n° 2003-06

The Effects of Payroll Tax Subsidies for Low Wage Workers on Firms Level Decisions

B. CRÉPON[1]
R. DESPLATZ[2]

Les documents de travail ne reflètent pas la position de l'INSEE et n'engagent que leurs auteurs.

Working papers do not reflect the position of INSEE but only the views of the authors.

The Effects of Payroll Tax Subsidies for Low Wage Workers on Firms Level Decisions

Bruno Crépon (CREST)
Rozenn Desplatz (Direction de la Prévision)

Abstract

We study the effect of the large increase of payroll tax subsidies for low-wage workers that occurred in France in 1995. We compute the ex ante changes in average labor costs in 1994 solely due to the changes in the tax subsidies and consider it as a treatment variable. We extend the Rubin causal framework to the case of a continuous treatment, define the ensuing parameters of interest and generalize the propensity score property. We find that payroll tax subsidies had a strong effects on employment and on other firms outcomes, stock of capital, share of unskilled workers and average labor cost.

1. Introduction

Payroll tax subsidies for low-wage workers were set up in France in 1993 in order to fight against the disappearance of unskilled jobs. The implementation of such policy is the opportunity to examine the effects on employment of policies aimed at skewing the wage distribution at the bottom of this distribution like changes in the minimum wage. The initial scheme of 1993 was considerably strengthened between 1995 and 1996. The purpose of this paper is to evaluate the effect of this large increase on employment and other firm level variables.

Detecting and measuring the effect of such policies as usually been difficult and their importance on labor market outcomes is still a matter of considerable debate (see the discussions surrounding the Card and Krueger', 1995 study). Some studies are based on the estimation of structural models as Meyer and Wise (1983) and Laroque and Salanié (2002), but most studies focus on policy changes that they treat as natural experiment. One important difficulty faced by researchers however is to find a proper control group. Some have used aggregate data on employment in some industries in different states as in Card and Krueger (1994). On the opposite Abowd, Kramarz, Margolis and Philippon, (2001), use individual employee data and define their control group as workers just above the future level of minimum wage. Kramarz Philippon (2001) uses the same idea to study the effect of payroll tax subsidies in France which is the purpose of our paper.

Our study is based at the firm level. We view this as a strong advantage as it allows us to account for substitutions between various types of workers, which is likely to be very important when implementing such a policy. This allows us also to examine the effect of the policy on employment as well as on other firm level variables. For example we examine the effect of the policy on capital stock and as far as we know this type of issue as never been examined empirically.

We follow the recent literature on program evaluation. The situation we face is however somewhat different. The tax rule changes between t_0 and t_1 from T_0 to T_1. For a variable of interest, say y_i, we would like to compare at date t_1 the output of the firm with tax rule T1, say $y_{i,t_1}(T_1)$, with what would have been its output if the tax rule had remained unchanged, say $y_{i,t_1}(T_0)$. Thus we would like to measure $c_i = y_{i,t_1}(T_1) - y_{i,t_1}(T_0)$, or similarly considering changes over the period t_1, t_0 $c_i = \left[y_{i,t_1}(T_1) - y_{i,t_0}(T_0) \right] - \left[y_{i,t_1}(T_0) - y_{i,t_0}(T_0) \right]$. This is thus very close to the usual framework of program evaluation: the causal effect is defined has the difference between the observed output $\left[y_{i,t_1}(T_1) - y_{i,t_0}(T_0) \right]$ and an unobserved one $\left[y_{i,t_1}(T_0) - y_{i,t_0}(T_0) \right]$. The identification problem in the program evaluation literature can be solved in various ways, depending on assumption.[1] One typical solution is to find for each individual who benefited from the policy another individual close enough who did not benefited from the policy. The corresponding assumption is that selectivity is only on observables. This assumption has received a lot of attention recently and various estimators have been proposed surrounding the previous matching idea, and their asymptotic properties have been extensively studied.[2] The situation we face is however different because in our case T_0 and T_1 are tax rules which apply to everybody and not treatment. Finding firms that where not affected by the policy is therefore difficult.

---

[1] See Heckman, Lalonde and Smith (1999) for a comprehensive survey
[2] See Rubin (1977), Rosenbaum and Rubin (1983), Heckman Ichimura Smith and Todd (1997,1998), Hahn (1998) and Hirano, Imbens and Ridder (2000).

To solve this problem, we compute for each firm in 1994 the potential change in average labor costs, namely the *ex ante* reduction in average labor costs, say t_i, solely due to changes in tax subsidies occurring between 1994 and 1997. This was computed using exhaustive wage information at the employee level. Given the change in the tax rule this variable is continuously distributed form 0 to almost 10 percent. We define firms not affected by the policy as firms with low values of the exante reduction. The unobserved counterfactual change $\left[y_{i,t_1}(T_0) - y_{i,t_0}(T_0) \right]$ is thus constructed on the basis of information concerning firms close enough to the firm considered and with arbitrary low values of this ex ante reduction.

This situation requires however to define a suitable framework to handle the case of a continuous treatment variable. This necessitates adapting the Rubin causal framework (Rubin 1974,1977). The binary causal model has already been extended to the case of multiple treatments by Lechner (2001) and Imbens (2000). However we need some more adaptations as our treatment variable is continuous. One of these adaptations is to introduce potential outcomes corresponding to each treatment values $\Delta y_i(t)$. The unobserved quantity of interest is therefore defined as $\left[y_{i,t_1}(T_0) - y_{i,t_0}(T_0) \right] = \Delta y_i(0)$ and the parameter of main interest is $E[\Delta y_i - \Delta y_i(0)]$. We show how it can be identified under the assumption of selectivity on observables. The function $E(\Delta y_i | x_i = x, t_i = t) = g(x,t)$ is identified by the data. Under the assumption of selectivity on observables this function is also equal to $E(\Delta y_i(t) | x_i = x)$ and the counterfactual is thus $E(\Delta y_i(0) | x_i) = g(x_i, 0)$. We also extend the Rosenbaum and Rubin (1983) propensity score property to this new setting and propose an estimation method based on the implementation of non parametric series estimators (Andrews (1991)).

Using a simple model of labor demand at the firm level and wage formation we show that the ex ante reduction is the right variable to assess the effect of the policy, as long as workers whose wages are directly concerned by the policy are perfectly complementary. Our model also allows us to deal with the issue of selectivity and justify the assumption of selectivity on observables on which our estimation will rely. One important aspect is that the policy apply to every firm : there is no self-selection in a program. The average level of wages, which is the result of a long process and is strongly dispersed among firms, acts as a random assignment to policy. The selectivity problem comes here from the fact that firms have characteristics that influences both the ex ante labor cost reduction and the evolution of employment. For example, low productivity firms are likely to be those most hit by international competition and thus are more likely to reduce their employment, but they are also likely to be the firms with lots of low paid workers benefiting therefore most from the policy.

The data used in our empirical analysis comes from the matching of two sources: the BRN (*Bénéfice Réels Normaux*) and the DADS (*Déclarations Annuelles de Données Sociales*). The BRN is a firm level file providing most accounting variables as well as employment. It accounts for 60% of the firms and 94% of the turnover. Firms with a turnover of more than 3.5 million francs are required to fill the corresponding declarations annually. The DADS is an employee level file. Firms fill monthly declarations about each of their employee, providing information about hours, wages, occupation, age and gender. Statistical processing of the DADS became exhaustive starting in 1993. The file at our disposal is the aggregation over the year of all the monthly declarations for a given employee-firm pair. At present, the employees covered by the DADS represent almost 80% of dependent employment. This file is used to compute the ex ante reduction of labor cost at the employee level which is then aggregated at the firm level. We also use this file to obtain information about the heterogeneity of the workforce using gender, skills and age criteria.

In the second part of this paper, we begin by setting out the statutory framework for the employer's payroll tax subsidies for the low-wage workers. In Section 3, we use a model of labor demand and wage equation to explain the relationship between potential outcomes and the *ex ante* labor cost reduction. In Section 4, we describe the data sources used to construct the samples as well as the definition of the variables of interest and the control variables chosen and present preliminary results, assuming linearity and homogeneity of the effect of the ex ante reduction in payroll tax subsidies. In Section 5, we present our extension of the Rubin causal model to the case of a continuous treatment. Results are presented in Section 6. We find that the payroll tax reduction for low wage workers have a strong effect on employment and examine the robustness of this finding in several dimension. We especially examine the sensitivity of our result to the set of control variables used and find that the key control variable is the level of total factor productivity. Results are used to provide an estimation of the number of jobs created or saved in France by the policy. Our methodology is applied to other firm level variables. We find numerous evolutions indicating that payroll tax subsidies had a strong effect on employment through substitutions between inputs. Section 7 concludes.

2. Payroll tax subsidies for low wage workers in France

2.1 Measures taken since 1993 to reduce payroll taxes for the lower paid
Payroll taxes are an increasing function of wages in France. The employer-paid contribution is substantial: at the minimum wage level it remained roughly stable over time, around 39.00% up to 1993. From then on, France implemented various measures to reduce payroll tax rates for low-wage workers (Figure 1). The first measure was introduced on 1 July 1993. It consisted of an exemption of 5.4 points in employers' payroll taxes for workers with monthly wage below 1.1 times the minimum wage and 2.7 points for those with monthly wage between 1.1 and 1.2 times the minimum wage. This program was extended in 1995. First, starting January 1, the two previous wage thresholds were increased to 1.2 and 1.3 times the minimum wage. Second, starting September 1, a new schedule was introduced. At the minimum wage, the total exemption was set to 18.2 points. From that point on, the subsidy decreased linearly for all wages up to 1.2 times the minimum wage, at which point the reduction was equal to 5.4 points. Between 1.2 and 1.3 times the minimum wage the subsidy was constant and equal to 2.7 points. The two programs were merged on October 1 1996 into a single one. The new reduction was defined as a linear decreasing function of the wage. The total exemption at the minimum wage remained unchanged, equal to 18.2 points. The subsidy then decreased linearly for all wages up to 1.33 times the minimum wage, at which point the reduction was zero. Initially planned to last only until 1 January 1998, this single degressive reduction has now been made permanent and has even been extended in the wake of the mandatory reduction of the workweek to 35 hours.

[Figure 1 – Rules for the reduction in payroll taxes form 1993 to 1998]

2.2 The ex ante reduction of average labor cost associated to payroll tax subsidies

For each firm we compute what we call   the average *ex ante* labor cost reduction, i.e., the changes in total 1994 compensation costs, directly induced by changes in payroll taxes over the 1994-1997 period. This variable measures the propensity for a firm to benefit from the 1995 and 1996 subsidies on the basis of its 1994 wage distribution. It does not depend on changes in average labor costs, which would be due to changes in employment structure or wages between 1994 and 1997 and only includes ex ante changes due to the change in the policy.

More precisely, using the gross wage $w_{j,i,94}$ of employee j in firm i in 1994, taken from the DADS (Annual Declarations of Social Data), we calculate two labor costs according to the payroll tax schedules for 1994 (L94) and 1997 (L97). That is:

$$c_{j,i,94}(L_{94}) = (1 + T_{94}(w_{j,i,94}))w_{j,i,94}$$
$$c_{j,i,94}(L_{97}) = (1 + T_{97}(w_{j,i,94}))w_{j,i,94}$$

where $T_{94}(w_{j,i,94})$ (and $T_{97}(w_{j,i,94})$) is the payroll tax associated with the gross wage level $w_{j,i,94}$ and the 1994 tax rule (respectively the 1997 tax rule). The *ex ante* reduction in the average labor cost at the firm level is then equal to:

$$t_i = \frac{\sum_{j \in i}(c_{j,i,94}(L_{97}) - c_{j,i,94}(L_{94}))}{\sum_{j \in i} c_{j,i,94}(L_{94})}$$

This reduction ranges from 0% for firms having no worker paid below 1.33 times the minimum wage in 1994 to 9.5% for firms all of whose workers are paid the minimum wage. It increases with the proportion of lower-paid workers in the firm.

Figure 2 and table 1 present different features of the distribution of the *ex ante* reduction in average labor cost for manufacturing firms (energy and agricultural sectors excluded) and non manufacturing firms (except financial sectors). They show that this distribution is heavily concentrated around low values for the *ex ante* reduction but that the reduction is nevertheless substantial (between 1 and 6%) for around half the firms. This is a nice feature of these distributions as the principle of our evaluation will be to find for each firm a firm close in some sense that received a marginal treatment. What the distribution shows is that the poll in which we will have to choose the counterfactual is very large.

[Table 1 - Ex ante reduction of labor cost]

[Figure 2 - Distribution of the ex ante reduction of labor cost in manufacturing and non manufacturing sectors.]

3. Identification of payroll tax subsidies effect using the ex-ante average labor cost reduction

The approach we adopt to identify the effect of payroll tax subsidies consists of regressing the evolution between 1994 and 1997 of a certain number of variables of interest (employment, average labor cost, skill structure, etc.) on the *ex ante* cost reduction and a set of control variables. In this section, we examine the conditions for the validity of such an approach.

To this purpose, we rely on a simple model of labor demand and wage equation with several types of workers. The model is set out in details in Appendix A. Assuming monopolistic competition in the product markets, the demand for labor is written as:

$$\log N = \Sigma_{N,N} \pi_N \log c_N + \Sigma_{N,K} \pi_K \log c_K + u$$

where N and c_N are the vectors of employment and costs of the different types of workers. π_N is the diagonal matrix of their cost shares relative to total cost (including the cost of capital) and $\Sigma_{N,N}$ and $\Sigma_{N,K,}$ are the matrices of Allen substitution elasticities reduced by the price elasticity of demand and u is the vector of perturbations that include technology and demand shocks.

The wage equations are simply written as mark-ups above reservation wages:
$$\log w = \log \overline{w} + \gamma u + v$$
Where w is the vector of gross wages and $\overline{w}$ is the vector of reservation wages. The mark-up is function of the technology shocks *u* and firm specific components. Mark-ups may change over time. For example we could figure that an increase in the tax rate leads to a reduction of the mark-up.

Costs and gross wages are linked by the relationship: $c_{N_j} = \left(1 + T(w_j)\right) w_j$, where $T(\ )$ is the payroll tax rate function. It is precisely this function which has been changed by the introduction of payroll tax subsidies for low wage workers.

We show in Appendix A that based on the assumption that the categories of employees affected by payroll tax subsidies are perfectly complementary, the changes in the variables of interest from 1994 to 1997, such as growth in gross wages or in the employment of different categories of workers, can be written as a function of the *ex ante* cost reduction t_i and firm specific characteristics:

$$\Delta y_i = \Delta y \left(\Sigma_{NNi}, \Sigma_{N,Ki}, \Delta \log(\overline{w}), \pi_i, \Delta \log(c_{Ki}), \gamma, \Delta u_i, \Delta v_i, t_i \right)$$
$$= \Delta y \left(\Phi_{\Delta yi}, t_i \right)$$

We also derive the set of factor entering in the *ex ante* labor cost reduction :

$$t_i = t \left(\pi_i, u_{i,94}, v_{i,94}, \gamma, \log \overline{w}_{94} \right)$$
$$= t \left(\Phi_{ti} \right)$$

We thus derive from this model the three following results:

   (i) the *ex post* variables of interest can be defined as functions of the *ex ante* average labor cost reduction once it is assumed that the categories of workers affected by payroll tax subsidies are complementary;
   (ii) the variables of interest depend on numerous factors that also affect the *ex ante* labor cost reduction
   (iii) there exist factors entering in the *ex ante* reduction but not in the *ex post* variable.

These are the individual firm effects $u_{i,94}$ and $v_{i,94}$ which influence the distribution of wages in 1994 and hence the *ex ante* labor cost reduction but have no direct effect on the *ex post* variables since the latter are taken in first difference.

These results show that it is possible to evaluate the effect of payroll tax subsidies on the changes of the variables of interest by comparing firms with different *ex ante* labor cost reduction. Given the existence of numerous factors that are common to the *ex ante* reduction and to the variables of interest, it is necessary to make comparisons between firms showing similar characteristics. In practice, we approach these characteristics through a set of observable variables. We thus make the following additional assumption:

Assumption (A-1): $\exists\, X$ such that $\Phi_{\Delta y} \perp \Phi_t | X$ and $l(t|X)$ is non degenerated

The first part of the assumption ensures that the control variables X_i make it possible to eliminate the selectivity bias. The second part states that the X_i variables do not totally explain the *ex ante* labor cost reduction. It thus ensures that there remains a source of random variation specific to the *ex ante* cost reduction. Firms with similar X can therefore benefit from different *ex ante* reductions.

One key issue to find control variables satisfying assumption A-1 concerns the unobserved terms in the demands and supplies of labor for various types of workers. With respect to demand it is likely that there is some tendency in these terms. This is consistent with the usual explanations of the decline in the demand for unskilled workers. We model them as $u_{it} = u_{it}^{(1)} + u_i^{(2)}t$ where $u_{it}^{(1)}$ is distributed as a random walk. This implies that there is a strong positive correlation between the residual in 1994 u_{i94} and its variation $u_{i97} - u_{i94}$ related to the firm-specific trend. To control for this common source of variability, we introduce the 1994 levels, as well as their past changes of the labor productivity and the capital labor ratio. We also introduce past values of the level and the changes in value added, operating income capital ratio and the share of labor cost in the set of control variables.

We assume that there is no correlation between the firm specific effect in the wage equation in 1994 v_{i94} and its variation between 1994 and 1997 $v_{i97} - v_{i94}$. We further assume that the term v_{i94} does vary substantially between firms. For this, we refer to the results of Abowd, Kramarz and Margolis (1999), on matched enterprise and employee data, which show that the firm-specific component of variance is a major source of variation in labor compensation costs. This is a central point for the identification problem: the existence of a great deal of variability in the wages paid by firms to apparently "identical" workers acts as a random assignment to the policy.

Besides the error terms, some factors are common to the two sets Φ_Δ and Φ_t. It is the case for the shares of the various types of workers. They can be directly measured at the firm level by the shares in total hours worked for 18 categories of employees, created using sex, skilled level (3) and age groups (3).

We introduce competition variables that may enter wage mar-ups. These are measured at individual level using the markup and at the sector level by the import and export ratios as well as by the entry and exit rates. Last we introduce the cost of capital that may also enter in the wage equation we measure it directly measured using firms accounting information, following Auerbach (1983).

We also add other firm level financial variables in the set of control variables: the share of debt in total financing and the *ex ante* variation in the cost of capital, measuring the variation induced uniquely by changes in taxation over the period. There was indeed a major change in taxation over the period 1994-1997 (mainly an increase in the corporation tax rate from 33.3% to 41.7%).

One other problem concern the existence of general equilibrium effects that may affect firms indirectly a change in $\overline{w}$. One type of general equilibrium effect that may be at work here is an equilibrium wage effect. Firms may be affected by the policy directly by changes in their labor cost but also indirectly by the fact that the labor cost of other firms was affected which may induce changes in reservation wages at the aggregate level. The upward pressure on the demand for low skill worker may induce an upward shift in their reservation wage, which will affect all firms. Similarly for more skilled workers the changes in their labor demand, which can be upward or downward depending on the relative strength of substitution and profitability effects, may induce changes in reservation wages.

What we assume here is that these equilibrium wage effects will not be observed during the short period covered by our study. Moreover recent attempt to assess the strength of these equilibrium wage effect have found them to be rather small (Blundell, Dias, Meghir and Van Reenen 2002). In examining our results we will find further evidence that in our case this effects should be small although we cannot identify them.

Given this assumption, $\Delta y_i\left(\Phi_{\Delta y}, 0\right) = \Delta y_i^0$, where Δy_i^0 is the evolution that would have been observed in the absence of the policy.

4. Data and preliminary results

4.1 Data
We mainly use two datasets: the BRN file (*Bénéfice Réels Normaux*) which is a firm level dataset and the DADS (*Déclarations Annuelles de Données Sociales*), which is an employee level dataset. The BRN declarations are filled annually by firms with a turnover of more than 3.5 million francs (1992 threshold) whose commercial, industrial or craftwork activity is carried out for lucrative purposes. The files record firms accounting information as well as total employment. It covers 60% of firms and accounts for 94% of all total turnover. The DADS are filled each month by any firm employing workers. They cover all employers and their employees with the exception of paid agricultural workers and civil servants. The available files are the aggregation over a given year of all declarations for a given firm employee pair. Statistical processing of the DADS became exhaustive starting 1993. However the quality of information related to hours, which we use, is low in 1993. We therefore only use information in these files starting 1994. In the most recent files, the employees covered by the DADS cover almost 80% of dependent employment.

4.1.1 Variables computed from the BRN
Employment
The measure of employment in the BRN file is employment at the end of the year. It is a full-time equivalent measure that accounts for part time workers.
Value added
Value added is defined as the difference between production and materials, added to production subsidies minus value added tax and other accrued taxes or production subsidies. It is divided by a value added price index at the two-digit level of the French industrial classification taken from national accounts. We therefore do not have a real measure of value added in volume, as price information at the firm level is not available.
Average labor cost
The average labor cost is equal to total labor compensation costs (wages and payroll taxes) divided by total employment.
Real capital stock

The real capital stock measure is computed as the gross book value of fixed assets including construction, technical installations and other fixed assets. It is adjusted for inflation assuming all the stock was bought in one time at a date computed as the difference between the considered year and the age of the capital stock. This age is itself defined as the product of an assumed life time of 16 years and the ratio of the net to gross book value.

Operating income

Operating income per unit of capital is computed as the difference between value added and total labor compensation.

Cost of capital

The measurement of the user cost of capital is firm-specific following Auerbach (1983). It is defined as the sum of terms reflecting economic depreciation of assets and inflation and the mean of the costs of debt and equity, weighted by their respective share in the firm's financial structure. The debt cost is measured as the observed average interest rate.

Ex ante change in the cost of capital

We calculate the change in the capital cost for each firm in 1994, which is solely due to changes in taxes over the period 1994-1997. It is obtained by comparing two capital costs, which are computed in 1994 from the firm accounts using the 1994 and 1997 tax rates.

Total cost

Total cost at the firm level is defined as the sum of total labor compensation and the product of the cost of capital and the capital stock. Using this information, we also calculate the wage share in total costs and the ratio between value added and total costs which approximates the mark-up.

Total factor productivity growth

Total factor productivity growth is defined as the difference between the growth rate of real value added and the cost share-weighted average of the growth rates of labor and real capital.

Entry and exit rates

Entry and exit rates are computed at a three digit level following, Dunne, Roberts and Samuelson (1988). We use BRN files for adjacent years and assume that real entry and exit are well measured by the entry in and exit from the files.

Variables from the DADS

The DADS files provide information for each employee in each firm, on total net nominal earnings during the year, the individual's sex, date of birth, occupation, number of days and hours worked during the calendar year. It also provides the status of the employee: full-time, part-time and intermittent workers. We selected only full-time and part-time employees. This information is used to compute the ex ante reduction in average labor cost detailed in Section 2 as well as the heterogeneity of the workforce.

The heterogeneity of the workforce is measured as the share in total hours worked of various types of workers. We consider 18 categories of employees, starting from three skill levels, three age levels and gender.

Occupation

The DADS provide information about occupation at a 2-digit level. We create three skill levels: high skilled workers (included business heads, senior executives and intermediate occupations), skilled workers (skilled blue- and white-collar workers) and unskilled workers (unskilled blue- and white-collar workers).

Age

We define three age groups: youngest employees (less than 25 years old), middle aged employees (between 25 and 49 years old) and oldest employees (50 years old and more).

4.1.3 Outcome variables

Outcome variables are the changes over the period 1994-1997 (the years immediately preceding and following the intensification of the payroll tax subsidies) of some variables of interest. They are calculated as the logarithmic differences of these variables in 1997 and 1994. We mainly focus on employment, but we also consider value added, average labor cost, capital labor ratio, labor productivity, productivity of capital and the share in total hours worked of unskilled workers.

4.1.4 Control variables

The control variables we introduce are the level of some variables in 1994 and, for some of them, their average changes over a past period. In this case, the changes considered are the changes between the first year a firm was present in the BRN files starting 1989, and 1994. For past characteristics of firms reflecting demand shocks, we consider the logarithm of value added and its average change. For past characteristics reflecting technical progress shocks, we take the logarithm of labor productivity as well as the average growth in total factor productivity, the logarithm and the average difference of the logarithm of the capital labor ratio. The competition variables are measured at individual level by the markup (absolute level and changes). At sector-level, we use the import and export ratios (two-digit level, available from the national accounts), along with the entry and exit rates (three digit level). The financial variables are the user cost of capital, the share of debt in total financing as well as the *ex ante* variation in the cost of capital. Finally, we introduce the shares in hours worked for the 18 categories of employees measured using the DADS in 1994. The set of control variables also includes the share of wages in total costs at individual firm level and the average cost and share of unskilled workers at the two-digit sector level. This last variable is obtained by aggregation of the DADS information.

The construction of the sample is set out in details in appendix C. It covers 87,720 firms, of which 34,371 (39%) are in manufacturing and 53,349 (61%) in non manufacturing. These firms employ a total of 3,772,941 people, of whom 2,053,777 (54%) work in manufacturing and 1,719,164 (46%) in non manufacturing.

4.2 Preliminary results on employment and payroll tax subsidies.

A first way to examine the effect of payroll tax subsidies on employment is to regress the growth rate of employment on the *ex ante* labor cost reduction and the different control variables:

$$(1) \qquad \Delta y_i = a\, t_i + x_i b + u_i,$$

where Δy_i represents the growth rate of employment between 1994 and 1997, t_i the average *ex ante* labor cost reduction, x_i the control variables and u_i the error term. In this initial specification, the impact of a marginal increase in the *ex ante* labor cost reduction is assumed to be constant (linear relationship) and identical from one firm to another.

Two parameters of interest can be deduced from regression (1). The first one is the slope parameter *a* which indicates the sensitivity of employment to the *exante* reduction in average labor cost. It has no structural interpretation. The elasticity of employment to the *ex ante* cost reduction combines various parameters of the labor demand that cannot be disentangled (see appendix A). The second parameter that can be deduced from regression (1) is an evaluation of the increase in the growth rate of employment induced by a payroll tax reduction : for each firm, the growth attributable to payroll tax subsidies is obtained by comparing the *ex post* situation of firms Δy_i with which that would have prevailed in the absence of tax subsidies $\Delta y_i(0) = x_i b + u_i$. It is then defined by:

$$(2) \qquad E[\varpi_i(\Delta y_i - \Delta y_i(0))] = a\, E[\varpi_i t_i]$$

where ϖ_i is a normalized weighting variable: $\varpi_i = N_i / E(N_i)$ and N_i denotes employment at the firm level in 1994.

Results are presented in table 2. The first two columns, labeled as "Elasticities", provide the slope parameters *a* for both manufacturing and non manufacturing sectors. The last two columns show the growth rates attributable to the tax subsidies. The results clearly show a positive and significant relationship between employment and the *ex ante* labor cost reduction. An increase in the *ex ante* reduction of 1% leads to a rise in employment of 1.6% in manufacturing and 1.8% in non manufacturing. The corresponding growth rates of employment attributable to the policy are respectively 1.3% and 2.4%.

We also apply this simple procedure to other outcome variables. Table 2 reports results for average labor cost and the share of unskilled workers. Results clearly show a sharp negative effect of the ex ante reduction on average labor cost as well as a positive effect on the share of unskilled workers. Both results thus indicate that the employment effect is probably linked to strong substitution effects among workers.

[Table 2 : Effect of the ex ante reduction in labor cost on employment and other firm variables between 1994 and 1997.]

We examine the sensitivity of our results to the choice of control variables. Table 3 present results when only a subset of the previous variables is introduced. We consider five types of variables: past values of firm-level variables (value added, labor productivity, capital labor ratio, ratio of operating income over capital, mark-up, wage share), past values of the changes of these variables (value added, capital labor ratio, total factor productivity, ratio of operating income over capital, mark-up, wage share), financial variables (cost of capital in level and evolution, debt ratio, *exante* variation in the cost of capital associated with fiscal policy change), the skill structure (share in hours of 18 types of workers) and sector-level variables (entry and exit rates, import and export ratios, average labor cost of unskilled workers, growth rate of unskilled workers). For each type of variables we consider two control sets. In the first one we only introduce the considered subset. The results appear in the column labeled as "Just with". In the second one, all variables of the total set except the considered subset are included. The corresponding results appear in the column labeled as "Without" in Table 3. For both sectors there is a substantial difference between the results obtained with subsets that include "*firm-level variables*" and the others. The value of the estimated slope parameter is small when this set is not introduced. Once introduced, the estimated slope parameters are substantially higher and do not depend strongly on the additional introduction of part of the other sets.

[Table 3 : Sensitivity of the slope parameter to control variables.]

Among the variables in level, labor productivity plays a key role. When introducing only labor productivity, the elasticity of the treatment variable in manufacturing is 2.24 and 1.29 in non manufacturing. These values are close to the results obtained with the whole set of control variables. Introducing only total factor productivity produces similar results (defined as the residual of the logarithmic regression of labor productivity on the capital-labor ratio). In the analysis sketched out in Section 2, the treatment depends on the level of productivity and the way in which wages are set. The changes in employment depend on the treatment and the changes in productivity. One interpretation of the key role of total factor productivity is that firms with low total factor productivity are likely to be

those for which things are getting worse and therefore those which are more likely to experience a reduction in their employment. Indeed the coefficient on total factor productivity in the previous simple regression of employment growth is positive.

5. Extension of the Rubin causal model to continuous treatment

In this section we extend the Rubin causal framework (Rubin (1974)) to the case of a continuous treatment and define the suitable parameters of interest. We also demonstrate their identification under an extended form of independence conditional on observables (Rubin (1977). We extend the propensity score property of Rosenbaum and Rubin (1983). Finally, we propose a semiparametric estimation method based on series estimators, using Andrews (1991) results.

5.1 Notations and definition of individual effects

Rubin's initial framework considers the effect of a binary treatment. The causal effect of the program on one variable of interest y_i is defined as the difference of two potential outcomes $y(1)_i$ and $y(0)_i$ corresponding to what would be the situation of individual if they receive or not the treatment As potential outcomes are not simultaneously observable, it is not possible to identify directly the individual causal effect.

In the statistical model considered here, firms i, $i = 1, \ldots, N$, can receive any treatment t falling in the interval $\left[\underline{t}, \overline{t}\right]$. The model introduces for each firm as many latent output variables $y_i(t)$ as there are possible treatments t. As in the one treatment case only one of these variables is observed, namely the one associated with the treatment that the firm has in fact received, i.e. $y_i(t_i)$. This can also be written as:

$$y_i = \int y_i(t)\delta_{t_i}(t)dt$$

Where δ_{t_i} is the Dirac function in t_i.

We consider various individual effects of the measure defined from the latent variables $y_i(t)$. By analogy with the Rubin causal model, one can compare for an individual the situations in which she benefits from treatments t_1 and t_0,

$$c_i(t_0, t_1) = y_i(t_1) - y_i(t_0).$$

A closely related parameter is the effect of a marginal increase in treatment t_0

$$d_i(t_0) = \partial y_i(t_0)/\partial t$$

We also consider the difference between the observed outcome and one potential outcome associated to a given common value of treatment: $y_i - y_i(t_0)$. Among these effects the comparison with the potential outcome associated to the treatment value zero is of specific interest:

$$e_i = y_i - y_i(0)$$

Last, we consider the effect of a uniform marginal increase in treatment:

$$f_i = \partial y_i(t_i)/\partial t.$$

As in the one treatment case, these effects are unobservable. It is necessary to make assumptions to identify and estimate the expectations of these various parameters, under specific assumptions.

We therefore defined the four following parameters of interest:

(3) $\quad E_1 = E(y_i(t_0))$ $\qquad$ Uniform Treatment (UT)

(4) $\quad E_2 = E(\partial y_i(t_0)/\partial t)$ $\qquad$ Marginal-Increase of a Uniform-Treatment (MIUT)

(5) $\quad E_3 = E((y_i - y_i(0)))$ $\qquad$ Average Treatment on the Treated (ATT)

(6) $\quad E_4 = E(\partial y_i(t_i)/\partial t)$ $\qquad$ Marginal-Increase of Treatment (MIT)

The first parameter E_1 (UT) represents the average output that would have been observed if all individuals had received a same treatment of intensity t_0. The second parameter E_2 (MIUT) represents the average effect of a marginal increase in the treatment when it has a constant value equal to t_0 throughout the population. The third parameter E_3 (ATT) represents the average effect of the treatment received $y_i(t_i)$ compared with the situation in which all individuals would have received a null treatment. It is called Average Treatment of the Treated because it can be seen as a weighted mean of the effect of a given treatment on these how received it: $ATT = E(E(y_i - y_i(0)|t_i))$. Interpreting this parameter as the effect of the tax subsidies therefore assumes that this latter situation $y_i(0)$ is identical to the one that would have prevailed in the absence of the program y_i^0, which amounts to assume that there are no indirect effects (Heckman, Lalonde and Smith (1999)):

Assumption (A-2) : $y_i(0) = y_i^0$

The last parameter E_4 (MIT) is the average effect of a marginal variation of the treatment t_i received by each individual.

5.2 Identification of the parameters of interest
One important type of identifying assumption in the case of a unique treatment is independence of potential outcomes $y(1)_i$ and $y(0)_i$ with treatment conditional on a set of observables, or under a weaker form the independence of $y(0)_i$ with treatment conditional on a set of observable. This identifying assumption has been analysed by Heckman, Ichimura and Todd (1997, 1999) and Heckman, Ichimura, Smith and Todd (1998). In these papers, they develop a kernel matching estimator and evaluate its performance using the experimental data available from the JTPA program.

We make the assumption that their exists a set of observable variables x_i conditional on which potential outcomes $y_i(t)$ and treatment t_i are independent:
Assumption (A-3) : $\exists x_i \quad$ such that $y_i(t) \perp t_i | x_i, \forall t \in [\underline{t}, \bar{t}]$

Proposition (P-1):
Under assumption (A-3), parameters E_1, E_2, E_3 *and* E_4 *are identifiable.*

Proof: The general idea underlying proposition (P-1) is simple. Consider for example parameter E_1. We have $E_1 = E(y_i(t_0)) = E_x[E(y_i(t_0)|x_i)]$. Given the independence assumption A-3 we also have: $E(y_i(t_0)|x_i) = E(y_i(t_0)|x_i, t_i = t_0) = E(y_i|x_i, t_i = t_0) = g(x_i, t_0)$, where $g(x_i, t_i) = E(y_i|x_i, t_i)$ is identifiable from the data. Thus, the parameter E_1 is simply $E_1 = E[g(x_i, t_0)]$. Similarly the second parameter can be rewritten as $E_2 = E[\partial g(x_i, t_0)/\partial t]$, the third one as $E_3 = E[y_i - g(x_i, 0)]$ and the last one as $E_4 = E[\partial g(x_i, t_i)/\partial t]$ and are all means of functions identifiable from the data.

Whereas in the previous approach, a particular form was given to the function g: $g(x, t) = a\,t + x\,b$, in this new approach no functional form is specified. The counterpart of this generalization is the practical difficulty of estimating function $g(x, t)$. However, we show that it is sufficient to estimate a function of smaller dimension $g(s(x_i), t_i)$ where $s(x_i)$ is the score, properly defined. For this purpose, we generalize the property of Rosenbaum and Rubin (1983) in the case of discrete treatment to the case of continuous treatment.

Proposition (P-2):
Let x be a vector of covariates satisfying assumption A-3 : $y_i(t) \perp t_i | x_i, \forall t \in [\underline{t}, \bar{t}]$
Let $s(x)$ *be a function of these variables such that*
$$l(t_i | x_i) = l(t_i | s(x_i))$$
where $l(t_i | x_i)$ *is the distribution of the treatment conditionally on* x_i,
Then assumption (A-3) holds for $s(x_i)$ *:*

$$y_i(t) \perp t_i | s(x_i), \forall t \in [\underline{t}, \bar{t}]$$

Proof: it is sufficient to show that $l(t_i | y_i(t), s(x_i)) = l(t_i | s(x_i))$ under the hypotheses $y_i(t) \perp t_i | x_i$ and $l(t_i | x_i) = f(t, s(x_i))$. We calculate the two quantities and show that they are equal. On the one hand we have:

$$l(t_i | y_i(t), s(x_i)) = \int l(t_i | y_i(t), x_i) l(x_i | s(x_i)) dx_i = \int l(t_i | x_i) l(x_i | s(x_i)) dx_i$$

$$= \int f(t_i, s(x_i)) l(x_i | s(x_i)) dx_i = f(t_i, s(x_i)) \int l(x_i | s(x_i)) dx_i = f(t_i, s(x_i))$$

since given the assumptions $l(t_i | y_i(t), x_i) = l(t_i | x_i) = f(t_i, s(x_i))$. Furthermore, we have:

$$l(t_i | s(x_i)) = \int l(t_i | x_i) l(x_i | s(x_i)) dx_i = \int f(t_i, s(x_i)) l(x_i | s(x_i)) dx_i$$

$$= f(t_i, s(x_i)) \int l(x_i | s(x_i)) dx_i = f(t_i, s(x_i))$$

since $l(t_i|x_i) = f(t_i, s(x_i))$.

In the case of a single treatment, the score is of dimension 1 and corresponds to the probability of treatment conditionally on the control variables. There is nothing to say that it would be of dimension 1 in the case of continuous treatment. We shall nevertheless make this assumption in the application to the payroll tax subsidies. We shall describe later the procedure we adopted to estimate the score. The function $g(s(x_i), t_i)$ is then bi-variate.

In the case of a unique treatment the conditional probability of being treated plays an important role both because of the Rosenbaum and Rubin property but also because parameters of interest can be estimated through weighting, weights being defined as functions of the conditional probability. This last type of estimation is interesting as it allows to reach the semiparametric efficiency bound (Hirano, Imbens, and Ridder (2001)). In the case of continuous treatment it is also possible to obtain an expression of some of the parameters of interest as a weighted mean of observation, weights being defined as a function of the conditional density. More precisely, we have the following property :

Proposition (P-3)
Let x be a vector of covariates satisfying assumption A-3 : $y_i(t) \perp t_i | x_i, \forall t \in [\underline{t}, \bar{t}]$

Let $f_t(t_0)$ *and* $f_t(t_0|x)$ *be respectively the unconditional and conditional distribution of the treatment t on* t_0, *then*

$$E(y(t_0)) = E\left(y \frac{f_t(t_0)}{f_t(t_0|x)} \bigg| t = t_0 \right)$$

Proof: Consider a given value of treatment t_0 and the Dirac function $\delta(t - t_0)$ defined by $\int g(t)\delta(t - t_0)dt = g(t_0)$ for any function $g(t)$. For any random variable z we have the following result : $E(z\delta(t - t_0)) = E(z|t = t_0)f_t(t_0)$.
Indeed we have

$$E(z\delta(t - t_0)) = \int z\delta(t - t_0)f_{z,t}(z,t)dzdt = \int z f_{z,t}(z,t_0)dz = \int z f_z(z|t_0)f_t(t_0)dz = E(z|t = t_0)f_t(t_0)$$

Now consider $E(y(t_0)\delta(t - t_0)|x)$, given the independence property, we have:

$$E(y(t_0)\delta(t - t_0)|x) = E(y(t_0)|x)E(\delta(t - t_0)|x)$$

and because of the preliminary result $E(\delta(t - t_0)|x) = f_t(t_0|x)$. Thus we have

$$E\left(\frac{y(t_0)}{f_t(t_0|x)}\delta(t - t_0) \bigg| x \right) = E(y(t_0)|x) \quad \text{which can be integrated over x to yield}$$

$$E\left(\frac{y(t_0)}{f_t(t_0|x)}\delta(t - t_0) \right) = E(y(t_0))$$

Applying now the preliminary result we get $E\left(\frac{y(t_0)}{f_t(t_0|x)}\delta(t - t_0) \right) = E\left(\frac{y(t_0)}{f_t(t_0|x)} \bigg| t_0 \right)f_t(t_0)$ and thus the desired result.

5.3 Estimation

Our estimators are defined as sample means of suitable functional transformation of a non-parametric estimate $\hat{g}(s,t)$ of $g(s,t) = E(y_i | s_i = s, t_i = t)$. We considered the following estimators for the parameters of interest UT, MIUT, TT, MIT defined by equations (3) to (6):

$$(7) \quad \hat{E}_1 = \frac{1}{N} \sum_{i=1}^{N} \hat{g}(s_i, t_0)$$

$$(8) \quad \hat{E}_2 = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial t} \hat{g}(s_i, t_0)$$

$$(9) \quad \hat{E}_3 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{g}(s_i, 0))$$

$$(10) \quad \hat{E}_4 = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial t} \hat{g}(s_i, t_i)$$

There are various ways of estimating the unknown function g non parametrically. Among them we could use the popular kernel estimator. However such a procedure would lead to intense computational burden as the function has to be estimated on all sample points. Instead, we used non-parametric series estimators to estimate the bivariate function $E(y_i | s_i, t_i) = g(s_i, t_i)$. The function g is approximated by a polynomial function of treatment and score whose degree increases with the size of the sample. We therefore consider a polynomial basis (P_k), used to define the set of regressors $P_i = (P_k(s_i) P_l(t_i))_{k+l \leq d_n}$, with k and l the degrees for each of the polynomials and d_n the maximum degree of the polynomials used (we use Chebytchev polynomial which are an orthogonal basis on $[-1,1]$ for the weighting function $1/\sqrt{1-x^2}$). We then consider the coefficients $\hat{\theta}' = (\hat{\theta}_{k,l})_{k+l \leq d_N}$ of the regression of variables y_i on this set of regressors in the sample: $\hat{\theta} = (P'P)^+ (P'Y)$ with $P = (P_i)$ and $Y = (y_i)$, $\forall i = 1, \ldots, N$. The estimators for functions g and $\partial g / \partial t$ are then given by

$$\hat{g}(s,t) = \sum_{k+l \leq d_n} P_k(s) P_l(t) \hat{\theta}_{k,l} \quad \text{and} \quad \frac{\partial}{\partial t} \hat{g}(s,t) = \sum_{k+l \leq d_n} P_k(s) \frac{\partial P_l}{\partial t}(t) \hat{\theta}_{k,l}$$

The characteristic of these estimators is that the degree d_n of the polynomials entering the approximation of functions g and $\partial g / \partial t$ increases with the size of the sample. This determines the number of regressors, which is equal to $(d_n + 1)(d_n + 2)/2$. In our application we will chose a degree between 1 and 6, which leads to the introduction of up to 28 regressors.

We refer to Andrews' work (1991) to study the asymptotic properties of the estimators of the parameters E_1 to E_4 (consistency and asymptotic normality). Parameters E_1, E_2, E_3 and E_4 are of the form $\Gamma(g)$,

for example $E(y(t_0)) = E(g(s,t_0))$. The estimators $\hat{E}_1$, $\hat{E}_2$, $\hat{E}_3$ and $\hat{E}_4$ can be written as $\Gamma_n(\hat{g})$, for example $\sum \hat{g}(s_i, t_0)/n$. The difference $\Gamma_n(\hat{g}) - \Gamma(g)$ can be written as $[\Gamma_n(\hat{g}) - \Gamma_n(g)] + [\Gamma_n(g) - \Gamma(g)]$. The asymptotic properties of the second term are straightforward. The main difficulties concern the first term which involves the estimation of the unknown function g.

Andrews (1991) gives conditions for the type of the functional Γ, function g and the family of functions used in the approximation $\hat{g}$ to describe the asymptotic behavior of the estimators. He examines several types of functional. One of them is $\Gamma_n(g) = \int D^\lambda g(t,s) d\eta(t,s)$, where η is a distribution of probabilities that can depend on the size of sample n and D^λ is any partial derivative operator (cf. case 6 of example 2.7 on page 310 of Andrews (1991)). Indeed, for the first parameter E_1 or the last one E_4 we have $\lambda = (0,0)$ and $d\eta = \delta_{t_0} \otimes d\hat{F}_n(s)$, with $d\hat{F}_n(s)$ the empirical distribution s and δ_{t_0} the delta function in t_0. For the second parameter E_2 we have $\lambda = (1,0)$ and $d\eta = \delta_{t_0} \otimes d\hat{F}_n(s)$, and for the last one E_4 $\lambda = (1,0)$ and $d\eta = d\hat{F}_n(t,s)$. Andrews deals precisely with the case of polynomial approximations of functions with compact support. The results (Theorems 1 and 2 and their application to Example II) make it possible to establish that under specific assumption concerning the regularity of function g our estimators are consistent and asymptotically normal when the degree of the polynomials increases with the size of the sample at a rate n^γ with $\gamma < 1/6$. On the other hand, Andrews' results do not make it possible to identify the convergence rates of the estimators. In the application used in our study we use bootstrap to calculate the standard deviations by a random drawing of 500 samples and apply the same estimation procedure to each drawing.

5.4 Weighted parameters
In order to obtain effects on the growth rate of aggregated employment, we also consider weighted means of individual parameters, mainly :

$$E_3^{\varpi} = E(\varpi_i(y_i - y_i(0))) \qquad \text{Weighted Treatment on the Treated (WTT)}$$

$$E_4^{\varpi} = E(\varpi_i \, \partial y_i(t_i)/\partial t) \qquad \text{Weighted Marginal-Increase of Treatment (WMIT)}$$

where ϖ_i is the normalized weighting variable i.e. $\varpi_i = N_i/E(N_i)$ in which N_i denotes employment. The identification and estimation of these parameters requires additional assumptions. For identification, the weighting variable has to be included in the list of conditioning variables.

Assumption (A-4): $y_i(t) \perp t_i | x_i, \varpi_i$, $\forall t \in [\underline{t}, \bar{t}]$.

Under this assumption it can be shown easily that the two weighted parameters are identified. However the estimation of the final parameters require the estimation of $E(y_i | s_i, t_i, \varpi_i)$ and not only $E(y_i | s_i, t_i)$ as previously. Although it is theoretically possible to consider non-parametric series estimators for these three variables, it is necessary in practice to make assumptions about the form taken

by this function. We chose to introduce simply the weighting variable as an additional regressor, i.e.

$$E\left(y_i \middle| s_i, t_i, \varpi_i\right) = \alpha \varpi_i + f\left(s_i, t_i\right).$$

5.6 The support

Heckman et al. (1998) pointed out the importance of the so called *support condition* in the case of single treatment when independence conditional on observable is assumed. This condition is also likely to be important in our case of a continuous treatment. The estimations of the parameters E_1 are to be compared with each other. They therefore are only of interest if it is possible to estimate quantities of the type $E\left(y_i(t_0) - y_i(t_1)\right)$, which necessitate estimation of both $E\left(y_i \middle| s_i, t_i = t_0\right)$ and $E\left(y_i \middle| s_i, t_i = t_1\right)$. For this, it is necessary that the score s_i belongs to the intersection of the supports of the distribution of the score conditional on treatments t_0 and t_1. Since one wants to make comparisons over an interval of treatment variables ($t \in \left[\underline{t}, \overline{t}\right]$), it is necessary to examine the support of the conditional distribution of the score, knowing treatment $f\left(s_i \middle| t\right)$, and to determine for the interval $\left[\underline{t}, \overline{t}\right]$ the common support of the corresponding distributions $S = \bigcap_{t \in \left[\underline{t}, \overline{t}\right]} \mathrm{Supp}\left(f\left(s_i \middle| t\right)\right)$. The parameters in which we are interested in this case are then the local parameters $E\left(y_i(t) \middle| s_i \in S\right)$ which can be compared two by two. In the case of parameter E_2, the support condition is automatically met as long as the conditional distribution of the score depends continuously on the treatment. However, if one wants to compare the effect of a marginal increase in the treatment for different values of the treatment, we have also to consider the local parameters $E\left(\partial y_i(t_0)/\partial t \middle| s_i \in S\right)$. In the case of parameter E_3 one wants to find for a firm with treatment t_i and score s_i a counterfactual $E\left(y_i(0) \middle| s_i\right) = E\left(y_i \middle| s_i, t_i = 0\right)$. This can be done only for firms with score belonging to $\mathrm{Supp}\left(f\left(s \middle| t = 0\right)\right) \cap \mathrm{Supp}\left(f\left(s \middle| t = t_i\right)\right)$. Thus the common support in this case is $S = \bigcup_{t \in \left[\underline{t}, \overline{t}\right]} \left\{\mathrm{Supp}\left(f\left(s \middle| t = 0\right)\right) \cap \mathrm{Supp}\left(f\left(s \middle| t = t_i\right)\right)\right\}$. The support associated with parameter E_4 is the total support.

6. Results

6.1 The estimation of the score

The implemantation of the semi-parametric estimators requires to compute the score. For this purpose, we assume that the distribution of the variable for the *ex ante* average labor cost reduction conditional on the control variables is a bivariate function of the *ex ante* cost reduction and an indicator of dimension 1 of the control variables, defined as a linear combination of these variables, i.e.

Assumption A-5 :
$$l\left(t \middle| x\right) = f\left(t, x\beta\right)$$

We consider several way of computing the score. The first one amounts to use the fact that for any function h, there exists a function $\widetilde{h}$ such that:

$$E\left(h(t) \middle| x\right) = \widetilde{h}\left(x\beta\right).$$

We consider the transformation $h(t) = \log(t/(0.10 - t))$ that defines the set of real values as support for the *ex ante* cost reduction and we assume that the corresponding $\tilde{h}$ function is the identity. The score used in estimation is then defined as: $s(x) = h^{-1}(x\hat{\beta})$ and therefore has values belonging to $[0,10\%]$.

The second way to estimate the score is to use the fact that under assumption A-5 their exist a function η unknown such that

$$E(t|x) = \eta(x\beta)$$

Estimation of $\hat{\beta}$ can be performed using semiparametric quasi maximum likelihood (SPQML) has proposed in Newey (1994), the unknown function being approximated by a polynomial function. More precisely, the estimation $\hat{\beta}$ of the parameter β is defined as

$$\hat{\beta} = \arg\min_{\beta} \sum_{i} (t_i - \hat{\eta}(x_i\beta))^2$$

$\hat{\eta}(x\beta)$ being defined as $\hat{\eta}(x\beta) = \sum_{k=0}^{K} \hat{\alpha}_k P_k(x\beta)$ with $\hat{\alpha}_k$ the coefficient of the linear projection of t_i on the set of functions $P_k(x_i\beta)$. We chose these last functions to be the Chebytchev polynomial functions of the logit transformation $2/(1 + \exp(-x\beta)) - 1$. As shown in Newey (1994) there is no loss of efficiency in the estimation of β and no correction to be made in the computation of its standard errors associated to the estimation of the unknown function η. In this case the score is defined as $0.10/(1 + \exp(-x\beta))$, and has also values belonging to the range 0-10%.

In the end using the first procedure, we selected some 40 variables out of the 50 or so initially introduced. The eliminated variables are a few sectoral competition indicators as well as some financial variables.

The estimated coefficients where used as starting values for the semiparametric estimation procedure with a polynomial approximation of degree 1, imposing the value of the intercept and the coefficient of one given variable (one sectoral dummy variable that was found strongly significant) to be the same as in the parametric procedure. We used estimated values as starting values for degree 2 and so on. Starting degree 3, convergence was reached in both sectors after only a few iterations with no sensible changes neither in the objective value nor in the estimated values of the parameters. We thus adopt degree 3 for the polynomial approximation in both sectors.[3] Results are qualitatively the same for manufacturing and non manufacturing and for both estimation procedure. In table D1 of appendix D we present results of the SPQML estimation.

6.2 Results of the semi-parametric estimation

The degree of the polynomial approximation used in the computation is chosen using the cross validation criteria:

---

[3] Cross validation could have been another way to determine the adapted degree of approximation and we will use it in the implementation of our estimation of function $g(s, t)$. However, in the case of the score estimation this would have led to a too intensive computational burden.

$$cv(d) = \frac{\sum_{j} \left(y_j - P_j(d) \hat{b}_{-j}(d) \right)^2}{N}$$

where $\hat{b}_{-j}(d)$ is the vector of the estimated coefficient of a degree d polynomial approximation of the unknown regression function using all observation but the jth and $P_j(d)$ the values for observation j of the vector of the polynomial functions involved in the approximation of degree d.

[Table 4 : Cross validation criteria in manufacturing and non manufacturing sectors]

Table 4 shows that the cross validation criteria is minimized in both sectors for degree 3 which we thus retain in the computation of our parameters of interest although we will examine the sensitivity of the results to this choice.

We first show the estimations of parameters E_1 (UT) and E_2 (MIUT), and examine various features such as the heterogeneity of the effect of the treatment, the accuracy of the estimations. We then go on to examine the effects of the payroll tax subsidies on the basis of parameters E_3 (TT) and E_4 (MIT). We also examine weighted forms of these last two parameters. We then discuss various robustness issues such as the support condition, the choice of the degree of the polynomial approximation, the method used to compute the score and the sensitivity to the choice of control variables. Finally we provide estimations for a broader set of outcome variables enabling us to shed light into the underlying mechanisms at work.

6.2.1 Heterogeneity of the effect

Figures in appendix E1 and E2 respectively shows the estimation results of parameters $E(y(t) - y(0))$ and $E(\partial y(t)/\partial t)$ for both manufacturing and non manufacturing sectors. These estimates are computed as $\sum_{i}[\hat{g}(s_i, t) - \hat{g}(s_i, 0)]/N$ and $\sum_{i} \partial \hat{g}(s_i, t)/\partial t /N$ where $\hat{g}$ is the estimation of the polynomial approximation of degree 3 of the unknown function g. Parameters are computed for 50 treatment values regularly spaced over the interval (0,0.10). The 5% confidence interval appearing in the figure (doted lines) are obtained with 500 bootstrap replications.

The effect of a marginal increase in treatment is quite constant in manufacturing with a value around 3%. It exhibits a slight increase for the largest values of the treatment but becomes very imprecise at this level. On the opposite, the marginal effect is not constant in the non manufacturing sector. Initially close to zero for small values of the treatment, it then rises sharply to reach a maximum value of 5% for an *ex ante* reduction of 2.5% and then decreases. Negative values for the parameter are obtained for the largest values of the treatment. As a result, parameter $E(y(t) - y(0))$ exhibits a clear upward sloping pattern.

Figure E3 plots simultaneously estimated values of parameters $E(y|t)$ as well as $E(y(0)|t)$. It clearly shows the strong bias that would arise when ignoring the simultaneity of treatment received and potential outcomes and also that firms that benefited from the most intensive treatment were also firms that would have experienced the most adverse evolution of their employment.

In both sectors estimations are fairly precise for small values of the treatment, where most observations are located. For large values of the treatment, estimation becomes very imprecise due to the smaller number of observations in this region.

6.2.2   Employment effect

In Table 5 we present the parameter estimates for the growth rate of employment obtained from a polynomial approximation of degree 3. We present results for both manufacturing and non manufacturing sectors. Four parameters are presented. These are the effect of the treatment on the treated E_3 (TT) and the effect of a marginal increase of treatment E_4 (MIT), both weighted and non weighted. Standard errors are obtained by bootstrap with 500 replications. All the parameters point to an important effect of the policy implemented. The effect of a marginal increase of treatment is around 3 depending on the sector. The aggregated effects on the growth rate of employment are respectively 2.24% in manufacturing and 3.15% in non manufacturing.

These new estimations can be used to provide an order of magnitude of the number of job created or safeguarded. Taking a basis of 5.7 millions employees in manufacturing and of 9.1 millions in non manufacturing we find that 415.000 jobs have been created of safeguarded with a 5% confidence interval of (270.00,560.000).

[Table 5 : Semi parametric estimation of Treatment Effect]

Figures 3a and 3b present results obtained for the weighted average effect of treatment on the treated WATT with different degrees of the polynomial approximation of the unknown function $g(s,t)$. The figure shows the point estimate of the parameter as well as its 5% confidence interval, computed using bootstrap with 500 replications. In the manufacturing sector the point estimate does not change much after degree three and increasing the degree only widens of the confidence interval. The picture is different in non manufacturing sector and less clear. Increasing the degree does not significantly affects the confidence interval and leads to fluctuation in the point estimate.

[Figure 3a : Sensitivity to the degree of the polynomial approximation - Manufacturing – Weighted effect of treatment on the treated]

[Figure 3b : Sensitivity to the degree of the polynomial approximation – Non Manufacturing – Weighted effect of treatment on the treated]

All the estimated values of the parameters exhibit impressive changes compared to results obtained using the previous direct procedure. For example, the weighted average growth rate of employment is increased form 1.3% to 2.3% in manufacturing. Similarly in non manufacturing the same parameter is increased from 2.4 to 3.2. Although there is also a substantial increase in standard errors, these changes in results point to the drawbacks associated to the assumption of linearity and homogeneity of effects.

Table 6 presents further results corresponding to various restrictions on the specification of the firm response. We examine the sensitivity of the results when imposing linearity of the response and both linearity and homogeneity. For each parameter TT and MIT, the first line recalls the results for the general unconstrained specification. The second line presents results when imposing linearity of the effect of the ex ante reduction in labor cost. The third line present results obtained when imposing both linearity and homogeneity. All unknown functions of the score entering these specifications are polynomial functions of degree three. Imposing linearity does not affect results in manufacturing, but imposing homogeneity does matter. This is not a surprise given the results of the estimation of parameters IUT presented in the previous section. The weighted average growth rate decrease from 2.3% to 1.9% when both linearity and homogeneity are imposed, but remains however higher than with the non constrained earlier specification in which controls enter linearly (1.3%). In non manufacturing on the opposite, both linearity and homogeneity matters strongly. When imposing linearity the estimation of the weighted effect of the treatment is strongly biased downward.

[Table 6 : Semi parametric estimation of Treatment Effect - Alternative specification]

6.2.3 Robustness Check
We examine the robustness of our results along different dimensions. We first examine the sensitivity of the results to the choice of the control variables. We then examine the effect of the restriction to the common support. Finally we examine the sensitivity to the way the score is computed.

To examine the sensitivity of the result to the choice of control the variables we proceed in the same way as in the first estimation method. Results are presented in Table 7. The set of control variables is divided into 5 subsets. For each of these subsets we perform the global analysis using only the considered subset as control variables (appearing as the columns labeled "Just with") and using all the variables except the considered subset (appearing as the columns labeled "Without"). In each case we presented the four main parameters of interest: ATT, WATT, MIT and WMIT. Again, results clearly show that the key subset is the firm variables in level. It also appears that in both sectors using only this subset lead to higher values of the parameters of interest.

[Table 7 : Sensitivity of the result to the choice of control variables]

We also examine the sensitivity of the results when restricting observations to the support adapted for the estimation of parameter E_3. The common support is determined using 20 treatment classes of identical length equal to 0.5%. In each of these classes, we calculate different percentiles. Figures Fa and Fb of appendix F display the results. These figures show the strong concentration of the score around low values for low values of treatment and the gradual widening-out as treatment increases. As a result, for a firm with a low value of the score, it is always possible to find similar firms with low treatment. On the opposite, when the score becomes higher, no similar firm with a low treatment may be found. Thus the common support is defined as the interval of zero score and a maximum value chosen on the basis of the distribution of the score for the low treatment class. We consider this maximum value to be 3% in manufacturing and 4% in non manufacturing. Results when restricting the computation of the parameters of interest to the common support are presented in Table 8. Regressions are performed on the whole sample but parameters of interest are computed using only observations on

the common support. The polynomial approximation is also of degree 3. We do not detect any substantial changes in the parameters of interest.

We also present in Table 8 the results obtained when using the crude score following the procedure described in section 7. A polynomial approximation of degree 3 is also used to estimate the unknown function g. No significant changes are observed in the manufacturing sector. However in the non manufacturing sector, changes are more substantial. The parameters obtained with the crude score are substantially higher, with an order of magnitude of one, two or even more than one standard error.

[Table 8 : Semi parametric estimation of Treatment Effect – Robustness to the support and score estimation]

6..2.4   Substitution and profitability effects

Our previous results point to strong employment effects of payroll tax subsidies for low wage workers. However, the method we use does not help in understanding the underlying effects of such a policy. The substantial evolutions in employment can be related to two broad types of mechanisms. The first one corresponds to substitutions between factors of production i.e. an increase of the unskilled-labor content of production. The second corresponds to a profitability effect, i.e. an increase in all factors of production due to an increase in demand if the reduction in production costs is passed through to prices. In this section we implement our methodology on a broader set of explanatory variables in order to obtain some insights about these underlying effects. Tables 9 and 10 present the results obtained for the semiparametric estimation of parameter E_3 (ATT and WATT) and E_4 (MIT and WMIT) for a broader set of variables. We examine the effect of payroll tax subsidies on the average labor cost, the share of unskilled workers, the capital labor ratio, the productivity of capital, the labor productivity, the stock of capital and value added.

There are drawbacks in some of these measures. First, the productivity of each factor as well as the value added are measured with accounting data. This data are is specified in nominal terms and we only use the aggregate price at the two digit level as deflator. These measures thus do not take into account the possible effect of payroll tax subsidies on individual prices. Second, unskilled labor is defined according to occupations and do not perfectly covers workers paid below 1.33 minimum wage as shown in Table 11.

Our results suggest that strong substitution effects may have been at work. This can be primarily seen in the negative effect on average labor costs, which may be related to composition effects within the workforce in favor of least paid workers. This can also be seen on the negative effect on the share of unskilled workers. Our results indicate that substitutions also operate between labor and capital, with the capital labor ratio falling under the impact of tax subsidies in each sector. We also observe a decrease in the productivity of capital and labor in both manufacturing and non manufacturing sectors. This is related in both cases to an increase in the volume of employment and capital but also to a stagnation and even a decrease in value added. To interpret these results assume a simple demand function with constant price elasticity $d \log q = -\varepsilon d \log p$. The firm's revenue is then simply $d \log pq = (1 - \varepsilon) d \log p$. Assuming a constant elasticity of substitution between labor and capital and constant return to scale we have $d \log q - d \log k = \sigma(d \log c - d \log p) = -\sigma d \log p$, assuming no effect on the cost of capital. Thus the changes in the variables of interest we can measure are $d \log k = (\sigma - \varepsilon) d \log p$, $d \log pq - d \log k = (1 - \sigma) d \log p$ and $d \log pq = (1 - \varepsilon) d \log p$. As a result,

we expect changes in these variables only if payroll tax subsidies are passed into prices. Our results on the capital stock and the productivity of capital are consistent with the previous computations if prices falls and the elasticity of substitution is small, i.e. below 1 and the price elasticity of demand. Notice that the comparison of changes in the capital labor ratio and changes in the average labor cost are consistent with an elasticity of substitutions around 0.5. To fix ideas, assuming such a value, the effect of payroll tax reduction on prices should be half the effect on the productivity of capital i.e. an average elasticity of 0.7, which is small given the strong effect on wages. However we also expect an increase in the value added, even not deflated as long as the price elasticity of demand is above 1. We observe a non significant but positive effect in manufacturing of payroll tax subsidies on value added which suggests that the value of the price elasticity of demand is above and close to 1. However, we observe negative effects in non manufacturing that may be significant (average slope parameter).

These results therefore show that the main underlying effects are probably substitution effects and mostly substitution among workers. Profitability effects may also have been at work but the price effect seems to be small. Changes in value added are of a small magnitude when they have the right sign. The only change that may be related to a profitability effect lies with capital. This may however receive an alternative explanation

[Table 9 : Semi parametric evaluation of a marginal increase of the exante reduction in labor cost]

[Table 10 : Semi parametric evaluation of growth rates due to payroll tax reduction for low wage workers.]

[Table 11 : Unskilled workers and minimum wage]

Indeed, all our results point to a strong and quick effect of the policy. We may ask whether it is reasonable to expect such a strong effect to occur so rapidly It is generally considered that substitution effects and profitability effects take time to materialize. Our results indicate that labor was substituted to capital and one may expect this substitution to occur only after a long period of time. Similarly we also observe evolution among the work force in favor of less skilled workers and this also requires time.

However when the counterfactual is an ongoing process of capital to labor substitution, of skilled to unskilled labor substitution and of a reduction in the capacity of production of unskilled intensive firms, then the effects of factor price changes can be more observed more quickly. Figures E3, show for both sectors the expectation of the counterfactual of employment evolution conditional to treatment: $E(y_i(0)|t)$ and the expectation of observed output variable conditional to treatment: $E(y_i|t)$. We clearly observe that the counterfactual is negative and sharply decreasing with treatment. On the opposite the expectation of employment conditional to treatment is not related to treatment. This suggests that the main effect of the policy was to stop an ongoing process of unskilled job destruction and not one of job creation, although there is probably a great deal of heterogeneity. This last conclusion is consistent with results obtained in Kramarz and Philippon (2001). Using individual employee data, they show, that the main effect of the policy is a reduction of employment to unemployment transition compared with employment to employment transitions. They do not observe an increase in the transition of unemployment to employment compared with employment to employment transitions.

This interpretation of our results is also consistent with macro-evolution at the time the policy took place. The payroll tax subsidies for low wage workers took place in a macroeconomic context that had worsened considerably (figure 4), following a long period of stagnation in activity since the beginning of the 1990s that lasted until 1994. During this period, employment had fallen sharply. The year 1994 signaled a break in this tendency with a substantial upturn in employment. Table 12 clearly shows that this upturn was associated with a decrease in the destruction rate of employment from 24.2% over the upswing period 1988-1991 and 24.5 over the recession period 1991-1994 to 20.5% over the period 1994-1997. Similarly, the decline in the productivity of capital stopped in 1994, as the job-content of growth began to rise (figure 2). The evolution in the productivity of capital clearly shows that the long run downward trend, almost uninterrupted since the beginning of the 1980s, was suddenly halted in 1993 and was even reversed starting in 1997. Another feature was a halt in the decline in the proportion of unskilled workers in total employment at the beginning of the 1990s (figure 5).

If this interpretation is right, then it is also likely the case that there was no upward pressure in the demand for un skilled workers and therefore that at least in this short run analysis equilibrium wage effects may not be important.

[Tables 12 : Employment creation and destruction rate[4]]

[Figure 4 – Aggregate evolutions between 1978 and 1996, private sector]

[Figure 5 - Share of unskilled workers between 1982 and 1996]

7 Conclusion

The aim of this paper is to evaluate the effect of the payroll tax reduction for low wage workers implemented in France in the middle 90's. The study is based on the calculation of the firm's *ex ante* labor cost reduction attributable to the extension of the original policy in 1995 and 1996. It is carried out using an extensive data source providing the wage distribution inside firms for each firm in 1994. The underlying principle is to compare the results for firms benefiting from different *ex ante* reductions. For this purpose, we develop a statistical model, based on the Rubin causal model, adapted to the continuous treatment case. We define the parameters of interest in such a framework and propose a semi-parametric estimation procedure using series estimators.

The identifying assumption we make is that the ex ante labor cost reduction (the treatment) and potential outcomes (the increase in employment when receiving a given treatment) are independent conditional on observables. We show that the exante reduction of labor cost has a clear interpretation in an economic model that we use to address the identification issue. We show that we need to control for the heterogeneity of the work force, heterogeneity of technological parameters, product market competition and unobserved demand and productivity shocks. We introduce a broad set of control variables including information on past performance in level and evolution, detailed information about the skill structure, various competition variables at the sector level and information about the financial structure and the cost of capital. The introduction of these control variables plays a key role and we show that among this broad set of information the past level of productivity is of primary importance.

---

[4] Defined following Davis and Haltiwanger 1992.

We find that payroll tax subsidies introduced between 1994 and 1997 had a strong effect on employment. The average growth rate attributable to the policy is 2.24 in manufacturing and 3.15 in non manufacturing and may have created or safeguarded 420,000 jobs in the economy over the period 1994-1997. We find that the effect on employment mainly reflects substitution between employee categories as well as between labor and capital. Profitability effects may also have occurred but are probably of less importance. These strong effects have occurred quickly because the counterfactual evolution of employment in firms that benefited most from the policy was a strong reduction in their employment.

Our results show that the impact of a marginal increase in tax subsidies is very heterogeneous within the population and may differ from one treatment value to another. They show the importance of tackling the question of evaluating subsidies in taxes in a suitable causal framework. Introducing control variables into a direct regression of output variables on the *ex ante* cost reduction would lead to an erroneous evaluation of its effect.

Our evaluation is not based on the specification and estimation of structural models. The effects we measure are a combination of different structural parameters (elasticities of substitution, demand elasticities), that cannot be separately identified. As a result, our evaluation does not require the estimation of elasticities of substitution between various categories of worker, nor the elasticity of the demand for labor to its cost. However, its main disadvantage is that it is only valid for the policy carried out over the 1994-1997 period. The evaluations cannot be used to study alternative policies such as, the extension of the tax cuts to a broader population or, on the contrary, the intensification of payroll tax subsidies for the population already targeted, or, changes in the minimum wage. Nevertheless, this study shows that attempts to change the distribution of earnings in the economy, especially for the lower paid, have substantial effects on employment.

Bibliographie

− Abowd J. M., F. Kramarz et D. N. Margolis, 1999, High Wage Workers and High Wage Firms, Econometrica, 67 (2), 251-333.

− Andrews D., 1991, Asymptotic normality of series estimators for non parametric and semiparametric regression models, Econometrica 59 (2), 307-345.

− Audric S., P. Givord et C. Prost, 2000, Estimation de l'impact sur l'emploi non qualifié des mesures de baisse de charges, Revue économique, vol. 51, no. 3.

− Brodaty T., B. Crépon and D. Fougère, 2001, Using matching estimators to evaluate alternative youth employment programs: evidence from France 1986-1988, in " The Evaluation of Active Labour Market Policies in Europe " M. Lechner eds Springer Verlag.

− Card, David, and Alan B. Krueger. 1995. Myth and Measurement: The New Economics of the Minimum Wage. Princeton, NJ: Princeton University Press.

− Hahn (1998) On the Role of Propensity score in Efficient Semiparametric Estimation of Average Treatment Effects Econometrica vol. 66 n°2 pp. 315-331

− Heckman J. 1979, Characterizing selection bias using experimental data, Econometrica 66(5) : 1017-1098.

− Heckman J., H. Ichimura, J. Smith and P. Todd, 1999, Characterizing selection bias using experimental data, Econometrica 66(5) : 1017-1098.

− Heckman, J. H. Ichimura, and P. Todd, 1997, Matching as an econometric evaluation estimator : evidence from evaluating a job training program, Review of Economic Studies 64(4) : 605 -654.

− Heckman J., H. Ichimura and P. Todd, 1998, Matching as an econometric evaluation estimator, Review of Economic Studies 65(2) : 261-294.

− Heckman, J., and Vytlacil, E., 2000, Local instrumental variables, NBER Technical Working Paper 252

− Hirano, K., G. Imbens and G. Ridder (2000) Efficient Estimation of Average Treatment Effect Using the Propensity Score NBER Technical Working Paper T0251

− Ichimura H., 1992, Semiparametric least squares estimation of single index models, Journal of Econometrics.

− Imbens, G. W., 2000, The Role of the Propensity Score in Estimating Dose-Response Functions", Biometrika, 87(3), 706-710.

− Imbens, G., and Angrist, J., 1994, Identification and estimation of local average treatment effects, Econometrica, 62, 467-476.

− Kramarz F. and T. Philippon, 2001, The impact of differential payroll tax subsidies on minimum wage employment, Journal of Public Economics, 82, 115-146.

− Laroque G. and B. Salanié (2002) "Labor Market Institutions and Employment in France" *Journal of Applied Econometrics* (2002), 17, 25-48.

− Lalonde R., 1986, Evaluating the econometric evaluations of training programs with experimental data, American Economic Review 76(4) : 604-620.

− Lechner, 2001, Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independance Assumption, in " The Evaluation of Active Labour Market Policies in Europe " M. Lechner eds Springer Verlag.

− Meyer, R. H., and D. A. Wise. 1983 The Effect of the Minimum Wage on the Employment Earnings of Youth *Journal of Labor Economics* 1, no. 1: 66-100

− Meyer, R. H., and D. A. Wise. 1983 Discontinuous Distributions and Missing Persons: The Minimum Wage and Unemployed Youth, Econometrica, Vol. 51, N°6, pp. 1677-1698

− Newey W., 1994, The Asymptotic Variance of Semi-parametric Estimator" Econometrica, vol 62 n°6 pp. 1349-1382

– Rosenbaum P. and D. Rubin, 1983, the central role of the propensity score in observational studies for causal effects, Biometrika 70(1) : 41-55.

– Rubin D., 1974, Estimating causal effects of treatments in randomized and non randomized Studies, Journal of Educational Psychology 66 : 688-701.

– Rubin D., 1977, Assignement to Treatment Group on the Basis of a Covariate, Journal of Educational Statistics Spring vol. 2 (1).

Figure 1 – Rules for the reduction in payroll taxes form 1993 to 1998
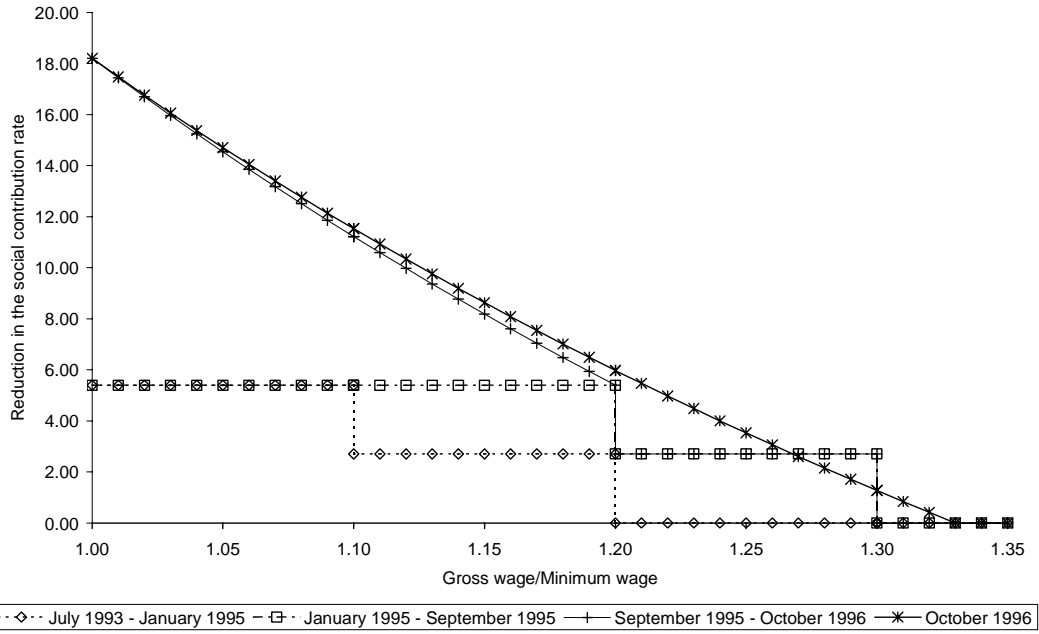
Figure 2 - Distribution of the ex ante reduction of labor cost in manufacturing and non manufacturing sectors.
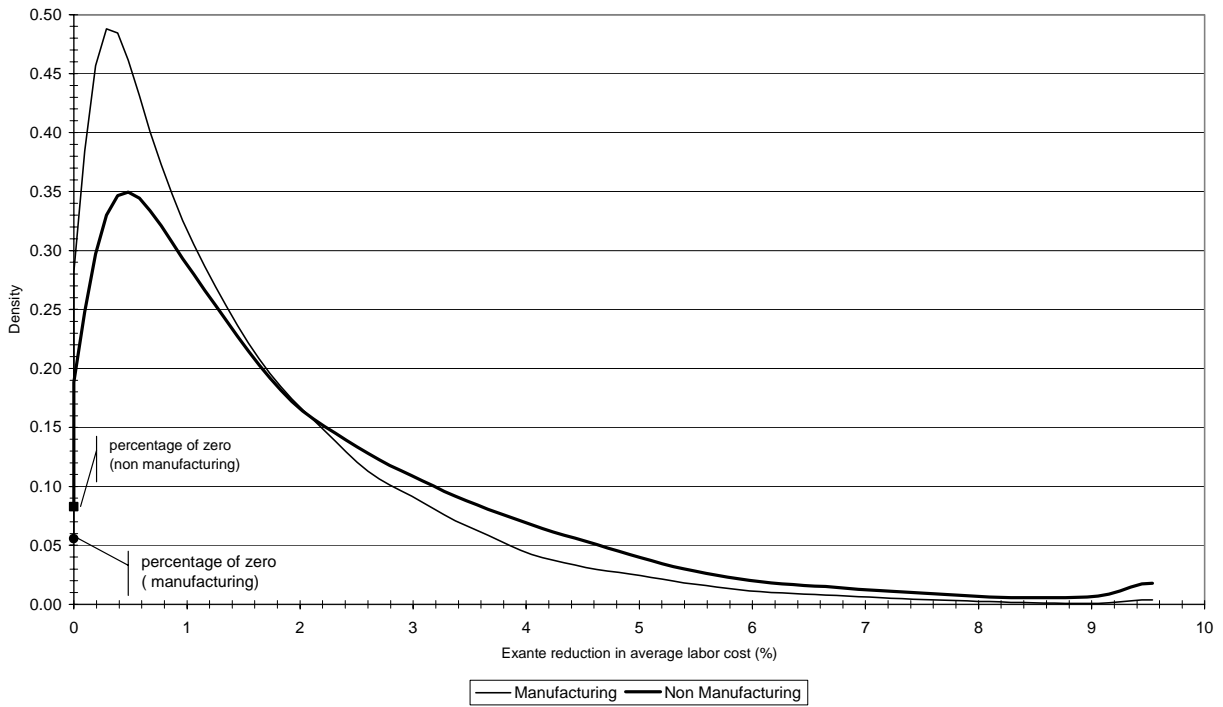
Note : Kernel estimates of the density for positive values of the ex ante reduction.

Figure 3a : Sensitivity to the degree of the polynomial approximation - Manufacturing – Weighted effect of average treatment effect on the treated
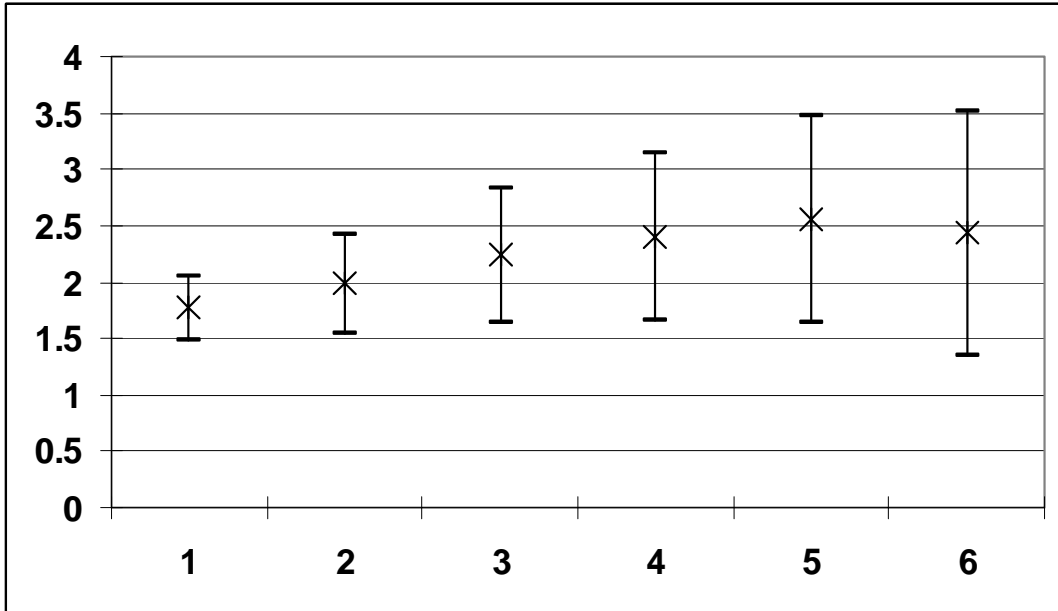
Figure 3b : Sensitivity to the degree of the polynomial approximation – Non Manufacturing – Weighted average treatment effect on the treated

Figure 4 – Aggregate evolutions between 1978 and 1996, private sector

- Labor productivity (right scale)
- Productivity of capital (left scale)

- Employment (left scale)
- Value added (right scale)
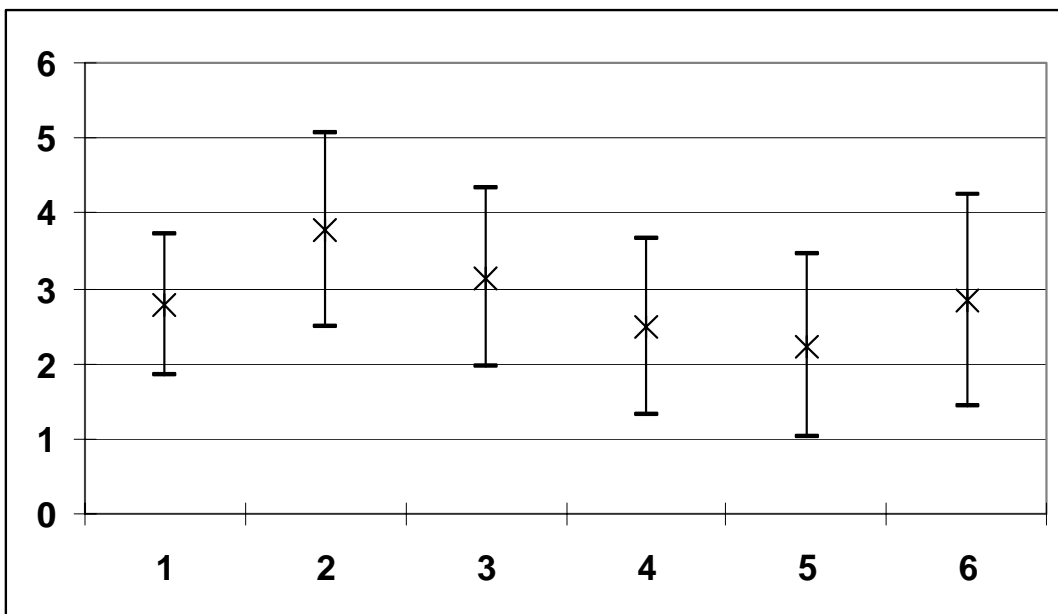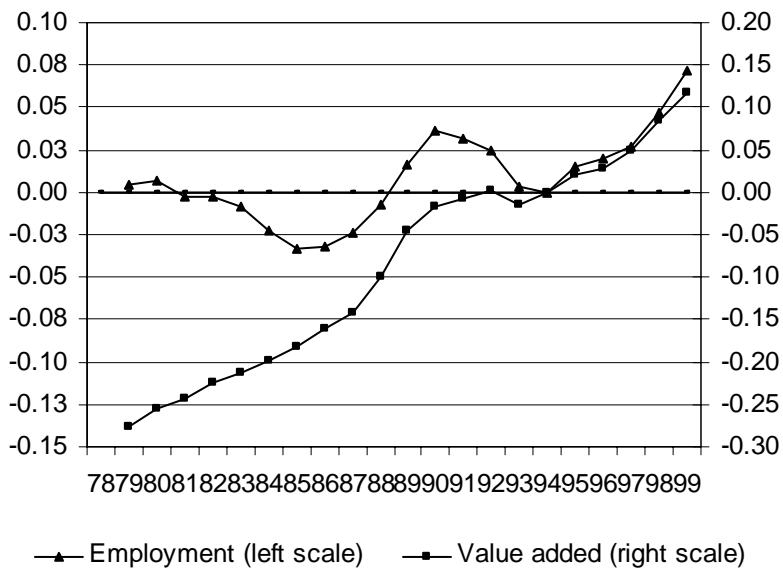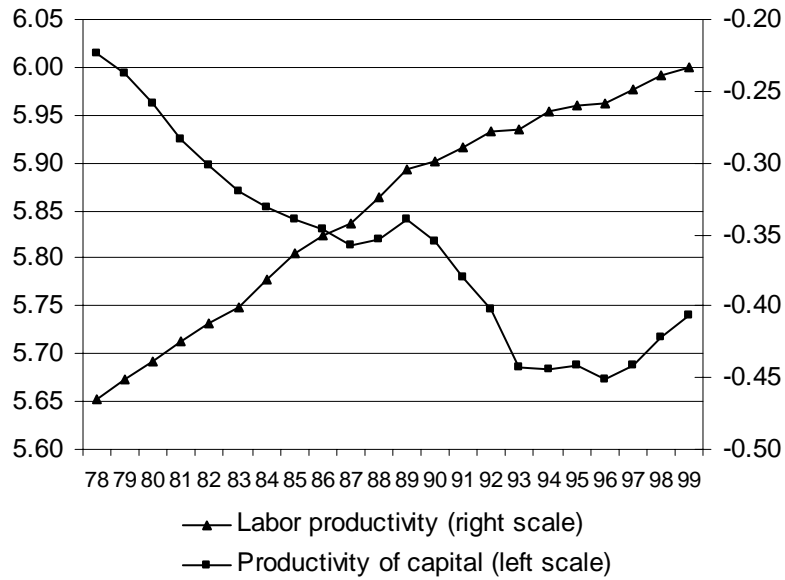
source : National Account (logarithm normalized to zero in 1994)

Figure 5 - Share of unskilled workers between 1982 and 1996
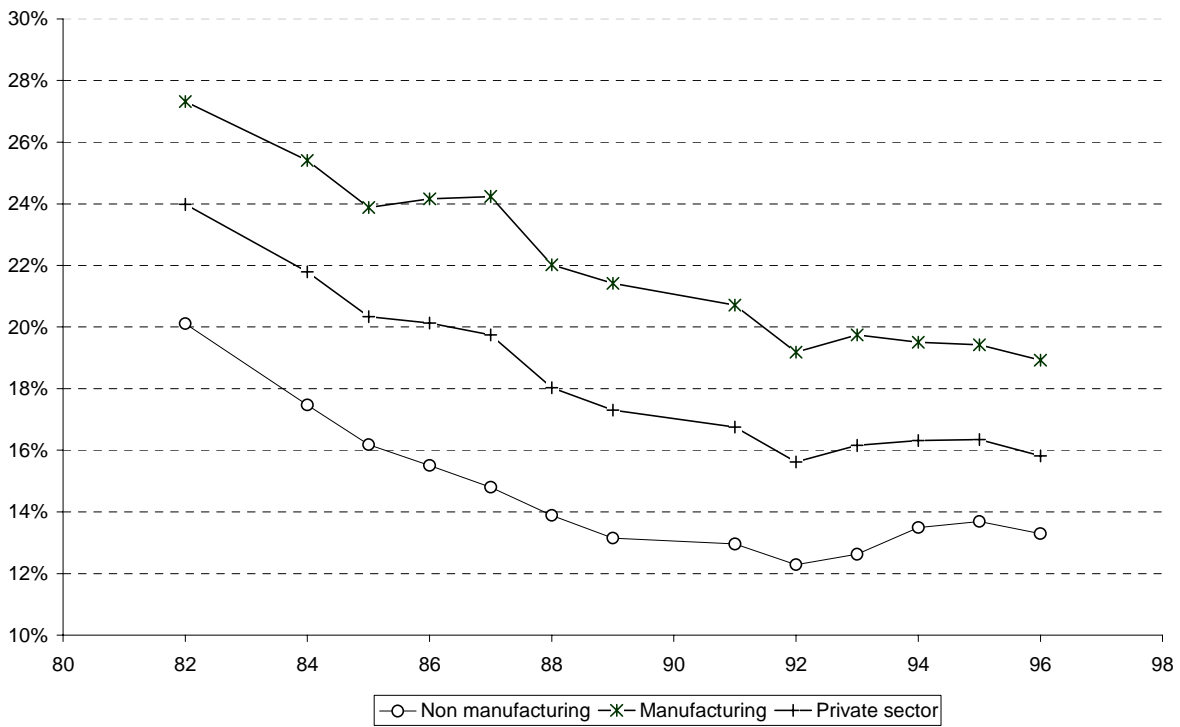
Source : Audric, Givord et Prost (2000) from the DADS : full time employees of the non agricultural private sector. Not available for 1983 and 1990.

Table 1 - Ex ante reduction of labor cost

	0%	0-1%	1-6%	6-9.5%	9.5%
Private sector					
Percentage of firms	7.2	39.3	50.2	2.6	0.6
Percentage of employees	1.0	65.6	32.7	0.6	0.0
Manufacturing					
Percentage of firms	5.6	45.9	46.9	1.5	0.1
Percentage of employees	0.8	73.3	25.6	0.3	0.0
Non Manufacturing					
Percentage of firms	8.3	35.1	52.4	3.3	0.9
Percentage of employees	1.3	56.4	41.3	0.9	0.1

Note : Obtained from the study sample involving 87,720 firms, of which 34,371 (39%) are in manufacturing sectors and 53,349 (61%) in non manufacturing. Employment in these firms is 3,772,941, of which 2,053,777 (54%) are in manufacturing sectors and 1,719,164 (46%) in non manufacturing.

Table 2 : Effect of the ex ante reduction in labor cost on some firm variables between 1994 and 1997.

Variables	Elasticities		Growth rate	
	Manufacturing	Non Manufacturing	Manufacturing	Non Manufacturing
Employment[a]	1.60	1.79	1.28	2.34
	(0.14)	(0.10)	(0.12)	(0.19)
Average labor cost[a]	-2.30	-2.25	-1.84	-2.96
	(0.10)	(0.09)	(0.09)	(0.20)
Share of unskilled workers	0.38	0.49	0.30	0.65
	(0.09)	(0.07)	(0.07)	(0.10)

Note : These results are obtained by the OLS regression of the variable of interest on the ex ante reduction in labor cost and a set of control variables in 1994 and for some of them in evolution over the past period. They are performed on 32,459 observations in manufacturing and 48,930 in non manufacturing. Firms with a zero ex ante reduction in labor costs were discarded. The [a] superscript means that the variable is expressed in logarithm

Table 3 : Sensitivity to control variables.

Variables	Manufacturing		Non Manufacturing	
	Just with	*Without*	*With*	*Without*
All	1.60	0.58	1.78	-0.00
	(0.14)	(0.11)	(0.10)	(0.08)
General in level	1.84	0.51	1.73	0.28
	(0.14)	(0.12)	(0.09)	(0.09)
General in difference	0.47	1.83	0.01	1.99
	(0.11)	(0.15)	(0.08)	(0.11)
Financial	0.39	1.60	0.06	1.77
	(0.12)	(0.14)	(0.08)	(0.11)
Share	0.51	1.69	0.17	1.79
	(0.13)	(0.14)	(0.09)	(0.10)
Sectoral	0.69	1.41	0.16	1.64
	(0.12)	(0.12)	(0.08)	(0.10)

Note : These results are obtained by the OLS regression of the variable of interest on the ex ante reduction in labor cost and a set of control variables in 1994 and for some of them in evolution over the past period. They are performed on 32,459 observations in manufacturing and 48,930 in non manufacturing. Firms with a zero ex ante reduction in labor costs were discarded. For each sector the column "Just with" correspond to the case where only the considered variables have been introduced as control variables. The second column "Without" correspond to the case where all variables have been introduced but the set of variables considered. Dummy variables at the one digit level have always been introduced.

Table 4 : Cross validation criteria in manufacturing and non manufacturing sectors

Degree	Number of regressors	Manufacturing	Non manufacturing
6	29	0.0636641	0.0829775
5	22	0.0636503	0.0829555
4	16	0.0636382	0.0829560
3	11	0.0636278	0.0829536
2	7	0.0636305	0.0829883
1	4	0.0637542	0.0830954

Table 5 : Semi parametric estimation of Treatment Effect

Variables	Manufacturing		Non Manufacturing	
Weight	1	Employment	1	Employment
Effect of a Marginal Increase of Treatment (MIT)				
Employment (log)	2.86	3.38	2.54	3.31
	(0.26)	(0.39)	(0.19)	(0.28)
Average Treatment Effect on the Treated (TT)				
Employment (log)	3.59	2.24	2.55	3.15
	(0.53)	(0.30)	(0.52)	(0.60)

*Note : **These figures are the semi parametric estimates of the parameter** $E_3^{\varpi} = E(\varpi_i(y_i - y_i(0)))$ **and** $E_4^{\varpi} = E(\varpi_i \partial y_i(t_i)/\partial t)$**, obtained with and without weighting firms by their employment. They are performed on 32.459 observations in manufacturing and 48.930 in non manufacturing. Firms with a zero ex ante reduction in labor costs were discarded.***

Table 6 : Semi parametric estimation of Treatment Effect - Alternative specification

Variables	Manufacturing		Non Manufacturing	
Weight	1	Employment	1	Employment

Effect of a Marginal Increase of Treatment (MIT)

	Manufacturing		Non Manufacturing	
Reference	*2.86*	*3.38*	*2.54*	*3.31*
	(0.26)	*(0.39)*	*(0.19)*	*(0.28)*
$g(s,t) = g_0(s) + tg_1(s)$	2.92	3.42	2.93	3.60
	(0.20)	(0.31)	(0.15)	(0.26)
$g(s,t) = g_0(s) + \lambda t$	2.32	2.32	2.15	2.15
	(0.16)	(0.16)	(0.12)	(0.12)
$g(x,t) = xb + \lambda t$	*1.60*	*1.60*	*1.79*	*1.79*
	(0.14)	*(0.14)*	*(0.10)*	*(0.10)*

Average Treatment Effect on the Treated (TT)

	Manufacturing		Non Manufacturing	
Reference	*3.59*	*2.24*	*2.55*	*3.15*
	(0.53)	*(0.30)*	*(0.52)*	*(0.60)*
$g(s,t) = g_0(s) + tg_1(s)$	3.65	2.28	4.56	3.95
	(0.23)	(0.18)	(0.24)	(0.67)
$g(s,t) = g_0(s) + \lambda t$	3.49	1.86	4.33	2.83
	(0.23)	(0.14)	(0.24)	(0.48)
$g(x,t) = xb + \lambda t$		*1.28*		*2.34*
		(0.12)		*(0.19)*

*Note : **See Table 5***

Table 7: Sensitivity of the result to the choice of control variables

	TT	WTT	MIT	WMIT	TT	WTT	MIT	WMIT
		Just with				Without		
				Non Manufacturing				
All	2.55	3.15	2.54	3.31	-	-	-	-
	(0.52)	(0.60)	(0.19)	(0.28)				
Level	3.23	4.70	2.76	4.61	0.71	0.59	0.10	0.09
	(0.45)	(0.81)	(0.17)	(0.33)	(0.38)	(0.32)	(0.16)	(0.23)
Difference	-0.02	-0.03	-0.04	-0.25	2.92	3.60	2.80	3.70
	(0.29)	(0.23)	(0.14)	(0.30)	(0.52)	(0.67)	(0.19)	(0.27)
Share	0.37	0.30	-0.15	-0.39	3.10	3.66	2.79	3.67
	(0.36)	(0.29)	(0.15)	(0.27)	(0.49)	(0.66)	(0.18)	(0.26)
Financial	0.49	0.30	0.17	-0.04	2.43	3.09	2.50	3.23
	(0.31)	(0.24)	(0.14)	(0.25)	(0.52)	(0.59)	(0.19)	(0.28)
Sectoral	0.94	0.58	0.52	0.33	2.23	3.03	2.36	3.17
	(0.32)	(0.28)	(0.15)	(0.38)	(0.49)	(0.57)	(0.19)	(0.27)
				Manufacturing				
All	3.59	2.24	2.86	3.38	-	-	-	-
	(0.53)	(0.30)	(0.26)	(0.39)				
Level	4.03	2.57	3.16	3.83	1.38	0.81	0.68	0.81
	(0.48)	(0.27)	(0.24)	(0.36)	(0.38)	(0.25)	(0.21)	(0.33)
Difference	0.96	0.51	0.58	0.63	4.11	2.50	3.12	3.66
	(0.30)	(0.20)	(0.19)	(0.32)	(0.52)	(0.30)	(0.26)	(0.38)
Share	1.27	0.74	0.56	0.61	3.57	2.40	3.16	3.82
	(0.37)	(0.24)	(0.21)	(0.32)	(0.51)	(0.29)	(0.25)	(0.37)
Financial	0.66	0.36	0.37	0.39	3.69	2.28	2.90	3.39
	(0.31)	(0.21)	(0.20)	(0.33)	(0.53)	(0.29)	(0.26)	(0.38)
Sectoral	1.30	0.70	0.69	0.59	3.36	2.07	2.61	3.08
	(0.31)	(0.21)	(0.20)	(0.31)	(0.59)	(0.30)	(0.26)	(0.38)

Note : For each row the results presented for the four main parameters are obtained using only the considered subset as explanatory variables in the score (left panel) and all the variables but the considered subset (right panel). Standard errors are obtained via 500 bootstrap replications.

Table 8 : Semi parametric estimation of Treatment Effect – Robustness to the support and score estimation

Variables	Manufacturing		Non Manufacturing	
Weight	1	Employment	1	Employment
Effect of a Marginal Increase of Treatment (MIT)				
Reference	*2.86*	*3.38*	*2.54*	*3.31*
	(0.26)	*(0.39)*	*(0.19)*	*(0.28)*
Common support	3.03	3.43	2.76	3.37
	(0.29)	(0.40)	(0.20)	(0.29)
Crude score	2.89	3.52	3.19	4.07
	(0.27)	(0.39)	(0.19)	(0.27)
Average Treatment Effect on the Treated (TT)				
Reference	*3.59*	*2.24*	*2.55*	*3.15*
	(0.53)	*(0.30)*	*(0.52)*	*(0.60)*
Common support	3.22	2.16	3.05	3.27
	(0.42)	(0.28)	(0.41)	(0.63)
Crude score	3.83	2.34	3.67	3.78
	(0.54)	(0.31)	(0.53)	(0.68)

Note : See Table 5. Rregressions are performed on the whole sample but parameters of interest are computed using only observation on the common support (0,3%) for manufacturing, (0,4%) for non manufacturing.

Table 9 : Semi parametric evaluation of a marginal increase of the exante reduction in labor cost

	Manufacturing		Non Manufacturing	
	TT	WTT	TT	WTT
Employment	2.86	3.38	2.54	3.31
	(0.26)	(0.39)	(0.19)	(0.28)
Average labor cost	-2.95	-3.02	-3.34	-4.27
	(0.21)	(0.31)	(0.15)	(0.23)
Share of unskilled workers	0.66	0.61	0.52	0.45
	(0.15)	(0.23)	(0.10)	(0.15)
Capital	1.22	1.65	0.92	1.08
	(0.29)	(0.43)	(0.21)	(0.32)
Capital labor ratio	-1.64	-1.72	-1.62	-2.23
	(0.33)	(0.51)	(0.24)	(0.39)
Productivity of capital	-1.17	-1.25	-1.36	-1.67
	(0.33)	(0.50)	(0.22)	(0.34)
Labor productivity	-2.81	-2.97	-2.98	-3.90
	(0.26)	(0.38)	(0.18)	(0.27)
Value added	0.04	0.40	-0.44	-0.59
	(0.29)	(0.43)	(0.18)	(0.27)

Note : These figures are the semi parametric estimates of the parameter $E_4^\varpi = E(\varpi_i \, \partial y_i(t_i)/\partial t)$, *obtained with and without weighting firms by their employment. They are performed on 32.459 observations in manufacturing and 48.930 in non manufacturing. Firms with a zero ex ante reduction in labor costs were discarded.*

Table 10 : Semi parametric evaluation of growth rates due to payroll tax reduction for low wage workers.

	Manufacturing		Non Manufacturing	
	TT	WTT	TT	WTT
Employment	3.59	2.24	2.55	3.15
	(0.53)	(0.30)	(0.52)	(0.60)
Average labor cost	-4.20	-2.28	-4.10	-4.23
	(0.43)	(0.24)	(0.47)	(0.73)
Share of unskilled workers	1.50	0.69	1.51	0.78
	(0.37)	(0.18)	(0.34)	(0.21)
Capital	1.41	1.02	2.29	1.48
	(0.58)	(0.32)	(0.57	(0.39)
Capital labor ratio	-2.19	-1.22	-0.25	-1.67
	(0.65)	(0.36)	(0.69)	(0.52)
Productivity of capital	-1.69	-0.95	-2.76	-1.98
	(0.65)	(0.36)	(0.62)	(0.47)
Labor productivity	-3.88	-2.17	-3.00	-3.65
	(0.52)	(0.29)	(0.53)	(0.69)
Value added	-0.29	0.07	-0.46	-0.50
	(0.57)	(0.31)	(0.49)	(0.31)

Note : These figures are the semi parametric estimates of the parameter $E_3^{\varpi} = E(\varpi_i(y_i - y_i(0)))$*, obtained with and without weighting firms by their employment. They are performed on 32.459 observations in manufacturing and 48.930 in non manufacturing. Firms with a zero ex ante reduction in labor costs were discarded.*

Table 11 Unskilled workers and minimum wage

%	Unskilled	Skilled	Total
Below 1.3 minimum wage	15.66	15.27	30.82
Above 1.3 minimum wage	7.17	62.00	69.18
Total	22.73	77.27	100

Source : Employment survey 1994.

Table 12 : Employment creation and destruction rate[5]

%	1988-1991	1991-1994	1994-1997
Growth rate	7,4	-0,8	3,0
Creation rate	31,6	23,7	23,5
Destruction rate	-24,2	-24,5	-20,5

Source : BRN 1988, 1991, 1994 et 1997

---

[5] Defined following Davis and Haltiwanger 1992.

Appendix A: Model of labor demand and wage equations with heterogeneous workers.

We use a simple model to: i)- identify the factors that affect the firm's ex ante labor cost reduction, ii)- assess the conditions under which the growth rate of the variables of interest can be defined as functions of the ex ante reduction and iii)- examine whether there exists a specific source of heterogeneity in the ex ante reduction. To keep notations as simple as possible, we omit time and individual indices.

The demand of labor

We consider a firm with a technology of production that incorporates L various types of workers N_j and capital K. We assume constant elasticities of substitution between the various factors and constant returns to scale.

We suppose that the demand addressed to the firm has a constant price elasticity ε: $Q = \delta P^{-\varepsilon}$ where P is the price and δ a scale parameter.

$N = (N_1,...,N_L)'$, $c_N = (c_{N_1},...,c_{N_L})'$ et $w = (w_1,...,w_L)'$ denote respectively the vectors of employment, costs and gross wages of the L types of workers and $\pi_N = \text{Diag}(\pi_{N_1},...,\pi_{N_L})$ the diagonal matrix of their cost shares relative to total cost (capital cost included) ($\sum_{j=1}^{L} \pi_{N_j} + \pi_K = 1$). Costs and gross wages are linked by the relationship: $c_{N_j} = (1 + T(w_j))w_j$, where $T(\ )$ is the payroll tax rate function. It is that function which has been changed by the introduction of payroll tax subsidies for low wage workers.

The firm level demand of workers can be written as:

$$\log N = \Sigma_{N,N} \pi_N \log c_N + \Sigma_{N,K} \pi_K \log c_K + u$$

where $\Sigma_{N,N} = [\sigma_{N_j,N_l} - \varepsilon]_{j,l=1,...,L}$ and $\Sigma_{N,K,} = [\sigma_{N_j,K} - \varepsilon]_{j=1,...,L}$ represent the matrices of Allen substitution elasticities reduced by the price elasticity of demand and u the vector of perturbations that include technology and demand shocks.

Wage equations

The wage equations are simply written as mark-up above reservation wages:

$$\log w = \log \overline{w} + \gamma u + v$$

Where w is the vector of gross wages and $\overline{w}$ is the vector of reservation wages. The mark-up is function of the firm productivity and other components specific to the firm. Mark-ups may change over time. For example we could figure that an increase in the tax rate lead to a reduction of the mark-up.

The exante reduction in labor costs

Using the wage equations above, we can derive in each firm the expression of wages by categories of workers at date 1994 with tax rules of 1994 and 1997. This permits to deduce the firm's ex ante labor cost reduction t, which is simply equal to the sum of the L workers-level ex ante subsidies. Indeed, we have $t = t_{N_1} + \ldots + t_N$, with t_{N_j} the ex ante labor cost reduction for workers of type j defined as $t_{N_j} = \tilde{\pi}_j [\log(1 + T_{97}(w_{j,94})) - \log(1 + T_{94}(w_{j,94}))]$, $\tilde{\pi}_j$ being the share in labor costs (capital cost excluded). The firm's ex ante labor cost reduction can thus be expressed as:

$$t = t(\overline{w}_{94}, \tilde{\pi}_{N,94}, u_{94}, v_{94})$$

It is thus a function of reservation wages, the shares of various types of workers, the productivity shocks and the firm specific component in the wage setting.

Ex post variation in gross wages and employment

The growth rate of the interest variables (mainly employment and wages of the various types of workers) over the period 1994-1997 are obtained by differentiating the wage and labor equations above. Assuming the cost- shares of the different categories of workers (π_{N_j}) are held constant the vector of these shares and $t_N = (t_{N_1}, \ldots, t_{N_L})'$ the vector of the ex ante labor cost reductions associated to the various types of workers, the growth rate of wages can be written as:

$$\Delta \log w = \Delta \log \overline{w} + \gamma \Delta u - \Delta v$$

Taking the demand for labor in difference over the two periods leads to

$$\Delta \log N = \Sigma_{N,N} \pi_N \Delta \log c_N + \Sigma_{N,K} \pi_K \Delta \log c_K + \Delta u$$

and the variation in labor cost can be written as

$$\Delta \log c_N = \Delta \log(1 + T(w)) + \Delta \log w$$

The term $\Delta \log(1 + T(w))$ can be split into two terms : the variation only due to the change in the tax rule (holding gross wages constant) and the variation due to the adjustment of gross wages between 1994 and 1997 (the tax rule being that of 1997):

$$\pi_N \Delta \log(1 + T(w)) = \pi_N [\log(1 + T_{97}(w_{97})) - \log(1 + T_{94}(w_{94}))]$$

$$= \pi_N [\log(1 + T_{97}(w_{94})) - \log(1 + T_{94}(w_{94}))]$$

$$+ \pi_N [\log(1 + T_{97}(w_{97})) - \log(1 + T_{97}(w_{94}))]$$

The first term $t_N = (1 - \pi_K)\tilde{\pi}_N[\log(1 + T_{97}(w_{94})) - \log(1 + T_{94}(w_{94}))]$ is simply the vector of the ex ante reductions corresponding to the different types of workers multiplied by $(1 - \pi_K)$.

Assuming then that the second term is as a function of the growth rate of gross wages i.e. $\pi_N[\log(1 + T_{97}(w_{97})) - \log(1 + T_{97}(w_{94}))] = \pi_N A \Delta \log(w)$, we have:

$$\Delta \log N = \Sigma_{N,N} \pi_N (I + A)\Delta \log w + (1 - \pi_K)\Sigma_{N,N} t_N + \Sigma_{N,K} \pi_K \Delta \log c_K + \Delta u$$

$$= (1 - \pi_K)\Sigma_{N,N} t_N + \Sigma_{N,N} \pi_N (I + A)(\Delta \log \overline{w} + \gamma \Delta u + \Delta v) + + \Sigma_{N,K} \pi_K \Delta \log c_K + \Delta u$$

To define the changes in employment as functions of the ex ante labor cost reduction, we have to assume some restrictions on the elasticities of substitution between the $\overline{l}$ categories of workers concerned by payroll tax subsidies :

$$\sigma_{N_j,N_k} = \sigma_{N_j,N_l} = \sigma_{N_j} \quad \forall k, \forall l \le \overline{l} \text{ et } \forall j = 1,\ldots, L \;^{6}$$

This is the case if workers concerned by tax subsidies are perfect complement (sufficient condition).

Given the definition of Allen Uzawa substitution elasticities $\sigma_{i,j} = CC_{i,j}/C_i C_j$ where C is the cost function, it is straightforward to see that the condition is satisfied as long as the cost function can be written as $C(c_{N_1},\ldots C_{N_{\overline{l}}}, C_{N_{\overline{l}+1}},\ldots C_{N_L}) = C(\lambda_{N_1} c_{N_1} + \cdots + \lambda_{N_{\overline{l}}} C_{N_{\overline{l}}}, C_{N_{\overline{l}+1}},\ldots C_{N_L})$

Under this assumption, the growth rates of gross wages and employment can be expressed in each firm as :

$$\Delta \log N = (1 - \pi_K)\varsigma_N t + \Sigma_{N,N} \pi_N (I + A)(\Delta \log \overline{w} + \gamma \Delta u + \Delta v) + + \Sigma_{N,K} \pi_K \Delta \log c_K + \Delta u$$

where ς_N is the $(L \times 1)$ vector equal to $\varsigma_N = \left[\sigma_{N_j} - \varepsilon\right]_{j=1,\ldots,L}$. The interest variables (employment, gross wages, capital, value added,…) can therefore be written as :

$$\Delta y = \Delta y(\Sigma, A, \pi_N, \pi_K, \Delta \log c_K, \gamma, \Delta u, \Delta v, \Delta \log \overline{w}, t)$$

Comparing factors affecting both the ex ante labor cost reduction and the interest variables, we can note that unobserved firm level components (technology and demand for the demand side and firm level wage component for wage equations) are specific to the ex ante reduction. The variables of interest are indeed taken in evolution and thus are not dependant of these specific effects. Common factors include structural parameters, shares of inputs in total cost, capital cost and demand and technology shocks. Finally, equations can be rewritten as:

$$\Delta y = \Delta y(\Phi_{\Delta y}, t) \text{ and } t = t(\Phi_t)$$

---

[6]These restrictions imply $\sigma_{N_l,N_k} = \sigma \quad \forall \quad l, k \le \overline{l}$ et $\sigma_{N_l,N_k} = \sigma_{N_l} \quad \forall \quad l > \overline{l}$ and $k \le \overline{l}$ where $l \in \{1,\ldots,\overline{l}\}$ is the set of worker types whose wages is below the upper threshold 1.33 times the minimum wages of the change in the tax rule (which is firm specific)

Appendix C : Elimination of outliers and data set construction

The data set construction includes several steps. We first build up a balanced sample of firms from the BRN source, selecting firms in Manufacturing and Non Manufacturing sectors (except the energy, agricultural and financial sectors) present in all five years 1993-1997. We also keep information from the date of firm's first appearance in the data source BRN to compute changes of some variables used as controls.

We then remove from the sample firms for which one of our main variables had erroneous values:
- firms with non positive values for value added, number of employees, capital and wage bill,
- firms for which the annual growth rates of the value added, capital and employement are less than the 1st percentiles of these distributions or more than the 99th percentiles in the sector (2-digit).
- firms which are extreme outliers in the distributions of the logarithms and the first differences of logarithms of labor productivity, capital labor ratio, labor and capital costs.

These restrictions yield to loose 50% of the firms in Manufacturing and 60% in Non Manufacturing. This strong rate of elimination is mainly due to "clean-up" on the capital cost.

We then merge the data from BRN with DADS, selecting only firms in both data sources for the years 1993 to 1997. The merging eliminate 9% of the remaining firms in both sectors. We then realize a last clean-up on our variables of interest and control variables, which induce an additional elimination of 14% of firms. The final sample is made of about 30% of the 295,118 firms, available in the BRN initial balanced sample. It has 87,720 firms, in which 34,371 firms (39%) belong to Manufacturing and 53,349 (61%) to Non Manufacturing. Those firms employ 3,772,941 individuals, 2,053,777 (54%) in Manufacturing and 1,719,164 (46%) in Non Manufacturing.

Despite an important elimination of observations, our analysis sample provides an overall growth rate of employment which is close to that of the national accounts (see table below). Both changes indicate a fall of employment in Industry and a rise in Services. Nevertheless, our sample gives a weight too much important to Industry and not enough to Services, as compared to national accounts. Moreover, it underestimates the decrease of employment in Industry (-1,2% over the period 1994-1997 against –2,6% in national accounts) and the increase in Services (+2,5% against +5,7%). For the building sector, the growth rates of employment from both sources are however quite similar (-5,3% against –5,6%). Such differences are frequent, when comparing data from firms and national accounts.

Appendix D : Semi Parametric Quasi Maximum Likelihood estimation of the score function (manufacturing, 32769 observations and non manufacturing, 49614 observations).

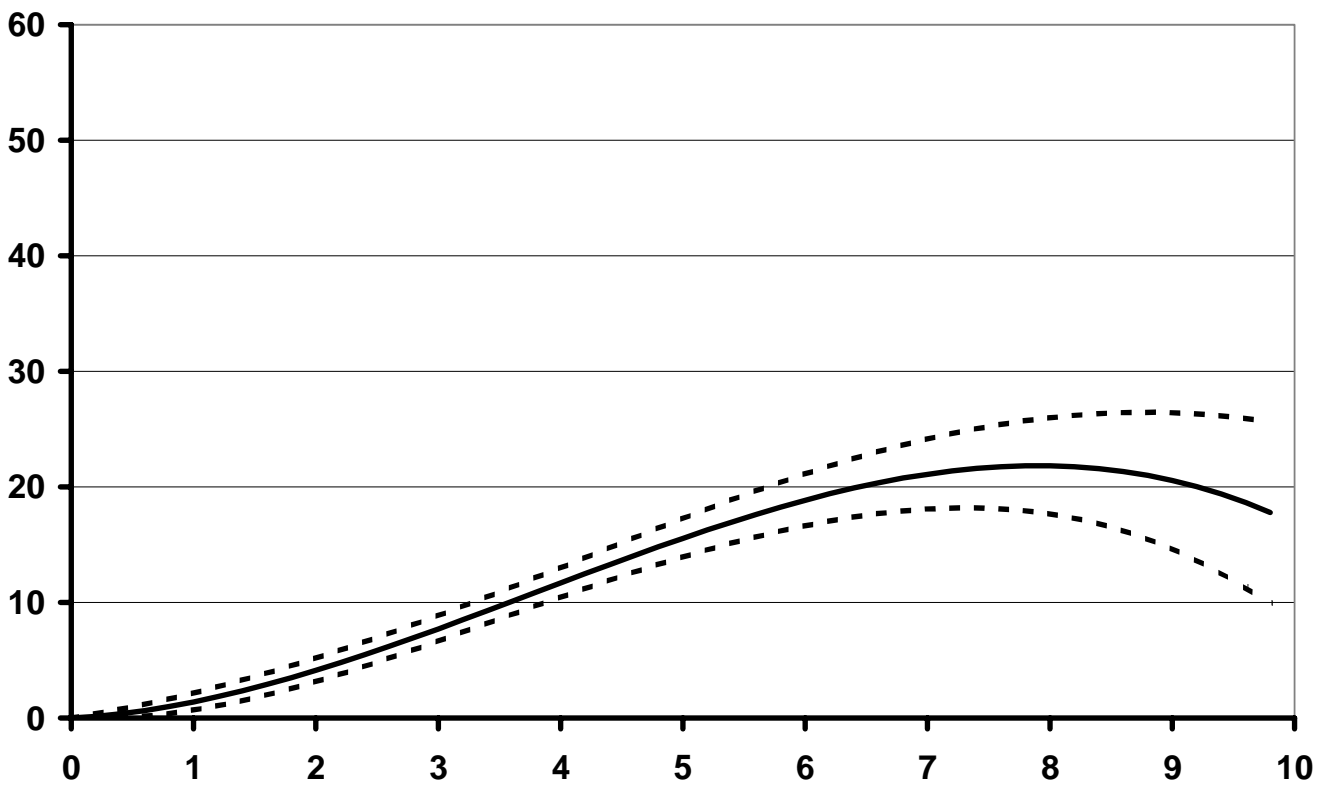| Variables	Manufacturing	Non Manufacturing
Firm level variables (general)		
Value added (log)	-0.311 (0.011)	-0.397 (0.010)
Labor productivity (log)	-0.940 (0.011)	-0.651 (0.014)
Capital labor ratio (log)	-0.081 (0.026)	-0.293 (0.019)
Operating income capital ratio	-0.070 (0.015)	-0.099 (0.012)
Firm mark-up	0.485 (0.015)	-0.604 (0.012)
Share of labor cost	-0.536 (0.023)	-0.615 (0.016)
Firm variables in difference (general)		
Value added (growth rate)	-0.064 (0.011)	-0.046 (0.008)
Total factor productivity (growth rate)	0.223 (0.013)	0.117 (0.011)
Capital labor ratio (growth rate)	0.043 (0.012)	0.061 (0.010)
Operating income capital ratio (difference)	-0.000 (0.012)	0.037 (0.008)
Firm mark-up (difference)	-0.079 (0.012)	0.059 (0.011)
Share of labor cost (difference)	0.037 (0.014)	0.059 (0.011)
Financial variables		
Cost of capital (log)	0.132 (0.013)	-0.013 (0.011)
Cost of capital (growth rate)	-0.016 (0.011)	-0.019 (0.010)
Debt ratio	0.049 (0.008)	0.017 (0.006)
Ex ante variation in the cost of capital	0.167 (0.011)	0.077 (0.010)
Work force heterogeneity		
Share of young unskilled men	0.068 (0.009)	0.075 (0.007)
Share of prime age unskilled men	-0.052 (0.009)	0.069 (0.008)
Share of old unskilled men	-0.018 (0.008)	0.028 (0.008)
Share of young skilled men	0.054 (0.010)	0.057 (0.007)
Share of prime age skilled men	----------------	----------------
Share of old skilled men	0.002 (0.009)	0.004 (0.008)
Share of young highly skilled men	-0.018 (0.011)	0.025 (0.007)
Share of prime age highly skilled men	-0.234 (0.011)	-0.160 (0.009)
Share of old highly skilled men	-0.130 (0.009)	-0.102 (0.008)
Share of young unskilled women	0.073 (0.011)	0.161 (0.007)
Share of prime age unskilled women	0.195 (0.010)	0.266 (0.009)
Share of old unskilled women	0.056 (0.010)	0.099 (0.007)
Share of young skilled women	0.059 (0.013)	0.069 (0.006)
Share of prime age skilled women	0.100 (0.009)	0.149 (0.008)
Share of old skilled women	0.022 (0.008)	0.051 (0.007)
Share of young highly skilled women	0.013 (0.008)	0.043 (0.006)
Share of prime age highly skilled women	-0.039 (0.011)	-0.007 (0.009)
Share of old highly skilled women	-0.022 (0.009)	-0.035 (0.008)
Sectoral variables		
Average labor cost of unskilled worker (log two digit level)	0.092 (0.012)	-0.180 (0.009)
unskilled worker (growth rate two digit level)	-0.021 (0.012)	-0.028 (0.008)
Entry rate (three digit level)	-0.142 (0.010)	0.167 (0.005)
Exit rate (three digit level)	0.044 (0.009)	-0.042 (0.007)
Import rate (two digit level)	-0.325 (0.022)	0.579 (0.030)
Export rate (two digit level)	0.103 (0.020)	-0.726 (0.031)

Note : Also include sector dummy variables at the one digit level. Level variables are taken in 1994, variables in difference or growth rate are taken over the longest period available when at the firm level and over the period 1990-1994 for variables at the sectoral level. All variables have been centered and reduced. The coefficient of the intercept and one dummy sectoral variables have been constrained to equal there estimated values in the linear estimation of a the logistic transformation of the treatment variable $\tilde{t} = \log(t/(0.10 - t))$.

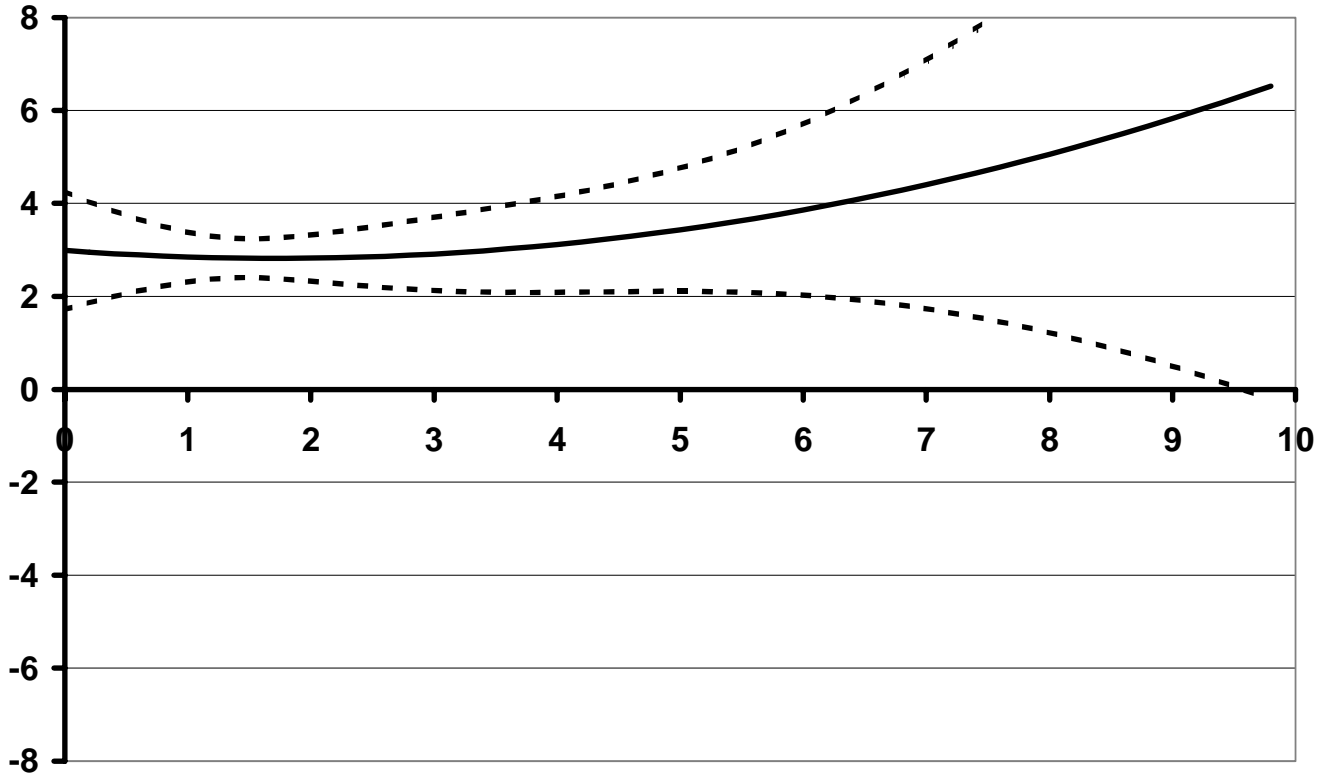## Appendix E1 : Estimation of parameter $E(y(t) - y(0))$
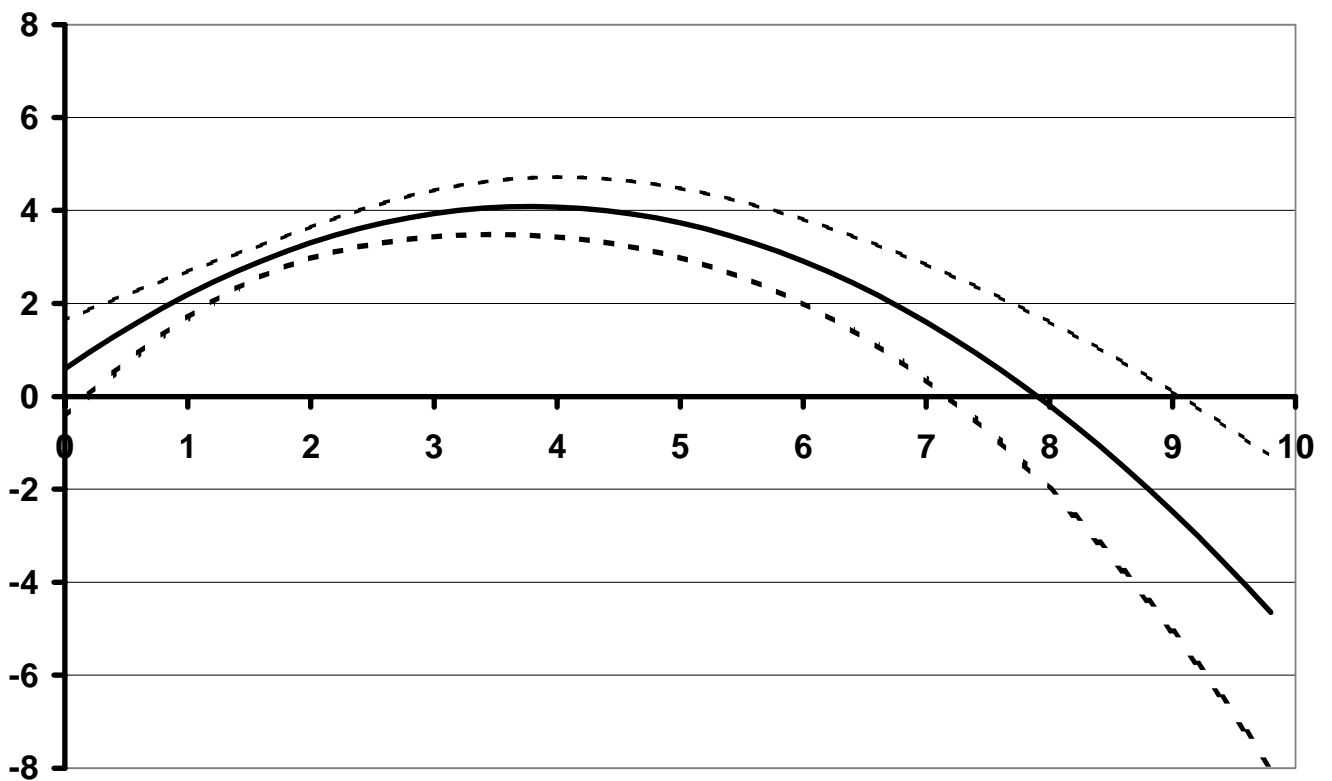### Manufacturing
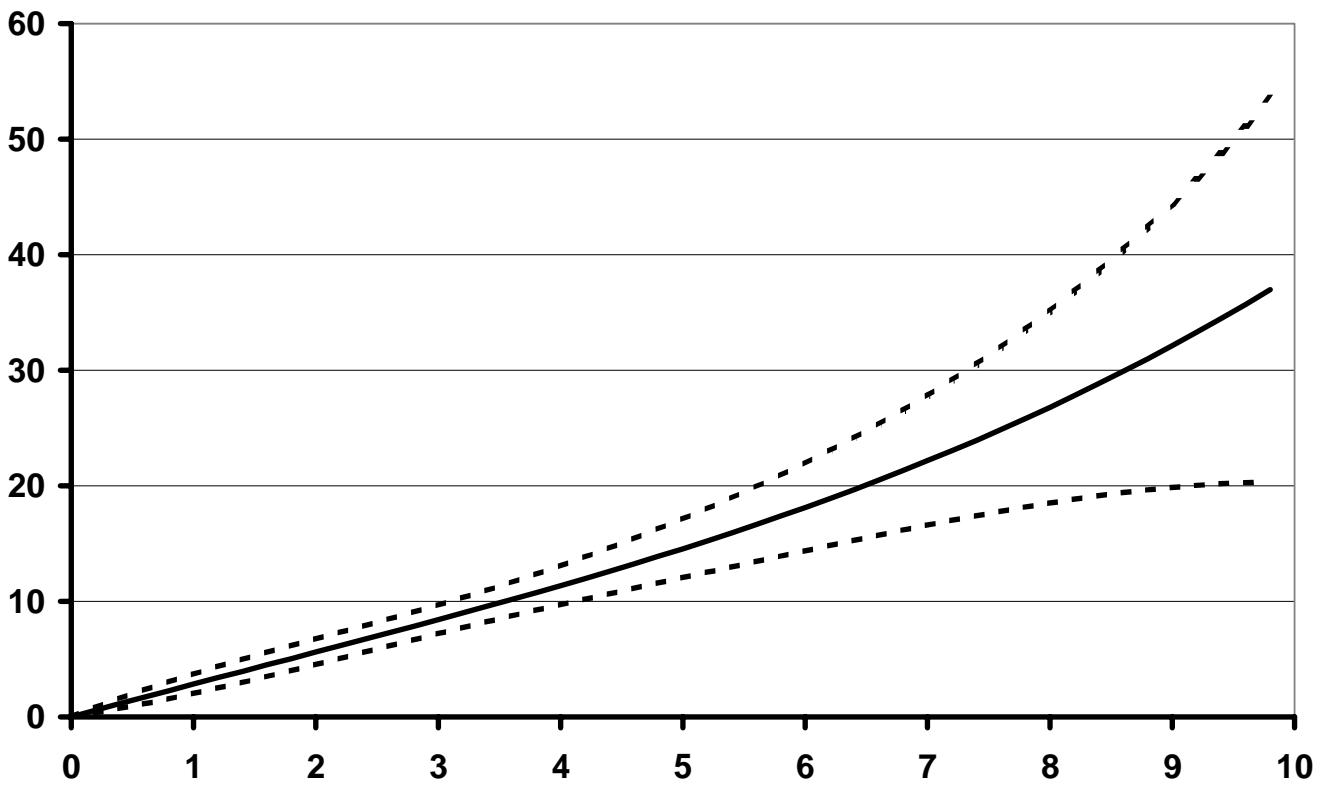
Non Manufacturing

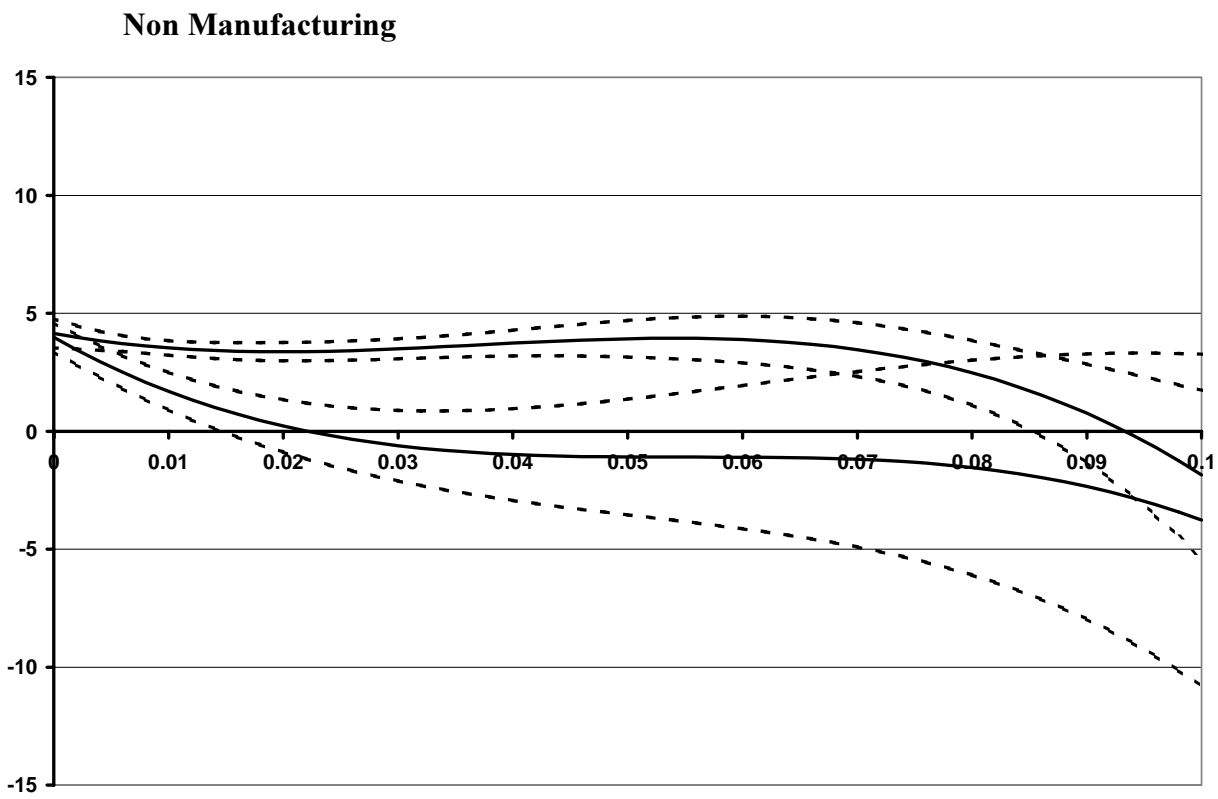# Appendix E2 : Estimation of parameter $E(\partial y(t)/\partial t)$

Manufacturing

Non Manufacturing

Appendix E3 : Estimation of parameter $E(y(0)|t)$ and $E(y|t)$

Manufacturing

Non Manufacturing
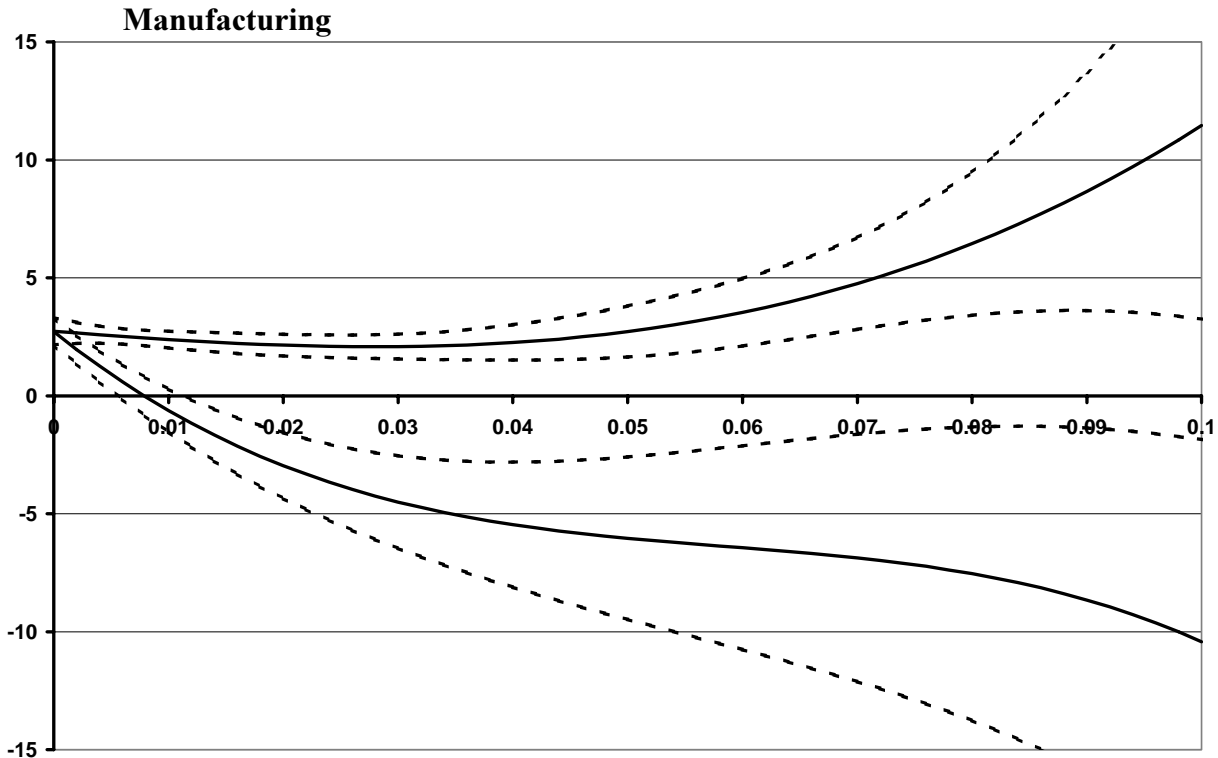
Note : Obtained with series estimator with a degree of polynomial approximation of 3. The solid line are the estimated values. The doted one are respectively the lower an upper bound at 5% obtained with 500 replication of the bootstrap. The upper solid line is the expectation of output given treatment. The lower solid line is the expectation of the counterfactual given treatment.
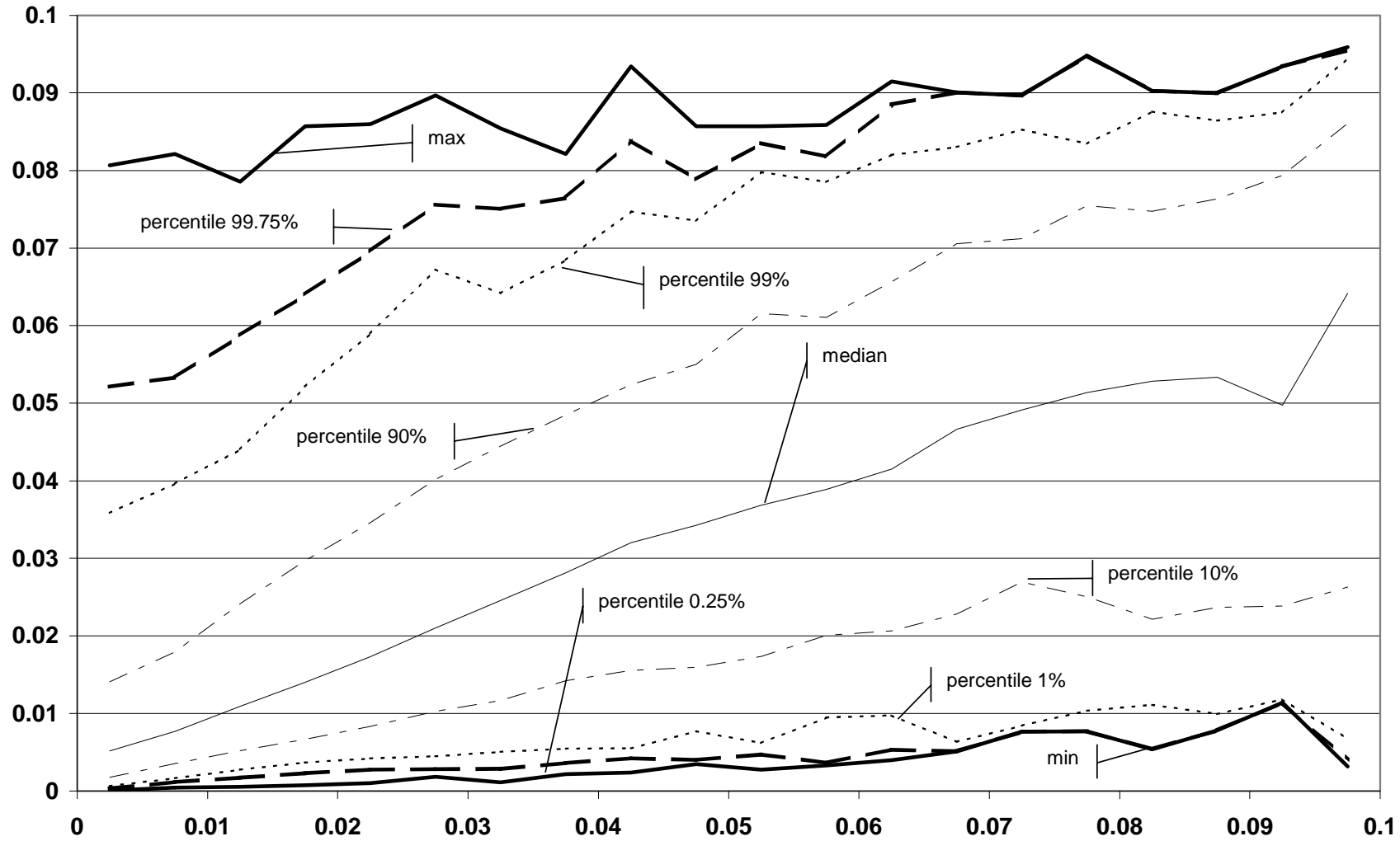
Appendix F : Common support

Figure Fa : Percentiles of the score distribution by class of treatment –manufacturing
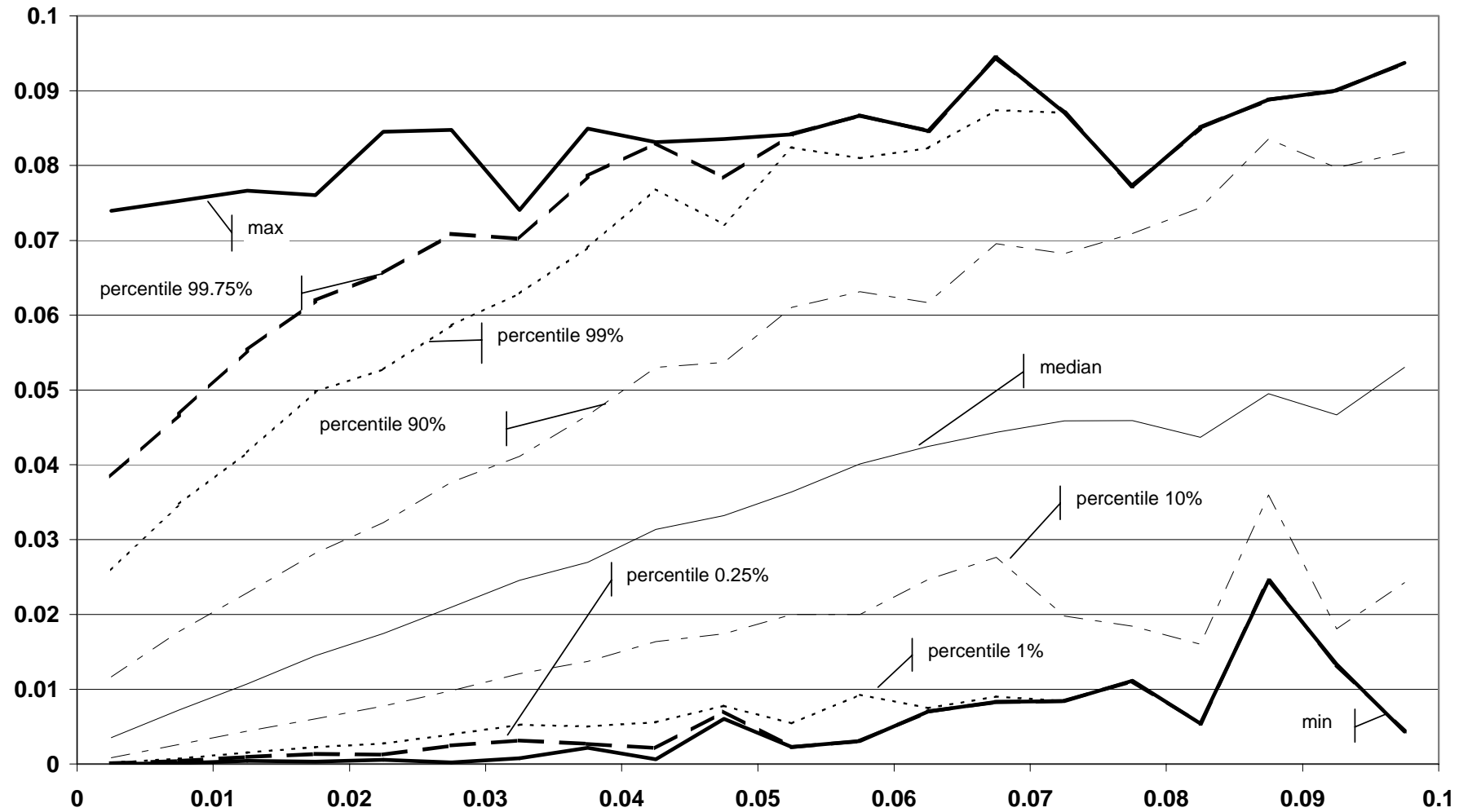
Note : 20 classes of treatment, width 0.5% : $[0,0.5\%[$, $[0.5\%,1\%[$, $[1\%,1.5\%[$, $[1.5\%,2\%[$, …

Figure F2a : Percentiles of the score distribution by class of treatment –non manufacturing

Note : 20 classes of treatment, width 0.5% : $[0,0.5\%[$, $[0.5\%,1\%[$, $[1\%,1.5\%[$, $[1.5\%,2\%[$, …