be viewed as consisting as $N$ multivariate observations $(z_i, t_i)$ in which $t_i$ is unobserved. For mixture models, $z_i$ is observation for firm $i$ and $t_i = (t_{i1}, ..., t_{iK})$ is defined as:

$$t_{ik} = \begin{cases} 1 & \text{if } z_i \text{ belongs to group } k \\ 0 & \text{otherwise} \end{cases}$$

The $t_i$’s are iid realizations from a multinomial distribution of one draw from $K$ types with probability $\lambda_1, ..., \lambda_K$. The complete-data likelihood (say $L_c$) is then:

$$L_c(\Psi) = \prod_{i=1}^{N} \sum_{k=1}^{K} [\lambda_k \phi_k(z_i; \mu_k, \Sigma_k)]^{t_{ik}}$$

At each iteration $j$, given $\Psi^j$, the EM algorithm alternate between two steps.

1. **E Step**: Compute the conditional expectation of the complete-data log-likelihood $L_c$ given the observed data and the current parameter estimated $\Psi^j$. From the linearity of $L_c$ in the unobservables $t_{ik}$, it gives for every $k = 1, ..., K$ and $i = 1, ..., N$:

$$\hat{p}_{ik}^j = \frac{\lambda_k^j \phi_k(z_i; \mu_k^j, \Sigma_k^j)}{\sum_{k'=1}^{K} \lambda_{k'}^j \phi_{k'}(z_i; \mu_{k'}^j, \Sigma_{k'}^j)}$$

$$\mu_k^j, \Sigma_k^j$$
2. M Step: Maximize the expected log-likelihood from the E step. Expressions are given explicitly by:

\[
\lambda_{j+1}^{k+1} = \frac{1}{N} \sum_{i=1}^{N} t_{ik}^j
\]

(24)

\[
\mu_{k+1}^j = \frac{\sum_{i=1}^{N} t_{ik}^j z_i}{\sum_{i=1}^{N} t_{ik}^j}
\]

(25)

\[
\Sigma_{k+1}^j = \frac{\sum_{i=1}^{N} t_{ik}^j (z_i - \mu_{k+1}^j) (z_i - \mu_{k+1}^j)'}{\sum_{i=1}^{N} t_{ik}^j}
\]

(26)

For mixture models, it is well-known that asymptotic theory does not provide reliable standard errors unless the number of observations is very large. For that reason, I use non-parametric bootstrap for standard errors.

5.4 Simulated Minimum Distance

I describe how I recover the structural parameters \( \theta = (\theta_1, ..., \theta_K, \theta_c) \) where \( \theta_k \) is type-specific and \( \theta_c \) is common across firms. \( \hat{\lambda}_k \) is an estimate of the proportion of each type. The proportion of each types \( \lambda_k \) are used to determined the number of firm of each types used in the estimation. I simulated an economy population with a number of firms for each type equal to \( N_k = round(\lambda_k N) \) where \( N \) is the sample size. The matched type-specific moments are the \( \hat{\mu}_k \). Note that the \( \hat{\mu}_k \) capture within firm variation contional on its type. Between-firm variation is accomodated via finite mixture modelling.

I use the MPEC approach recently advocated by Su and Judd (2008). Usually minimum distance methods involve two nested steps: 1. given an value of the structural parameters \( \theta \), the model is solved and theoretical moments are computed 2. theoretical moments are compared to observed moments. The MPEC approach consists in minimizing the distance between observed and simulated moments with respect to the structural parameters and the policy function subject to the constraint that the policy function is optimal.

Here the policy functions are fairly simple and can be summarized by two parameters for each type of firms \( k \): \( \overline{y}_k \) and \( y_k \). Let \( G_k(\theta_c, \theta_k, \overline{y}_k, y_k) = 0 \) summarize the optimality conditions.

For a given simulation \( s \), let the simulated moments \( \Psi^s(\theta_k, \theta_c, \overline{y}_k, y_k) \) be a function the type-specific structural parameters \( \theta_k \), the common parameter \( \theta_c \) and the policy \( (\overline{y}_k, y_k) \).

The estimator of \( \theta_k \) for every \( k \in K \), \( \theta_c \) and \( (\overline{y}_k, y_k) \) minimizes:

\[
\sum_{k=1}^{K} \lambda_k \left\| \frac{1}{S} \sum_{s=1}^{S} \Psi^s(\theta_k, \theta_c, \overline{y}_k, y_k) - \hat{\mu}_k \right\|_{\Omega_k}
\]

(27)

subject to the constraints:

\[
G_k(\theta_c, \theta_k, \overline{y}_k, y_k) = 0 \text{ for every } k \in K
\]

(28)
where $\Omega_k$ is weighting matrix given by a bootstrapped estimation of the variance-covariance matrix of the $\hat{\mu}_k$ and $S$ is the number of simulation.

Testing for the number of components in a mixture is an important but difficult problem which has not been completely resolved. My understanding of the literature is that the BIC is a commonly used criteria for choosing the number of components. This is defined by

$$BIC = 2 \times \log L(\hat{\Psi}) - \dim(\hat{\Psi}) \times N$$

where $\hat{\Psi}$ and $L(\hat{\Psi})$ denote respectively the estimated parameters values and the maximized likelihood.

The variance of the estimator is estimated by bootstrap. In the $b^{th}$ bootstrap repetition, a new set of type-specific moments $\hat{\mu}_k^b$ is produced by randomly selecting observations from the original data with replacement. An estimator $\theta_b$ is found by minimizing the weighted distance between the re-centered simulated moments

$$\sum_{k=1}^K \lambda_k \left\| \frac{1}{S} \sum_{s=1}^S (\Psi_s^*(\theta_k^b, \theta_c, \bar{y}_k, y_k) - \Psi_s^*(\theta_k, \theta_c, \bar{y}_k, y_k)) - (\hat{\mu}_k^b - \hat{\mu}_k) \right\|_{\Omega_k}$$

subject to the constraints

$$G_k(\theta_c, \theta_k^b, \bar{y}_k, y_k) = 0 \text{ for every } k \in K$$

Each type has a specific and exogenous wage rate $w_k$ estimated in the auxiliary model. The matched and type-specific moments taken from the data are: $\tilde{z}_k = [p_k, RR_k, \sigma_{pk}, \sigma_{\Delta n, \Delta ok}, INA_k]$. $p_k$ is the average value of productivity, $\sigma_{pk}$ is its variance. Note that this is within-firm variation in productivity. Between-firm variation is accomodated via finite mixture modelling. $\sigma_{\Delta n, \Delta ok}$ is the covariance between first-difference employment and first-difference output. $INA_k$ is the inaction rate. $RR_k$ is the reallocation rate (job creation - job destruction rate). To compute job creation and destruction rates as defined by Davis and Haltiwanger (1999), I use $A_t = A_{t-1} u_t$ with $u_t = e^{-\epsilon_t}$. Then:

$$\frac{n_t - n_{t-1}}{\frac{n_t + n_{t-1}}{2}} = \frac{y_t - y_{t-1} u_t}{\frac{y_t + y_{t-1} u_t}{2}}$$

A key idea behind an auxiliary model is that it should be an econometric model that is easy to estimate and parsimonious. For that reason, I do not include the inaction rate in the estimation of the finite mixture. It is by nature a discrete variable: in our dataset it can take only integer between 0 and 6. There are ways to include such a variable in the estimation process (see McLachlan and Peel (2000)) but it significantly increases both computing time and the number of parameters. Then to compute group-specific inaction rate, I use:

$$INA_k = \frac{\sum_{i=1}^N INA_i \hat{I}_{ik}}{\sum_{i=1}^N \hat{I}_{ik}}$$
The discount factor $\beta$ is not identified and fixed. Type-specific parameters are: $\theta_k = [\alpha_k, c_k, \sigma_k]$. Standard deviation of measurement error in log-output $\sigma_o$ and the productivity trend $\mu$ is not considered as type specific.

5.5 Small MonteCarlo Experiments

I simulate $K$ types of firms with type-specific parameters $\theta_k$ and common parameters $\theta_b$. Using the simulated data, I estimate a finite mixture model.

If the number of types $K$ is known, the finite mixture modelling provides unbiased and precised estimates of the proportion of each type and type-specific moments.

The BIC criteria does a poor job at determining the number of types. Increasing the number of types until one can not distinguish 2 different types does a better job. It has some degree of arbitrariness though. In practice I construct confidence interval by using nonparametric bootstrap and recentering. I increased the number of types until the confidence intervals overlap for each type-specific means between two groups.

5.6 Comments

A standard simulated minimum distance estimator of the parameters would consist in simulating an economy with $K$ types of firms. Fitting a mixture of normal on simulated data and then choosing structural parameters to match the mixture parameters in the simulated data and the real data.

What are the advantages of the proposed method. First of all, the computing time is significantly reduced. To give an example in the model that I estimate, finding the structural parameters using my method take at least 10 hours (I use a simulated annealing algorithm). Using the other method, I would need to estimate a mixture at each iteration. There is then three nested loops: 1. minimizing the objective function, 2. solving the model and 3. estimating the finite mixture using the EM algorithm. Furthermore, getting bootstrap standard errors of the structural parameters would basically be intractable unless one parallelize the problem.

Second, estimation of the proportion of each type $\lambda$’s is immediate. When simulating an economy, the identification of the proportion of each type is not trivial. For example, I could impose that there is entry of firms up to the point where the expected value of entry equals the entry cost. But nothing guarantees that it holds in the data. And I cannot estmimate the entry cost because I have no informations on entry / exit.

Thirdly, Heckman and Singer (1984) have emphasized the flexibility of the discrete distribution in that, for sufficient mass points, any distribution can be approximated to a high degree of accuracy. If I simulate an economy with $K$ firms, I basically need to assume that there are effectively $K$ types of firms in the economy. Lastly, identification is somehow more transparent. Finite mixture modelling accommodate between firms differences. I can then focus on identifying the parameters separately for each type of firms. Thus, non-diagonal terms of the weighting matrix have by construction more zeros with the proposed method.
The main drawback of the proposed method is that it assumes that finite mixture modelling provides consistent estimators of the proportions of each type and of type-specific means. This is a drawback with respect to standard minimum distance estimators where one can get consistent structural parameters estimates even if the estimated auxiliary parameters are not consistent. Yet, Monte Carlo experiments show that this assumption is not too strong in our particular application.

6 Structural Estimation: Results

6.1 Auxiliary Model

The four variables chosen for estimating the mixture are: wages, labor productivity, reallocation rate and within-firm variance of productivity. The maximum BIC value occurs for the 3 group model as shown in table 4.

<table>
<thead>
<tr>
<th>Number of Groups</th>
<th>LogLikelihood</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4300.41</td>
<td>8711.64</td>
</tr>
<tr>
<td>2</td>
<td>-4013.01</td>
<td>8255.58</td>
</tr>
<tr>
<td>3</td>
<td>-3943.56</td>
<td>8235.41</td>
</tr>
<tr>
<td>4</td>
<td>-3896.73</td>
<td>8260.48</td>
</tr>
<tr>
<td>5</td>
<td>-3862.31</td>
<td>8310.39</td>
</tr>
</tbody>
</table>

Table 4: Analysis to determine the number of groups

The improvement in the fit of the model of considering 3 types of firm is noticeable: the BIC criteria goes from $-8711$ to $-8325$. As is usually found in the estimation of finite mixture, a small number of types is able to capture unobserved heterogeneity. Table 5 reports the moments that I use in the estimation. 90% confidence interval were constructed by using nonparametric bootstrap: I treat the sample data as if they were the population and carry out a Monte Carlo simulation in which the data are sampled randomly 200 times with replacement. Each bootstrap sample is used to compute a bootstrap estimate. Then the confidence interval is found by considering the quantile of the probability measure induced by bootstrap sampling conditional on the actual sample data.

Firms of type 2 (3) pay high (low) wages, have a high (low) labor productivity, create a lot (few) of jobs and have a high (low) variance. Firms of type 1 are in an intermediary situation. The four moments considered are strongly correlated across types. Basically, high wages firms have a high labor productivity and attract a lot of workers. The structural model will then give an answer to the following questions:

1. Is high labor productivity ($p$) due to high wages ($w$) or to high market power and/or highly decreasing returns ($\alpha$)?

2. Is the high variance of labor productivity ($\sigma_p$) due to high adjustment costs ($c$), a volatile environment ($\sigma$) or simply measurement error ($\sigma_o$)
<table>
<thead>
<tr>
<th>Type</th>
<th>Moments</th>
<th>Average</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Wages</td>
<td>142.9078</td>
<td>138.1042</td>
<td>146.9050</td>
</tr>
<tr>
<td></td>
<td>Reallocation Rate</td>
<td>0.0738</td>
<td>0.0683</td>
<td>0.0810</td>
</tr>
<tr>
<td></td>
<td>Inaction Rate</td>
<td>0.1659</td>
<td>0.1558</td>
<td>0.1724</td>
</tr>
<tr>
<td></td>
<td>Labor Productivity</td>
<td>275.7239</td>
<td>263.7812</td>
<td>286.7397</td>
</tr>
<tr>
<td></td>
<td>Within Variance</td>
<td>29.5564</td>
<td>26.7493</td>
<td>32.1631</td>
</tr>
<tr>
<td></td>
<td>Mixing Proportion</td>
<td>0.5180</td>
<td>0.4295</td>
<td>0.5862</td>
</tr>
<tr>
<td>2</td>
<td>Wages</td>
<td>168.5614</td>
<td>163.0550</td>
<td>174.6642</td>
</tr>
<tr>
<td></td>
<td>Reallocation Rate</td>
<td>0.0811</td>
<td>0.0745</td>
<td>0.0872</td>
</tr>
<tr>
<td></td>
<td>Inaction Rate</td>
<td>0.1617</td>
<td>0.1531</td>
<td>0.1686</td>
</tr>
<tr>
<td></td>
<td>Labor Productivity</td>
<td>383.1015</td>
<td>362.4326</td>
<td>403.7290</td>
</tr>
<tr>
<td></td>
<td>Within Variance</td>
<td>59.1810</td>
<td>52.5581</td>
<td>67.5873</td>
</tr>
<tr>
<td></td>
<td>Mixing Proportion</td>
<td>0.3060</td>
<td>0.2289</td>
<td>0.3882</td>
</tr>
<tr>
<td>3</td>
<td>Wages</td>
<td>127.3323</td>
<td>122.1634</td>
<td>133.3623</td>
</tr>
<tr>
<td></td>
<td>Reallocation Rate</td>
<td>0.0552</td>
<td>0.0494</td>
<td>0.0645</td>
</tr>
<tr>
<td></td>
<td>Inaction Rate</td>
<td>0.2006</td>
<td>0.1942</td>
<td>0.2127</td>
</tr>
<tr>
<td></td>
<td>Labor Productivity</td>
<td>226.6031</td>
<td>215.8291</td>
<td>237.5684</td>
</tr>
<tr>
<td></td>
<td>Within Variance</td>
<td>15.5056</td>
<td>14.1715</td>
<td>17.2028</td>
</tr>
<tr>
<td></td>
<td>Mixing Proportion</td>
<td>0.1760</td>
<td>0.1220</td>
<td>0.2538</td>
</tr>
</tbody>
</table>

Table 5: Average Moments and Bootstrapped 90% Confidence Interval - Auxiliary Model - Number of Firms: 2770 - Number of Periods: 7

6.2 Structural Parameters

[Results are not available yet]

7 Policy Experiments in General Equilibrium

The goal of this section is to revisit the findings of the seminal paper of Hopenhayn and Rogerson (1993) that use calibration to evaluate the effects of firing costs in general equilibrium. I use the estimate of the structural parameters as inputs.

7.1 Framework

The framework is based on Hopenhayn and Rogerson (1993). Two simple modifications are the introduction of firm ex-ante heterogeneity and the random walk assumption. I focus on the stationary equilibrium of the economy. The stationary distribution of firm may not exist in this economy without the introduction of entry/exit in the model. This is because of the random walk assumption that may cause the variance of the distribution to be infinite.

Exit of firm is exogenous: firms die with a probability $\lambda$ that is calibrated. New establishments can also be created, though it is costly. Specically, in each period a new establishment can be created by paying a cost of $c_e$. After paying this cost a realization of the firm level initial condition ($A_0$) and type parameter $k$ is drawn from a distribution $G(A_0, k)$. 

20
Potential entrants enter until the expected value of doing so is driven to zero. This is the free entry condition:

\[ E[V(A, 0; k)] = c_e \]  

(34)

where the expectation is taken with respect to the profitability \( A \) and firm type \( k \).

Let \( \mu(A, n, k) \) be the measure of firms with state \((A, n, k)\). The economy is inhabited by a single representative household. It is comprised by a continuum of members of total size 1. She supplies labor inelastically and switches at no costs between firms (possibly of different types). Preferences over consumption \( c_t \) are given by \( \sum_{t=0}^{\infty} \beta^t u(c_t) \).

A stationary equilibrium consists of a mass of entrants \( M^* \geq 0 \), a measure of firm \( (\mu(A, n, k)) \) such that 1. demand equals supply in the labor market, 2. the state of the economy is reproduced in each period and 3. the free entry condition is satisfied.

7.2 Results

[Results are not available yet]

8 Conclusion

[to be written]
References


23
Appendices

A Solving the Model

In this appendix I consider a more general model with both job creation and destruction costs. In the main text, there is job destruction costs only because it will appears that both kind of costs impact similarly labor demand. The objective function writes:

\[ V(A_t, n_t) = \max_{d_j, j \geq t} E \left\{ \sum_{j=t}^{\infty} \beta^{j-t} \left[ A_j^{1-\alpha} (n_j + d_j) - w(n_j + d_j) - \varphi d_j^+ - \varphi d_j^- \right] \mid n_t, A_t \right\} \]

where \( x^+ = x \) if \( x \) is positive and zero otherwise and \( x^- = -x \) if \( x \) is negative and zero otherwise. Simple calculations shows that the objective function exists if \( \frac{1}{\beta} < \mu \). Then the problem falls in the class of problems anaylsed in Chapter 5 of Stokey et al. (1989). The Bellman equation writes:

\[ V(A, n) = \max_d \left\{ A^{1-\alpha} (n + d)^\alpha - w(n + d) - \varphi d^+ - \varphi d^- + \beta \int V(Ae^\epsilon, n + d) dZ(\epsilon) \right\} \]

with \( \epsilon \to N(\mu, \sigma^2_\epsilon) \). Define \( y = \frac{n}{A} = p^{\frac{1}{\alpha-1}} \) where \( p \) is labor productivity.

\[ V(A_1, n_0) = \max_{n_t, t \geq 1} E_1 \sum_{t=1}^{\infty} \left\{ A_t^{1-\alpha} n_t^\alpha - wn_t - \varphi(n_t - n_{t-1})^+ + \varphi(n_t - n_{t-1})^- \right\} \]

\[ = \max_{n_t, t \geq 1} E_1 \sum_{t=1}^{\infty} A_t \frac{A_t}{A_1} \left\{ \left( \frac{n_t}{A_t} \right)^{-\alpha} - w \frac{n_t}{A_t} - \varphi \left( \frac{n_t}{A_t} - \frac{n_{t-1}}{A_1 e^{\epsilon t}} \right)^+ + \varphi \left( \frac{n_t}{A_t} - \frac{n_{t-1}}{A_1 e^{\epsilon t}} \right)^- \right\} \]

with the constraint that \( \frac{y_0}{e^{\epsilon t}} = \frac{n_0}{A_1} \). Consider \( \lambda \geq 0 \).

\[ V(\lambda A_1, \lambda n_0) = \lambda A_1 \max_{y_t, t \geq 1} E_1 \sum_{t=1}^{\infty} A_t \frac{A_t}{A_1} \left\{ y_t^{-\alpha} - wy_t - \varphi \left( y_t - \frac{y_{t-1}}{e^{\epsilon t}} \right)^+ + \varphi \left( y_t - \frac{y_{t-1}}{e^{\epsilon t}} \right)^- \right\} \]

because \( \frac{y_0}{e^{\epsilon t}} \) and \( \frac{A_t}{A_1} \) for any \( t \) are not modified by such a transformation. Then, \( V \) is homogeneous of degree 1.

Define \( v(x) = V(1, \frac{n-1}{A}) \) with \( x = \frac{n-1}{A} \). \( v \) satisfies the Bellman equation:

\[ V(A, n_{-1}) = \max_n \left\{ A^{1-\alpha} n^\alpha - wn - c |n - n_{-1}| + \beta \int V(Ae^\epsilon, n) dZ(\epsilon) \right\} \]

\[ AV(1, \frac{n-1}{A}) = A \max_n \left\{ \left( \frac{n}{A} \right)^\alpha - wn - c \left| \frac{n}{A} - \frac{n_{-1}}{A} \right| + \beta \int e^\epsilon V(1, \frac{n}{A} e^{-\epsilon}) dZ(\epsilon) \right\} \]

\[ v(x) = \max_y \left\{ y^\alpha - wy - c |y - x| + \beta \int e^\epsilon v(ye^{-\epsilon}) dZ(\epsilon) \right\} \]
The target $\bar{y}$ and $\underline{y}$ satisfy:

$$\alpha \bar{y}^{\alpha - 1} - w + \bar{c} + \beta \int v'(\bar{y}e^{-\epsilon})dZ(\epsilon) = 0$$

$$\alpha \underline{y}^{\alpha - 1} - w - \underline{c} + \beta \int v'(\underline{y}e^{-\epsilon})dZ(\epsilon) = 0$$

and it translates in the optimal policy:

$$y(x) = \begin{cases} 
\frac{y}{x} & \text{if } x < y \\
\frac{y}{x} & \text{if } y < x < \bar{y} \\
\bar{y} & \text{if } x > \bar{y}
\end{cases}$$

from which can be recovered the employment policy:

$$n(A, n_{-1}) = \begin{cases} 
\frac{yA}{n_{-1}} & \text{if } n_{-1} < yA \\
\frac{n_{-1}}{yA} & \text{if } yA < n_{-1} < \bar{y}A \\
\bar{y}A & \text{if } n_{-1} > \bar{y}A
\end{cases}$$

Define $d(y) = \int e^{-\epsilon}v(ye^{-\epsilon})dZ(\epsilon)$. $d'(y) = \int v'(ye^{-\epsilon})dZ(\epsilon)$. Express $v$ and $v'$ as a function of $x$ and the optimal policy:

$$v(x) = \begin{cases} 
\frac{\alpha x^{\alpha - 1} - w \bar{y} - \underline{c}(y - x) + \beta d(\bar{y})}{\underline{c}} & \text{if } x < y \\
x^{\alpha - w}x + \beta d(x) & \text{if } y < x < \bar{y} \\
\bar{y}^{\alpha - w} - \bar{c}(x - \bar{y}) + \beta d(\bar{y}) & \text{if } x > \bar{y}
\end{cases}$$

$$v'(x) = \begin{cases} 
\frac{\underline{c}}{\alpha x^{\alpha - 1} - w + \beta d'(x)} & \text{if } x < y \\
\frac{-\underline{c}}{\bar{y}^{\alpha - w} - \bar{c}(x - \bar{y}) + \beta d(\bar{y})} & \text{if } x > \bar{y}
\end{cases}$$

$$v'(y^{-\epsilon}) = \begin{cases} 
\frac{\underline{c}}{\alpha y^{\alpha - 1}e^{(1-\alpha)} - w + \beta d'(y^{-\epsilon})} & \text{if } \epsilon > \log \left(\frac{y}{\underline{y}}\right) \\
\frac{-\underline{c}}{\alpha y^{\alpha - 1} (1-\alpha) - w + \beta d'(y^{-\epsilon})} & \text{if } \epsilon < \log \left(\frac{y}{\underline{y}}\right)
\end{cases}$$

It follows:

$$d'(y) = \underline{c} \left(1 - \Phi \left(\frac{\log \left(\frac{y}{\underline{y}}\right) - \mu}{\sigma}\right)\right) - \bar{c} \Phi \left(\frac{\log \left(\frac{y}{\bar{y}}\right) - \mu}{\sigma}\right)$$

$$+ \alpha \bar{y}^{\alpha - 1} \int_{\log \left(\frac{\underline{y}}{\bar{y}}\right)}^{\log \left(\frac{\bar{y}}{\underline{y}}\right)} e^{(1-\alpha)\phi} \left(\frac{\epsilon - \mu}{\sigma}\right) d\epsilon$$

$$- w \left[\Phi \left(\frac{\log \left(\frac{y}{\underline{y}}\right) - \mu}{\sigma}\right) - \Phi \left(\frac{\log \left(\frac{y}{\bar{y}}\right) - \mu}{\sigma}\right)\right]$$

$$+ \beta \int_{\log \left(\frac{\underline{y}}{\bar{y}}\right)}^{\log \left(\frac{\bar{y}}{\underline{y}}\right)} d'(y^{-\epsilon}) \phi \left(\frac{\epsilon - \mu}{\sigma}\right) d\epsilon$$
Define \( g(y, \epsilon) = d'(ye^{-\epsilon}) = \int v'(ye^{-\epsilon})e^{-\epsilon^2}d\epsilon \). \( g \) is the convolution of a normal density and an integrable function. Using Theorem 9 on p. 59 in Lehmann (1986), which applies to any distribution belonging to the exponential family, it follows that \( g \) is analytic. One can write:

\[
g(y, \epsilon) = \sum_{i=0}^{\infty} b_i(y)e^{i\epsilon(1-\alpha)}
\]

in which the functions \( b_0, b_1, \ldots \) are defined on the real line.

\[
\sum_{i=0}^{\infty} b_i(y) = \zeta \left( 1 - \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) \right) - \overline{\epsilon} \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right)
\]

\[
+ \alpha y^{a-1} \int_{\log \left( \frac{y}{\hat{y}} \right)}^{\log \left( \frac{y}{\hat{y}} \right)} e^{i\epsilon(1-\alpha)} \phi \left( \frac{\epsilon - \mu}{\sigma} \right) d\epsilon
\]

\[
- w \left[ \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) - \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) \right]
\]

\[
+ \beta \sum_{i=0}^{\infty} b_i(y) \int_{\log \left( \frac{y}{\hat{y}} \right)}^{\log \left( \frac{y}{\hat{y}} \right)} e^{i\epsilon(1-\alpha)} \phi \left( \frac{\epsilon - \mu}{\sigma} \right) d\epsilon
\]

\[
\sum_{i=0}^{\infty} b_i(y) = \zeta \left( 1 - \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) \right) - \overline{\epsilon} \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right)
\]

\[
+ \alpha y^{a-1} \int_{\log \left( \frac{y}{\hat{y}} \right)}^{\log \left( \frac{y}{\hat{y}} \right)} e^{i\epsilon(1-\alpha)} \phi \left( \frac{\epsilon - \mu}{\sigma} \right) d\epsilon
\]

\[
- w \left[ \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) - \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) \right]
\]

\[
+ \beta \sum_{i=0}^{\infty} b_i(y) \int_{\log \left( \frac{y}{\hat{y}} \right)}^{\log \left( \frac{y}{\hat{y}} \right)} e^{i\epsilon(1-\alpha)} \phi \left( \frac{\epsilon - \mu}{\sigma} \right) d\epsilon
\]

\[
\sum_{i=0}^{\infty} b_i(y) = \zeta \left( 1 - \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right) \right) - \overline{\epsilon} \Phi \left( \frac{\log \left( \frac{y}{\hat{y}} \right) - \mu}{\sigma} \right)
\]

\[
+ \alpha y^{a-1} K_1(y)
\]

\[
- w K_0(y)
\]

\[
+ \beta \sum_{i=0}^{\infty} b_i(y) K_i(y)
\]
\[
\sum_{i=0}^{\infty} b_i(y)(1 - K_i(y) \beta) = \epsilon \left( 1 - \Phi \left( \frac{\log \left( \frac{y}{\bar{y}} \right) - \mu}{\sigma} \right) \right) - \overline{\epsilon} \Phi \left( \frac{\log \left( \frac{\bar{y}}{\bar{y}} \right) - \mu}{\sigma} \right) \\
+ \alpha y^{a-1} K_1(y) \\
-w K_0(y)
\]

with

\[
K_0(y) = \left[ \Phi \left( \frac{\log \left( \frac{y}{\bar{y}} \right) - \mu}{\sigma} \right) - \Phi \left( \frac{\log \left( \frac{\bar{y}}{\bar{y}} \right) - \mu}{\sigma} \right) \right]
\]
\[
K_1(y) = \int_{\log(\frac{y}{\bar{y}})}^{\infty} e^{\epsilon(1-\alpha) \phi \left( \frac{\epsilon - \mu}{\sigma} \right)} d\epsilon \\
= e^{(1-\alpha)\mu + (1-\alpha)\sigma^2/2} \left[ \Phi_{\mu(1-\alpha)\sigma^2, \sigma^2}(\log \left( \frac{y}{\bar{y}} \right) ) - \Phi_{\mu(1-\alpha)\sigma^2, \sigma^2}(\log \left( \frac{\bar{y}}{\bar{y}} \right) ) \right]
\]

Substituting into the FOC for destroying jobs:

\[
\alpha y^{a-1} - w + \overline{\epsilon} + \beta b_0(\bar{y}) = 0
\]
\[
\alpha y^{a-1}(1 - K_0(\bar{y}) \beta) - w(1 - K_0(\bar{y}) \beta) + \overline{\epsilon}(1 - K_0(\bar{y}) \beta) \\
+ \beta \epsilon (1 - \Phi_{\mu, \sigma}(\log(r))) - \beta \overline{\epsilon} \Phi_{\mu, \sigma}(0) + \beta \alpha y^{a-1} K_1(\bar{y}) - \beta w K_0(\bar{y}) = 0
\]
\[
\alpha y^{a-1}(1 + (K_1(\bar{y}) - K_0(\bar{y})) \beta) - w + \overline{\epsilon}(1 - (\Phi_{\mu, \sigma}(0) + K_0(\bar{y}) \beta) \\
+ \beta \epsilon (1 - \Phi_{\mu, \sigma}(\log(r))) = 0
\]

Substituting into the FOC for creating jobs:

\[
\alpha y^{a-1} - w - \epsilon + \beta b_0(\bar{y}) = 0
\]
\[
\alpha y^{a-1}(1 - K_0(\bar{y}) \beta) - w(1 - K_0(\bar{y}) \beta) - \epsilon(1 - K_0(\bar{y}) \beta) \\
+ \beta \epsilon (1 - \Phi_{\mu, \sigma}(0)) - \beta \overline{\epsilon} \Phi_{\mu, \sigma}(-\log(r)) + \beta \alpha y^{a-1} K_1(\bar{y}) - \beta w K_0(\bar{y}) = 0
\]
\[
\alpha y^{a-1}(1 + (K_1(\bar{y}) - K_0(\bar{y})) \beta) - w + \epsilon(1 + K_0(\bar{y}) - \Phi_{\mu, \sigma}(0) \beta - 1) \\
- \beta \overline{\epsilon} \Phi_{\mu, \sigma}(-\log(r)) = 0
\]

\[
\alpha y^{a-1}(1 + (K_1(\bar{y}) - K_0(\bar{y})) \beta) = w + \epsilon(1 - (1 - \Phi_{\mu, \sigma}(-\log(r))) \beta) + \beta \overline{\epsilon} \Phi_{\mu, \sigma}(-\log(r))
\]