Explaining Firms efficiency in the Ivorian Manufacturing sector: A robust nonparametric approach.

Nolwenn Roudaut[*]
University of South Brittany, IREA

Anne Vanhems[†]
University of Toulouse, GESCT

March 4, 2009

Abstract

In this work we use a one-step nonparametric methodology to estimate and explain technical efficiency levels of Côte d'Ivoire firms. We consider external environment factors that migh influence the production process and estimate conditional technical efficiency levels using a robust nonparametric method (the conditional expected frontier of order m).We analyze the sensitivity to categorical environmental factors by comparing conditional and unconditional frontiers with bootstraps methods.

1 Introduction.

Measurement of efficiency in production allows to understand better the production process, to identify obstacles to firm functioning and give possible ways to enhance firms performance and development. Since the seminal works of Koopmans (1951), Debreu (1951) and Farrell (1957), a considerable amount of studies have proposed methods to estimate firms technical efficiencies. Generally speaking, the idea is to analyze how firms combine their inputs to produce their outputs in an efficient way. The maximal achievable level of outputs for a given level of inputs defines the production frontier. The technical efficiency of a particular firm is then characterized by the distance between its level of outputs and the optimal level that should be produced if efficient (see Shephard 1970 for more details). In literature on efficiency analysis, the nonparametric approach has received a great interest, mainly because it does not require the specification of a functional form for the frontier. Data Envelopment Analysis (DEA) and Free Disposal Hull (FDH) are among the most known and applied nonparametric techniques (see Cooper, Seiford and Tone, 2000 for a general overview). Nevertheless, these traditional nonparametric approaches present some severe limitations that should be carefully considered. In particular, results are very sensitive to outliers and extreme values. In addition, unsatisfactory techniques are used for the introduction of environmental or external variables in the estimation of efficiency (see Daraio and Simar 2007). Robust nonparametric approaches have recently been proposed to overcome the sensitivity effect: order-m frontiers (Cazals, Florens and Simar 2002) and α-quantile frontiers (Aragon, Daouia, Thomas-Agnan 2005). Then these

[*]nolwenn.roudaut@univ-ubs.fr
[†]a.vanhems@esc-toulouse.fr

methods have been adapted to estimate conditional frontier models by Daraio and Simar (2005, 2007).

The evaluation of the influence of external-environmental factors on efficiency of producers is a relevant issue related to the explanations of efficiency and the identification of economic conditions that create inefficiency. A great majority of methods using DEA technics are based on a two-stage approach, with extensions to three-stage and four-stage analysis (see Fried, Schmidt and Yaisawarng 1999 or Fried, Lovell, Schmidt and Yaisawarng 2002). As pointed out by Simar and Wilson 2007, usual inference on the obtained estimates of the regression coefficients in the second stage is not valid and need to be corrected. Moreover, these multi stage approaches rely on a very strong structural separability condition between the input-output space and the space of environmental factors. The nonparametric robust conditional frontier model developed by Daraio and Simar 2005 overcomes these drawbacks by offering a one stage robust method of estimation.

Daraio and Simar (2005) propose a one-step fully nonparametric procedure to take into account environmental variables but their method is only valid for continuous environmental variables. Since we consider mainly dichotomic environmental variables, we adapt their method to discrete variables using the kernel function introduced by Aitchison and Aitken (1976) and applied in Li and Racine (2004). We then analyse the influence of factors using bootstrap technics developed by Florens and Simar (2005). We pay a particular attention to the sensitivity of our estimator to smoothing parameters, relying on Simar (2003) analysis and Badin, Daraio and Simar (2009) paper.

The objective of this work is to estimate and explain technical efficiency levels of Côte d'Ivoire firms using this recent method of robust conditional frontier. We analyse a database from Côte d'Ivoire collected in 1995 and 1996 within the RPED framework ("Regional Program on Enterprise Development in Africa"). Efficiency measurement have been performed on this database in at least two papers. Roudaut (2006) studies the impact of business environment on technical efficiency using a stochastic parametric method (see Battese and Coelli 1995 for an overview on these methods) and interpret the effect of each business environnement variable via the estimated parameters. On the same database, Chapelle and Plane (2005) apply nonparametric DEA technics and stress the importance of size in managerial performance. Their method in four steps allows to capture three effects: managerial, scale of production and technological effect. The DEA methodology allows for a more general form for the production frontier estimator (see Seiford 1996 for example).

The paper is organized as follows. The methodology is presented in Section 2. We introduce the database and the main variables of interest in Section 3 and results and comments are given in Section 4.

2 Methodology.

In this section, we present the methodology developed by Daraio and Simar 2005 and adapted to discrete environmental variables case. The choice of the smoothing parameters and the analysis of sensitivity to environmental factors by bootstrap technics are particularly emphasized.

2.1 The conditional expected frontier of order-m.

In what follows, we consider the nonparametric conditional expected frontier of order m estimator developed by Cazals, Florens, Simar (2002) (CFS hereafter) and Daraio and Simar (2005). This estimator is robust to extreme values because it does not envelop all the data cloud. It is based on quite easy practical computations and benefits from a relatively high rate of convergence[1]. Note that in our setting the environmental factors are discrete variables.

Although the method can be applied in multivariate settings (multi-inputs multi-outputs technology), for sake of simplicity and to be closer to our empirical work, we only present the case of a single output produced by several inputs. For a more general and complete definition, we refer to CFS 2002. Let X denotes a vector of size $p \geq 1$ of input levels, Y the production level and Z denotes a vector of size $r \geq 1$ of environmental factors. The factors Z provide additional information, they are exogenous to the production process but may explain part of it. One way to introduce this additional information in a one-step estimation is to condition the production process to a given value of Z. In other words, the joint distribution of (X, Y) conditional on $Z = z$ defines the production process if $Z = z$.

In what follows, we present the conditional measure developed for the output oriented case. The Farrell-Debreu output efficiency score can be adapted to the conditional efficiency score in the following way. Consider F the distribution function of (X, Y, Z) defined on $\mathbb{R}_+^p \times \mathbb{R}_+ \times \mathbb{R}^r$ and $S_{Y|X,Z}(y\,|x, z) = \Pr(Y \geq y\,|X \leq x, Z = z)$[2]. We obtain the definition of the output conditional full-frontier efficiency measure: $\lambda(x, y\,|z) = \sup\left\{\lambda \,\middle|\, S_{Y|X,Z}(\lambda y\,|x, z) > 0\right\}$.

Consider a fixed integer $m \geq 1$. The conditional order-m output efficiency measure can be computed as:

$$\lambda_m(x, y\,|z) = \int_0^\infty \left[1 - \left(1 - S_{Y|X,Z}(y\,|x, z)\right)^m\right] du \tag{1}$$

By definition (see Farrell 1957), a firm is efficient if its efficiency parameter λ is equal to 1. It is inefficient if $\lambda > 1$. Contrary to standard methods of frontier estimation, the definition of expected frontier of order m allows some points to be above the frontier. This property is a direct consequence of the robustness of this estimator. Some firms can be characterized as super-efficient with $\lambda_m < 1$. It can be shown that $\lim_{m\to\infty} \lambda_m(x, y\,|z) = \lambda(x, y\,|z)$.

2.1.1 Nonparametric estimation.

Consider a *iid* sample of size n (X_i, Y_i, Z_i) derived from the random vector (X, Y, Z). A nonparametric estimator is then given by:

$$\widehat{\lambda}_m(x, y\,|z) = \int_0^\infty \left[1 - \left(1 - \widehat{S}_{Y|X,Z}(y\,|x, z)\right)^m\right] du \tag{2}$$

[1]Note that there exists another recent nonparametric method which presents the same characteristics as the expected frontier of order m. The alternative estimator is the conditional quantile-based estimator, introduced by Aragon, Daouia and Thomas-Agnan (2005). Both estimators are equivalent (see Daraio and Simar 2007).

[2]Inequalities involving vectors are defined on an element-by-element basis, e.g. for any $(x, x') \in \left(\mathbb{R}_+^p\right)^2$ such that $x' \geq x$, some elements of both vectors are equal and some elements of x' are greater than the corresponding elements of x.

where $\widehat{S}_{Y|X,Z}$ represents a nonparametric estimator of $S_{Y|X,Z}$:

$$\widehat{S}_{Y|X,Z}(y\,|x,z) = \frac{\sum_{i=1}^{n} \mathbb{1}\,(X_i \le x, Y_i \ge y)\,K(z, Z_i, h_n)}{\sum_{i=1}^{n} \mathbb{1}\,(X_i \le x)\,K(z, Z_i, h_n)} \tag{3}$$

where K is the kernel function with compact support and h_n is the bandwidth of appropriate size (with the property $h_n \to 0$ when $n \to +\infty$). Since $\widehat{\lambda}_m$ converges to the FDH estimator when m goes to infinity, the order-m conditional efficiency score can be viewed as a robust estimator of the conditional efficiency score λ. For finite m, the corresponding attainable set will not envelop all the data points and is more robust to extremes or outliers. (For asymptotic properties, see CFS 2002).

In practice, equation (2) is approximated using Monte-Carlo methods (see Daraio Simar 2005 for more details). Note that the number of replications B to adjust approximation quality is usually fixed greater or equal to 200 for most empirical applications.

2.1.2 Smoothing parameters choice.

In this part, we define the kernel function K associated to categorical variables (non ordered or ordered) and set the values of the bandwidth parameter h_n and the smoothing parameter m.

Kernel function choice. All environmental variables we consider are categorical variables, the age with ordered categories, and all others dichotomic variables. Therefore, we need to use an adapted kernel function K. Aitchison and Aitken (1976) first introduced a kernel method adapted to multivariate binary variables, and Li and Racine (1994) extended it to ordered and non ordered categories. Following their notations, the kernel function for a non ordered univariate categorical variable Z is defined by:

$$K_{no}(z, Z, h_n) = \begin{cases} 1 & if \quad Z = \quad z \\ h_n & if \quad Z \neq \quad z \end{cases}$$

In case of an ordered univariate categorical variable Z with c different values, the associated kernel function is defined by:

$$K_o(z, Z, h_n) = h_n^s$$

where $|z - Z| = s$ and $0 \le s \le c$. In equation (3), we will replace the general kernel function K by either K_{no} or K_o (or the product of both if necessary).

Smoothing parameter h_n. It is known that the choice of smoothing parameters is of crucial importance in nonparametric kernel estimation. Following Daraio and Simar (2005), we choose h_n in order to minimize a cross-validation criterium based on the marginal density of Z. An alternative would be to use the method developed by Badin, Daraio and Simar (2009): an adaptive data-driven method which optimizes the estimation of the conditional survival function $S_{Y|X,Z}$.

Choice of m. As for any nonparametric estimator, the empirical choice of the smoothing parameter m is very important. It affects in particular the shape of the estimator and its rate of convergence. When m is too small, a lot of points are not considered in the frontier estimation and are above the frontier. When m is large, the frontier takes into account almost all the points but is less smooth and robust, it converges to the FDH frontier (see Deprins et all 1984). CFS provide an asymptotic optimal expression of $m(n)$ but no empirical rule of thumb.

In the application, we decide to fix m in order to keep a sufficiently large number of points below the production frontier. The points that are not taken into account in the efficiency estimation, above the m frontier, can be outliers or points from the DGP that are structurally excluded. In order to take into account this distinction, we apply the methodology developed by Simar (2003). According to expected frontier definition, some observations are located above the estimated frontier, even for large values of m. Even so, being located above the frontier did not necessarily define an extreme value: one has to fix a threshold value t in order not to exclude observations that are above but somehow close to the frontier. In practice two issues emerge : which values of m one must choose and from which t observations can be considered as outliers? Simar method does not offer straight answers and relies on a sensibility analysis of results. The author proposes the following "semi-automatic" method. First, one must choose some arbitrary threshold values t. For each value, the proportion of observations for which efficiency index is inferior to $1 - t$ is evaluated. The calculus of this proportion for each threshold and for various values of m must enable to detect extreme observations. If the sample exhibits no extreme values then this proportion must be linearly decreasing with m values. Hence any deviation from the linear form, or the presence of an "elbow", can reveal the presence of potential extreme values.

We provide in Appendix a Table measuring the sensibility to extreme values, and finally chose $t = 0.1$, such that, below this level, there is almost no more super-efficient firms.

2.2 Analysis of sensitivity to environmental factors.

Daraio and Simar (2005) developped a useful methodology allowing to detect the effect of some environmental factor on the performance on firms. The idea is to analyze the ratio of the conditional efficiency scores over the unconditional scores as a function of the conditioning factors Z. The nonparametric estimator defined above $\widehat{\lambda}_m$ is consistent when the sample size goes to infinity and the smoothing parameters go to zero quickly enough (see CFS or Daraio and Simar for more details). But the rate of convergence is deteriorated as the dimension in the input/output space increases. Therefore, in order to take into account our small sample size, we analyze the effect of each environmental variable independently from the others.

We then calculate an efficiency score for each factor Z and consider the ratio: $Q_z = \frac{\widehat{\lambda}_m(x,y|z)}{\widehat{\lambda}_m(x,y)}$ as a function of z. For a dichotomic variable Z, we calculate the mean ratio for $Z = 1$ et $Z = 0$ and compare both using bootstrap technics. We apply the naive bootstrap methodology developed in Florens and Simar (2005)[3]. For the ordered variable, we regress the ratio Q_z with respect to Z and analyze its trend.

[3] Florens and Simar 2005 justify in their article why the naive bootstrap can lead to good results in this particular case of m-frontier estimation.

3 Database.

Our empirical analysis is performed on a dataset of a representative sample of manufacturing firms in Côte d'Ivoire for the years 1994 and 1995. The data come from a survey conducted in the framework of the World Bank project RPED (Regional Program on Enterprise Development). Due essentially to missing values, our final sample restricts to 195 firms in 1994 and 174 in 1995[4] .The survey was built in order to represent four sectors of production: agro-industries, textile, wood and metal working. Inside each sector, there exists a huge heterogeneity in terms of technology. Since our objective is to estimate efficiency via production functions, it is important to define subsectors which are as homogeneous as possible in terms of technology. Note that some studies keep the initial segmentation (like Chapelle and Plane 2005)). However, in this work, we follow the methodology of Roudaut 2006, and distinguish small activities close to craft industries ("Low technology") from the more technologically intensive ones ("High technology"). The subsectors with High levels of Technology and investment (HT) are the textile industry, the timber industry, agro industry and bakeries. Low technology (LT) subsectors include wood furniture making, confection and metal-working.

3.1 Production variables.

Output level vector Y is evaluated by sales values, corrected with stock variations. The matrix X refers to the 2 inputs used: K (capital), L (labor). Labor is evaluated as the total number of hours worked in the firm per week. Capital definition is made on a permanent inventory basis.

3.2 Business environment variables.

In this section we describe business environment characteristics of firms. Firms, even in the same country and sectors, can be confronted to constraints of different kind and magnitude: various laws and taxes, labor regulations, reputation, networks, domestic and foreign competition, among others, all of which is likely to affect a firm's production behavior and so its technical efficiency. Paying no heed to these features can lead to misleading estimations of frontiers and of efficiency levels.

Summary statistics of the business environment variables used in estimations are presented in Appendix in Table (1).

- Formal vs informal.

In Côte d'Ivoire, as in most developing countries, the industrial sector is composed of a modern sector (often called the *formal* sector) and an *informal* sector. The criterion used here to define the informal sector is the one of legality; others definitions can be used (small size, low capital intensity...), but the legacity criterium is the most often used. This survey is particularly interesting because contrary to usuals studies on African firms, it contains information on both formal and informal sectors, and thus allows to make comparisons between them. Both sectors have very different characteristics: in terms of technology of production, and in terms of economic environment constraints. Informal firms use basic technologies, produce simple goods and are not exposed to laws, taxes and various regulations whereas formal firms use more

[4] Initially 234 firms were surveyed in 1994 and 230 in 1995.

capital-intensive technologies and face regulation constraints. Remark that as the informal firms use mainly basic technology, they are not represented in the HT sector..The dummy variable DFORM takes the value 1 if the firm belongs to the formal economy.

- Unions.

At the time of the study, one main union (UGTCI) is involved in all industrial sectors in Côte d'Ivoire. However because this union was affiliated to the ruling party, the PDCI (Azam and Morrisson, 1994), it has been accused of having a fairly cooperative behavior. Almost 80% of the formal firms have union member workers.The unionization rate is high: more than 95% of firms with union workers have an unionization rate of 100%. The dummy variable DSYND takes the value 1 if at least one worker of the firm belongs to an union.

- Market conditions: external competition and demand.

Level of competition is obviously different in each market. As the level of domestic competition is hard to define with our data, we choose to focus only on external competition. In our sample, between 50 et 60% of the formal firms are exporters. Few informal firms declare being exporters, but it is really a marginal feature. Most of firms are not exporter specialists : the average level of the propensity to export is about 60%. The proportion of firms which export to Africa is similar to that of firms which export to the rest of the world. Concerning the relation between efficiency and exports, the intuition is that foreign oriented firms are more efficient, because of the *self selection* and *learning by exporting* effects. DPROPEX is a dummy variable which takes the value one if the firm exports a share of its production.

- Ownership

Informal firms are all sole proprietorships (entrepreneurial firms) but formal firms may be sole proprietorships or have a more complex organizational form (managerial firms). These two types of structure for formal firms can give rise to different problems of monitoring, for example agency costs, which can be linked to efficiency. The dummy variable DMAN takes the value 1 if the owner is also the manager.

- Age

4 Empirical application.

For the application, we first provide a sensibility analysis with respect to the parameters of bootstrap and m. We finally present the results with $B = 500$ replications for the score calculus and the confidence interval around the mean ratio. About the choice of the smoothing parameter m, we present the results for $m = 2.n_x$, with n_x the number of firms that use less than a value x of inputs (n_x can be calculated for each firm of the sample). The choice of m is motivated by the last tables in Appendix. The value corresponds to the threshold with 5% extreme values below 0.9. Let's note that as m increases, the effects of Z disappear as noted in Daraio and Simar (1997).

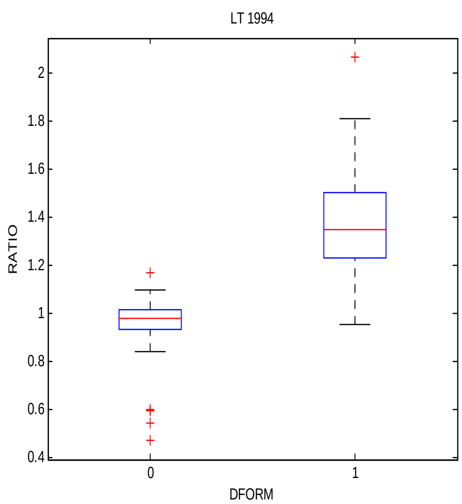
We analyze the behavior of the ratio Q_Z for the four subsamples (LT 1994, LT 1995, HT 1994, HT 1995) and each environmental factor Z. When Z is a dichotomic variable, we provide boxplots of Q_Z for $Z = 0$ and $Z = 1$. In order to compare both mean ratios, we present bootstrapped confidence intervals with different risk values. At last, for the variable Age, We plot the score ratio with respect to the factor Z with a nonparametric regression (using Li and Racine 2004).

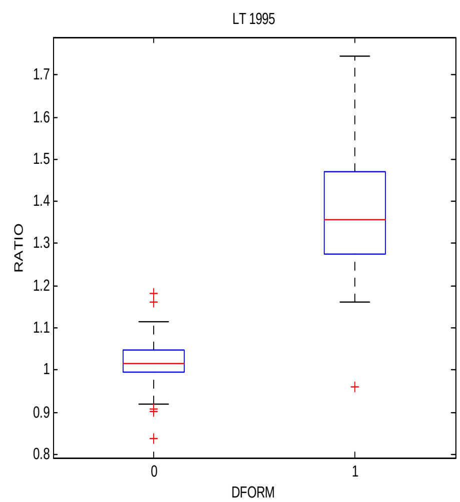
The interpretation of the effect of Z is the following: an increasing ratio Q_Z demonstrates a favorable environemntal factor, " which means that the environmental variable operates as a sort of extra input freely available: for this reason, the environment is favorable to the production process" (Daraio Simar 2007).

In case of unfavorable Z, the environmental variable penalizes the production of output, and the ratio Q_Z is decreasing.

4.1 Formal/informal effect.

The distinction betwenn formal and informal firms appears only for the Low Technology sector. Lets analyze the impact of formal sector on productivity. The bootstrapped confidence intervals show that with a 10% risk, both mean ratios are different and Q_Z is greater in mean for formal firms. This conclusion is robust for both years 1994 and 1995. It is not surprising at all that informal firms are on average less efficient than formal firms, it was also the conclusion given by Roudaut (2006). Note that bootstrapped intervals are quite close, and when m increases, there is no more difference beween both mean ratios.





Dform LT 1994

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.4298	0.5740	0.6189	**0.9690**	1.1937	1.2579	1.3975
Z=1	0.9338	1.1503	1.2096	**1.3771**	1.7703	1.8452	1.8996

Dform LT 1995

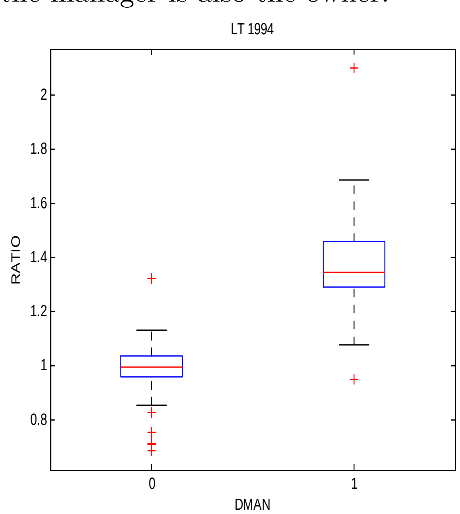
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.7059	0.7492	0.8075	**1.0098**	1.1956	1.2308	1.2835
Z=1	1.1457	1.2248	1.2755	**1.3716**	1.7059	1.7445	1.7797

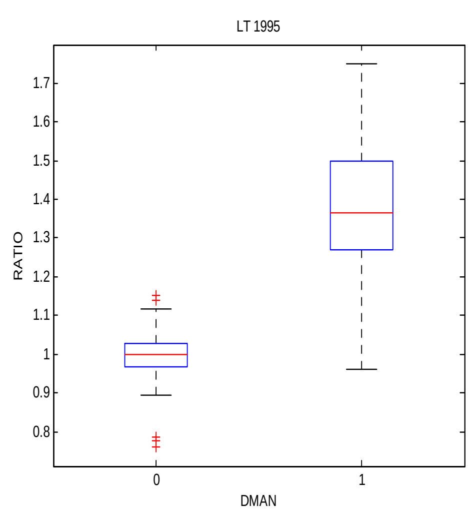
4.2 Ownership effect.

We now analyze the effect of the manager owing the firm or not. We detail the analysis for both sectors LT and HT.

4.2.1 Low technology.

It appears that for the low technology sector, we obtain significantly different mean ratios, the result is again robust for the two years of analysis. With a 10% risk, the efficiency increases when the manager is also the owner.





Dman LT 1994

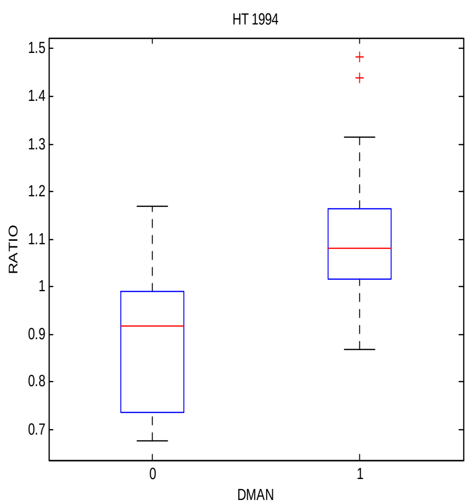
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.5501	0.5978	0.6766	**0.9860**	1.2052	1.2341	1.2723
Z=1	1.0512	1.1320	1.2205	**1.3878**	1.8551	1.9212	2.0825

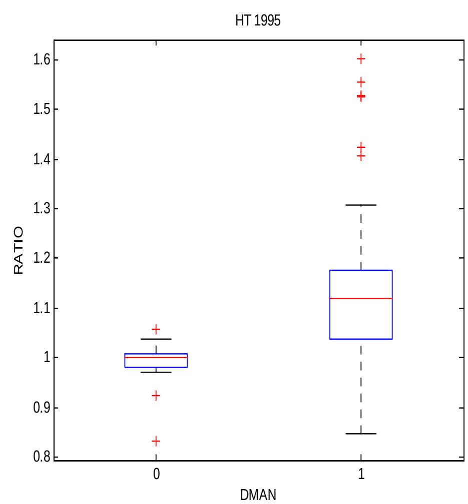
Dman LT 1995

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.5181	0.7122	0.7782	**0.9915**	1.1406	1.1654	1.2178
Z=1	0.9514	1.2226	1.2818	**1.4052**	1.8565	1.9224	2.0261

4.2.2 High technology.

On the contrary, for high technology sector, there seem to be no significant effect of ownership on efficiency. The large deviaton in each subsample seems to explain this result.





Dman HT 1994

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.3736	0.4575	0.5101	**0.8896**	1.1387	1.1687	1.2265
Z=1	0.7366	0.8038	0.8648	**1.1110**	1.4048	1.4868	1.5827

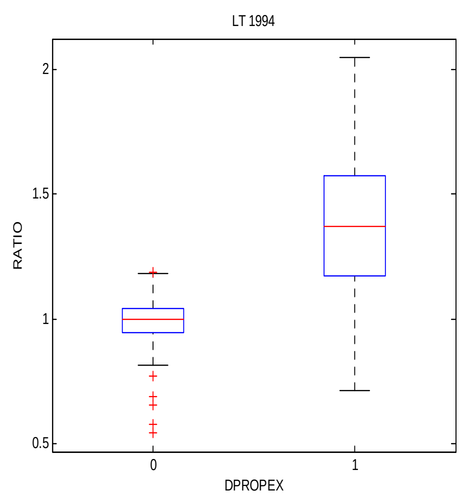
Dman HT 1995

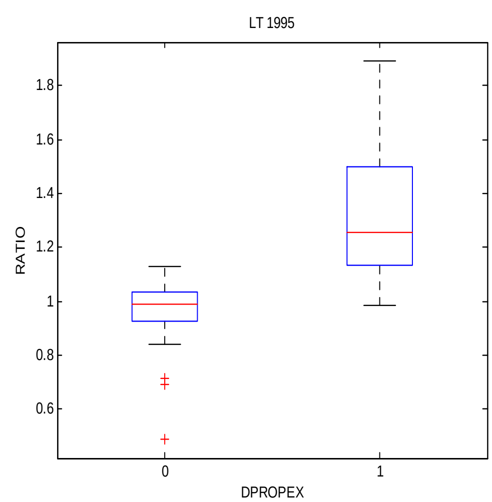
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.7804	0.8192	0.8655	**0.9841**	1.1692	1.1944	1.2924
Z=1	0.6456	0.7784	0.8284	**1.1151**	1.5482	1.6884	1.7893

4.3 Exportation effect.

Again we notice a positive effect of exportation on efficiency for law technology sector and no effect for high technology sector. The result is robust for both years.

4.3.1 Low Technology.





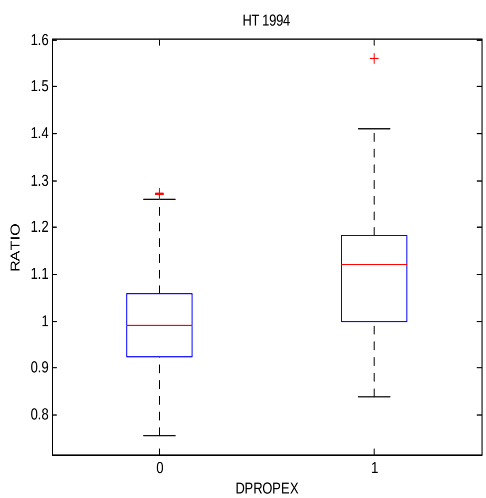
Dpropex LT 1994

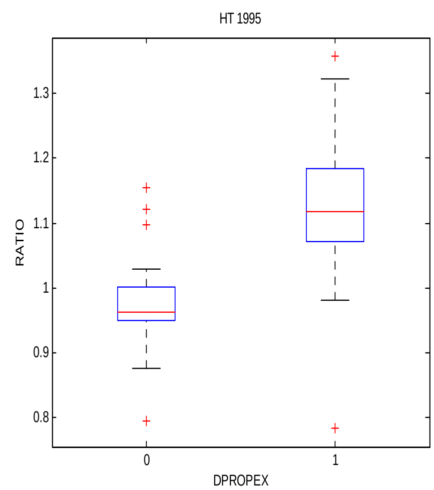
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.6424	0.6847	0.7436	**0.9849**	1.1312	1.1590	1.1924
Z=1	0.9117	1.0785	1.1557	**1.3274**	2.0547	2.1805	2.3054

Dpropex LT 1995

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.6980	0.7779	0.8229	**0.9646**	1.1350	1.1468	1.1640
Z=1	1.0877	1.1681	1.2437	**1.3905**	1.9831	2.0768	2.2764

4.3.2 High Technology.





Dpropex HT 1994

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.4603	0.5639	0.6113	**0.9894**	1.3084	1.4293	1.4799
Z=1	0.7751	0.8555	0.9110	**1.0776**	1.3491	1.3776	1.4502

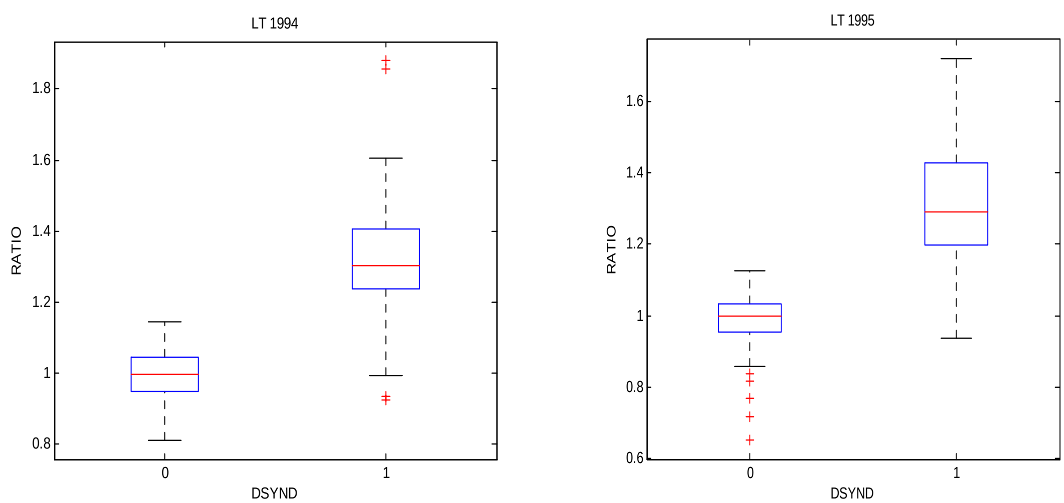
Dpropex HT 1995

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.6313	0.6836	0.6964	**0.9618**	1.4306	1.5081	1.6194
Z=1	0.8184	0.9162	0.9734	**1.1296**	1.4639	1.5065	1.5875

4.4 Union effect.

We observe the same positive effect of union on efficiency for low technology, with a 10% risk, and no effect for high technology. It seems that environmental variables are more influent on low technology sector.

4.4.1 Low Technology.



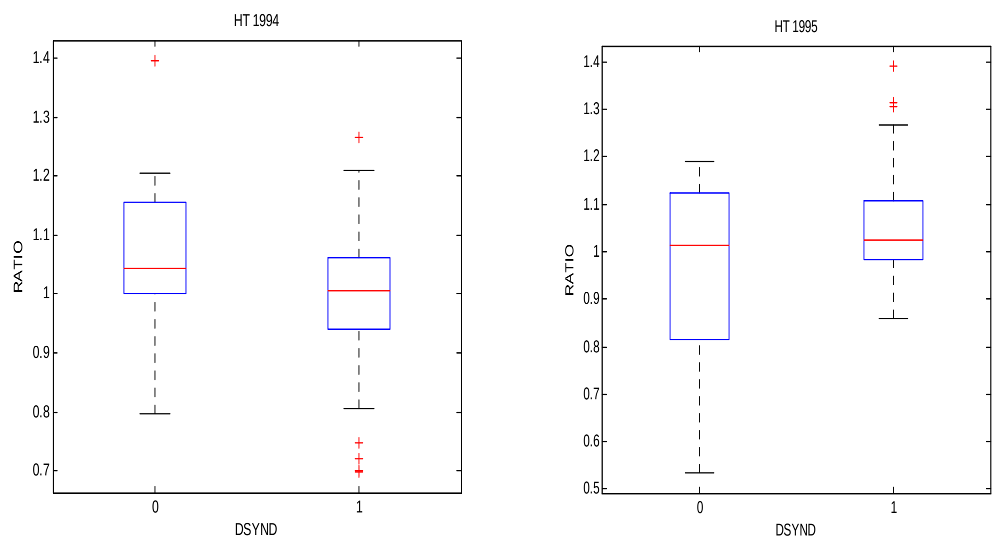
Dsynd LT 1994

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.4599	0.6170	0.6730	**0.9770**	1.1800	1.2390	1.2772
Z=1	1.0821	1.1133	1.1889	**1.2583**	1.6389	1.7283	1.7424

Dsynd LT 1995

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.7127	0.8180	0.8458	**0.9756**	1.1584	1.1846	1.2041
Z=1	1.0355	1.1203	1.1797	**1.3079**	1.7666	1.8250	2.0039

4.4.2 High Technology.



Dsynd HT 1994

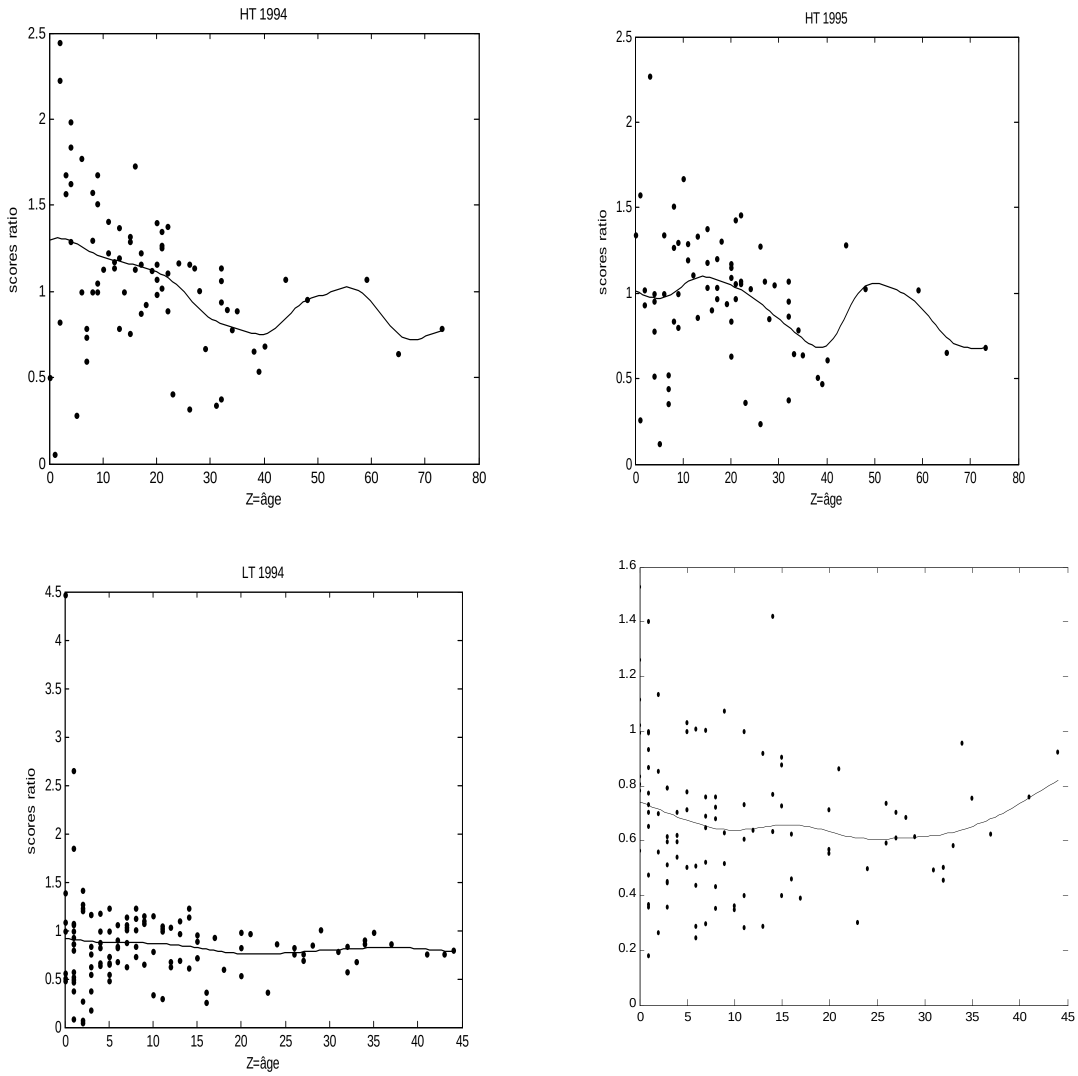
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.2208	0.5152	0.5960	**1.0548**	1.3789	1.4493	1.5515
Z=1	0.5963	0.6879	0.7271	**1.0124**	1.3656	1.4138	1.4534

Dsynd HT 1995

	1%	5%	10%	mean	10%	5%	1%
Z=0	0.5248	0.5877	0.6675	**0.9526**	1.3246	1.4227	1.6269
Z=1	0.7150	0.7931	0.8385	**1.0345**	1.4023	1.5120	1.6027

4.5 Age effect.

The effect of Age on efficiency is less clear on the graphical representations. It seems more important on high technology sector, and negative (but we know by descriptive statistics that firms from high technology sector live longer).



5 Conclusion.

We have presented in this work an application of recent nonparametric conditional production frontiers method. This method has the advantage to be more robust to outliers and a one-step estimation procedure. Using Li and Racine (2004) work, we adapted it to categorical and discrete environmental variables. The empirical work is performed on a survey of manufacturing firms of Côte d'Ivoire for two years 1994 and 1995. Except for the age variable, it seems that all the environmental variables have a positive impact on efficiency for low technology sector. These low technology firms seem very sensitive to external factors. On the contrary, the age of the firm has a negative impact for high technology sector and no apparent impact on low technology sector. We note that the results are robust for both years 1994 and 1995.

We are aware of losing information by not analyzing jointly the effects of environmental factors. Only the small sample size prevents us from performing a multivariate analyzis. Moreover, the range of bootstrapped confidence intervals is quite large, which may be explained by the heterogeneity of the database, besides the subsampling between low technology and high technology.

References

[1] Aitchison, J., Aitken, C.G.G (1976): "Multivariate binary discrimination by the kernel method", *Biometrika*, 63, 413-420.

[2] Aragon, Y., Daouia, A., and Thomas-Agnan, C. (2005) : "Nonparametric Frontier Estimation: a Conditional Quantile-based Approach", *Econometric Theory,* 21, 358-389.

[3] Azam, J.-P., and Morisson, C. (1994) : "The Political Feasibility of Adjustment in Côte d'Ivoire and in Morocco", Development Centre Studies, OECD, Paris.

[4] Badin, L., Daraio C. and Simar, L. (2009): "Optimal Bandwidth Selection for Conditional Efficiency Measures: a Data-driven Approach", *Working Paper*

[5] Battese, G.E., and Coelli, T.J. (1995) : "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function and Panel Data", *Journal of Empirical Economics,* 20(2), 325-332.

[6] Charnes, A., Cooper, W.W., Rhodes, E. (1978) : "Measuring the inefficiency of decision making units", European Journal of Operational Research, 2, 429-444.

[7] Cazals, C., Florens, J-P. and Simar, L. (2002) : "Nonparametric Frontier Estimation: a Robust Approach", *Journal of Econometrics,* 106, 1-25.

[8] Chapelle, K. and Plane, P. (2005) : "Productive Efficiency in the Ivorian Manufacturing Sector: an Exploratory Study using a Data Envelopment Analysis Approach", The Developing Economies, 43(4), 450-471.

[9] Cooper, W.W., Seiford L.M., and Tone K. (2000): "Data Envelopment Analysis: A comprehensive text with Models, Applications, References and DEA-Solver Software", Kluwer Academic Publishers, Boston.

[10] Daouia, A. and Simar, L. (2007) : "Nonparametric Efficiency Analysis: a Multivariate Conditional Quantile Approach", *Journal of Econometrics,* 140, 375-400.

[11] Daraio, C. and Simar, L. (2005) : "Introducing Environmental Variables in Nonparametric Frontier Models: a Probabilistic Approach", *Journal of Productivity Analysis,* 24, 93-121.

[12] Daraio, C. and Simar, L. (2007): "Conditional Nonparametric Frontier Models for Convex and Non Convex Technologies: A unifying Approach", *Journal of Productivity Analysis*, 28, 13-32

[13] Daraio, C. and Simar, L (2007) : "Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications", Springer Verlarg.

[14] Debreu, G. (1951): "The Coefficient of Resource Utilization", *Econometrica*, 19, 273-292.

[15] Deprins, D., Simar, L. and Tulkens, H. (1984) : “Measuring Labor Inefficiency in Post Offices”, In *The Performance of Public Entreprises: Concepts and Measurements*, M. Marchand, P. Pestieau and H. Tulkens (eds.), Amsterdam, North Holland, 243-267.

[16] Farrell, M.J. (1957) : "The measurement of productive efficiency", *Journal of the Royal Statistical Society*, A, 120, 253-281.

[17] Florens, J-P., and Simar, L. (2005) : “Parametric Approximations of Nonparametric Frontiers”, *Journal of Econometrics*, 124(1), 91-116.

[18] Fried, H.O., Lovell, C.A.K., Schmidt, S.S. and Yaisawarng, S. (2002): "Accounting for environmental effects and statistical noise in Data Envelopment Analysis", *Journal of Productivity Analysis*, 17, 157-174.

[19] Fried, H.O., Schmidt, S.S. and Yaisawarng, S. (1999): "Incorporating the operating environment into a nonparametric measure of technical efficiency", *Journal of Productivity Analysis*, 12, 249-267.

[20] Gattoufi, S., Oral, M., Reisman, A. (2004): "Data envelopemtn analysis literature: a bibliography update (1951-2001)", *Socio-Economic Planning Sciences* 38, 159-229.

[21] Koopmans, T.C. (1957): "An Analysis of Production as an Efficient Combination of Activities", in Koopmans T.C. (Ed.), *Activity Analysis of Production and Allocation*, Cowles Commision for Research in Economics, Monograph N°.13, Wiley, NewYork.

[22] Li, Q., Racine, J. (2004): "Nonparametric Estimation of Regression functions with Both Categorical and Continuous Data", *Journal of Econometrics*, 119, 99-130

[23] Marchat, J-M. (1998) : “Concurrence and Structure du Marché Intérieur des Produits Manufacturés en Côte d’Ivoire”, RPED Discussion paper, No. 80.

[24] Roudaut, N. (2006) : “Influences of the Business Environment on Manufacturing Firms Technical Efficiencies : The Côte d’Ivoire Case”, *Journal of Productivity Analysis,* 25, 93-109.

[25] Seiford, L.M. (1996): "Data Envelopment Analysis: The Evolutions of the State of the Art (1975-1995)", *Journal of Productivity Analysis*, 7, 99-138.

[26] Shephard, R.W. (1970): "Theory of cost and production function", Princeton University Press, Princeton, NJ.

[27] Simar, L. (2003) : “Detecting Outliers in Frontier Models: A Simple Approach”, *Journal of Productivity Analysis,* 20, 391-424.

[28] Simar, L., and Wilson, P.W. (2007) : “Estimation and Inference in Two-Stage, Semiparametric Models of Production Process”, *Journal of Econometrics*, 25, 31-64.

6 Appendix.

6.1 Descriptive statistics.

Table 1: Data summary of Business Environment variables

Year	1994		1995	
Sector	Low	High	Low	High
Frequencies	118	77	104	70
Production	641.60 (3,329.7)	6,631 (12,729)	839.11 (3,328.3)	8,381 (15,360)
Capital	725.07 (3,191.6)	961,180 (8,291,300)	389.43 (1,257.3)	1,002,700 (8,261,400)
Labour	1.404 (2.518)	16.118 (50.73)	1.376 (2.559)	13.15 (20.22)
Age	10.76 (10.88)	19.23 (14.71)	10.6 (10.7)	19.73 (15.14)
Unions presence	32.2	81.8	31.73	85.7
Managerial firms	40.7	79.2	41.35	77.1
Exporters	20.3	61.0	24.04	64.3
Formal firms	45.8	100	45.19	100

Production, capital and labor are expressed in Millions of CFA Francs, Age in years (Mean (standard error)).
Categorical variables are in percentage.
Source : RPED

6.2 Sensibility to the parameter m.

Tables of numbers of superefficient firms depending on the level of m and t.

6.2.1 Low Technology 1994.

LT 1994					
118	**Values of sensibility parameter t**				**m=a.nx**
nb of points below or on	**0< <0,001**	**0,001< <0,01**	**0,01< <0,1**	**>0,1**	**a**
94	0	0	3	21	0.95
94	0	0	3	21	1
94	0	0	10	14	1.5
94	0	0	20	4	2
94	0	3	21	0	3
94	0	8	16	0	3,5
94	1	14	9	0	4
95	2	16	5	0	5
103	7	8	0	0	6

6.2.2 Low Technology 1995.

LT 1995					
104	Values of sensibility parameter t				$m=a.n_x$
nb of points below or on	0< <0,001	0,001< <0,01	0,01< <0,1	>0,1	a
70	0	0	7	27	0.95
69	0	1	9	25	1
70	0	1	18	15	1.5
72	0	0	26	6	2
72	0	8	24	0	3
72	0	16	16	0	3,5
73	2	21	8	0	4
76	6	21	1	0	5
82	11	11	0	0	6

6.2.3 High Technology 1994.

HT 1994					
77	Values of sensibility parameter t				$m=a.n_x$
nb of points below or on	0< <0,001	0,001< <0,01	0,01< <0,1	>0,1	a
57	0	0	4	16	0.95
57	0	0	3	17	1
58	0	1	8	10	1.5
59	0	0	14	4	2
59	0	2	16	0	3
59	0	7	11	0	3,5
59	0	13	5	0	4
60	2	15	0	0	5
64	2	11	0	0	6

6.2.4 High Technology 1995.

HT 1995					
70	Values of sensibility parameter t				$m=a.n_x$
nb of points below or on	0< <0,001	0,001< <0,01	0,01< <0,1	>0,1	a
46	0	1	9	14	0.95
46	0	1	10	13	1
48	0	1	14	7	1.5
49	0	0	17	4	2
49	0	9	12	0	3
49	1	11	9	0	3,5
49	1	13	7	0	4
49	9	12	0	0	5
57	6	7	0	0	6