

Estimation de l'efficacité des entreprises en Côte d'Ivoire: une approche non paramétrique robuste

N. Roudaut et A. Vanhems

JMS 2009

Plan

- Introduction aux notions de frontière de production et efficacité technique
- Une estimation non paramétrique robuste de l'efficacité
- La base de données
- Les résultats

Introduction-1:

Frontières de production

- *Processus de production*: génère des vecteurs **X** (inputs) et **Y** (outputs) tq **$(X, Y) \in \Psi \subset \mathbb{R}_+^p \times \mathbb{R}_+^q$**

- *Ensemble de production*:

$$\Psi = \{ (x, y), \text{ x permet de produire y} \}$$

- *Frontière de Production (output-oriented)*:

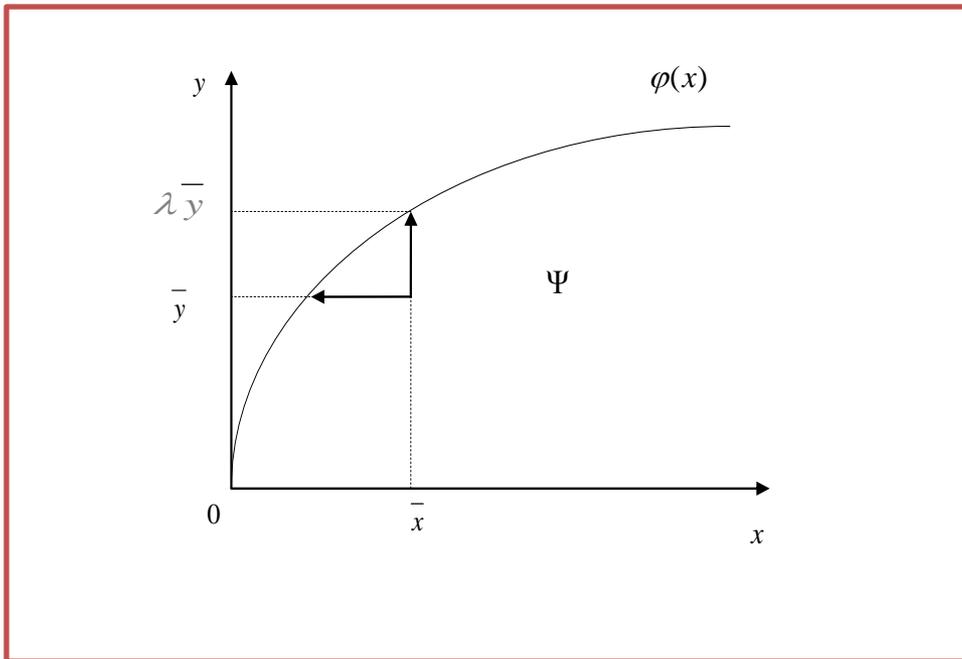
$$\varphi(x) = \{ y \mid (x, y) \in \Psi, (x, \lambda y) \notin \Psi, \forall \lambda > 1 \}$$

- *Mesure d'efficacité de Farrell-Debreu*:

$$\lambda(x) = \sup \{ \lambda \mid (x, \lambda y) \in \Psi \}$$

Introduction-2

- Exemple - Cas univarié



$\lambda = 1$: entreprise **efficace**

$\lambda > 1$: entreprise **inefficace**

Introduction-3: Traduction probabiliste

- Sous hypothèse de **free-disposabilité** de Ψ ,

$$\varphi(\mathbf{x}) = \sup\{\mathbf{y} \mid F_{Y|X}(\mathbf{y} \mid \mathbf{x}) < 1\}$$

avec $F_{Y|X}(\mathbf{y} \mid \mathbf{x}) = \text{Prob}(Y \leq \mathbf{y} \mid X \leq \mathbf{x})$.

$$\lambda(\mathbf{x}) = \sup\{\lambda \mid S_{Y|X}(\lambda \mathbf{y} \mid \mathbf{x}) > 0\}$$

avec $S_{Y|X}(\mathbf{y} \mid \mathbf{x}) = \text{Prob}(Y \geq \mathbf{y} \mid X \leq \mathbf{x})$.

Introduction-4: état de la littérature

- Méthodes paramétriques/ **non paramétriques**
(*stochastiques*) (déterministes)

Méthodes DEA, FDH, sensibles aux valeurs extrêmes

Farrell 1957, Charnes, Cooper and Rhodes 1978, Deprins, Simar Tulkens 1984

Méthodes **m-frontières**, α -quantile

Cazals, Florens, Simar 2002

Aragon, Daouia, Thomas-Agnan 2005

- Introduction de variables exogènes

Méthodes d'estimation en plusieurs étapes

Fried, Schmidt, Yaisawarn 1999, 2002, Simar, Wilson 2003

Méthodes en une étape

Banker, Morey 1984, Daraio, Simar 2005, 2007

Méthode d'estimation-1

- M-frontière conditionnelle: soit (X, Y, Z) un vecteur inputs, outputs, variables exogènes

X=travail, capital, Y=production, Z=âge de l'entreprise, présence d'un syndicat...

- Soit $m \geq 1$ et (Y_1, \dots, Y_m) échantillon *iid* de $Y | X \leq x$

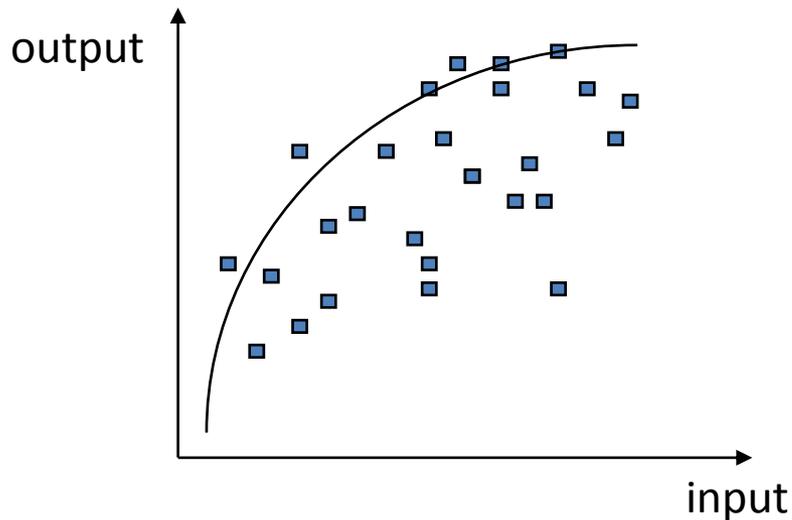
$$\varphi_m = E(\max(Y_1, \dots, Y_m) | X \leq x, Z = z)$$

$$\varphi_m(x|z) = \int_0^{+\infty} [1 - F_{Y|X,Z}(y|x, z)^m] dy$$

$$\lambda_m(x, y|z) = \int_0^{+\infty} [1 - F_{Y|X,Z}(uy|x, z)^m] du$$

Méthode d'estimation-2

- Interprétation:



- Le maximum est maintenant calculé sur un sous échantillon de taille m
- Approximation plus robuste aux valeurs extrêmes
- Qui tient compte de l'environnement extérieur

- $\lambda_m = 1$: entreprise **efficace**
- $\lambda_m > 1$: entreprise **inefficace**
- $\lambda_m < 1$: entreprise **super-efficace**

Méthode d'estimation-3

Observations $(X_i, Y_i, Z_i), i=1, \dots, n$

- Estimation non paramétrique:

$$\hat{\lambda}_m(x, y|z) = \int_0^{+\infty} [1 - \hat{F}_{Y|X,Z}(uy|x, z)^m] du$$

$$\hat{F}(x, y|z) = \frac{\sum_{i=1}^n 1(X_i \leq x, Y_i \leq y) K(z, Z_i, h_n)}{\sum_{i=1}^n 1(X_i \leq x) K(z, Z_i, h_n)}$$

- Approximation par méthode de Monte-Carlo

Méthode d'estimation-4

- *Choix des paramètres de lissage:*
 - Noyau **K** adapté au type de **Z**: Catégoriel ordonné ou non ordonné, continu (Aitchison, Aitken 1976, Li, Racine 1994)
 - Fenêtre **h_n**: solution d'un critère de minimisation par CV sur la densité marginale de **Z** (Daraio Simar 2005)
 - Paramètre **m**: analyse de sensibilité et calcul des outliers en fonction de m (Simar 2003)
 - **m = a . n_x**
 - Étude du % de points tels que **$\lambda < 1 - t$**

La base de données-1

- Enquête RPED 1994, 1995 sur un échantillon d'entreprises du secteur manufacturier (Côte d'Ivoire)
 - Analyse par année et par technologie: **HT** et **LT** (Roudaut 2006)
 - **Y**=Production; **X**= Capital, travail
 - **Z₁** = Formel/informel
 - **Z₂** = Présence de syndicat
 - **Z₃** = Exportateur
 - **Z₄** = Structure managériale
 - **Z₅** = Age
- Variables dichotomiques
- Variable catégorielle ordonnée

La base de données-2

Table 8: Data summary of Business Environment variables

Year	1994		1995	
	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)).

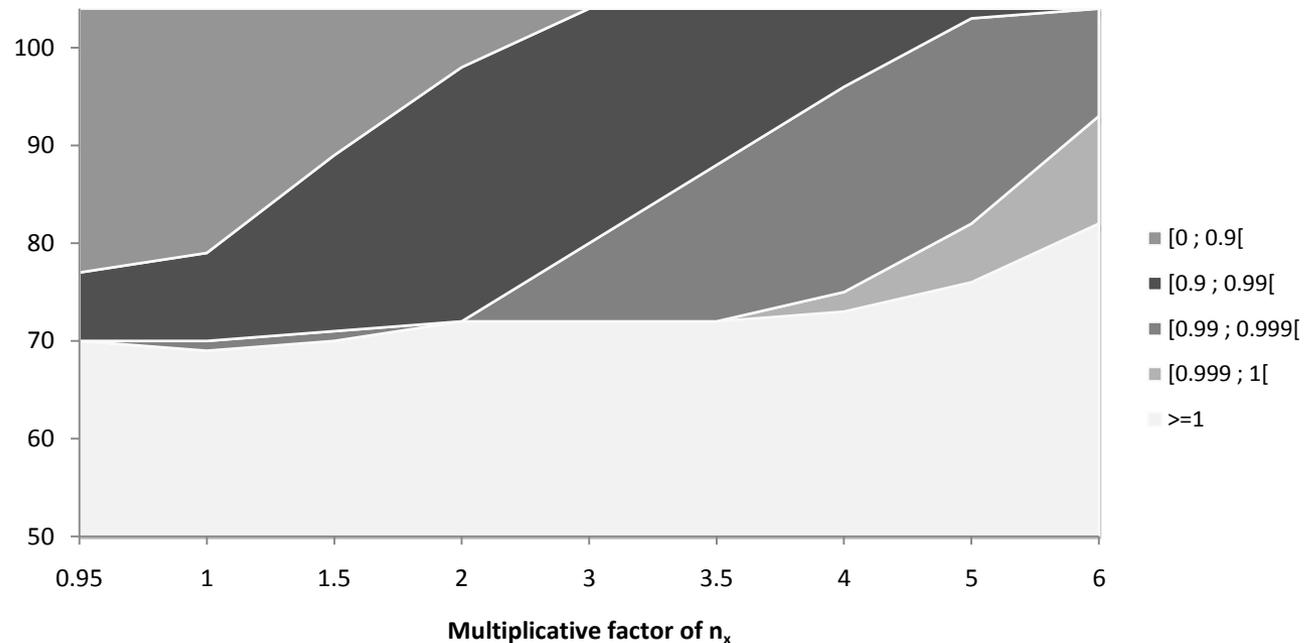
Categorical variables are in percentage.

Source : RPED

Résultats-1

- *Analyse de sensibilité*: $m = 2$, n_x , $t=0,1$ et **5%** d'efficacités inférieures à 0,9.

Exemple: Cas LT 1994; $n = 104$



Résultats-2

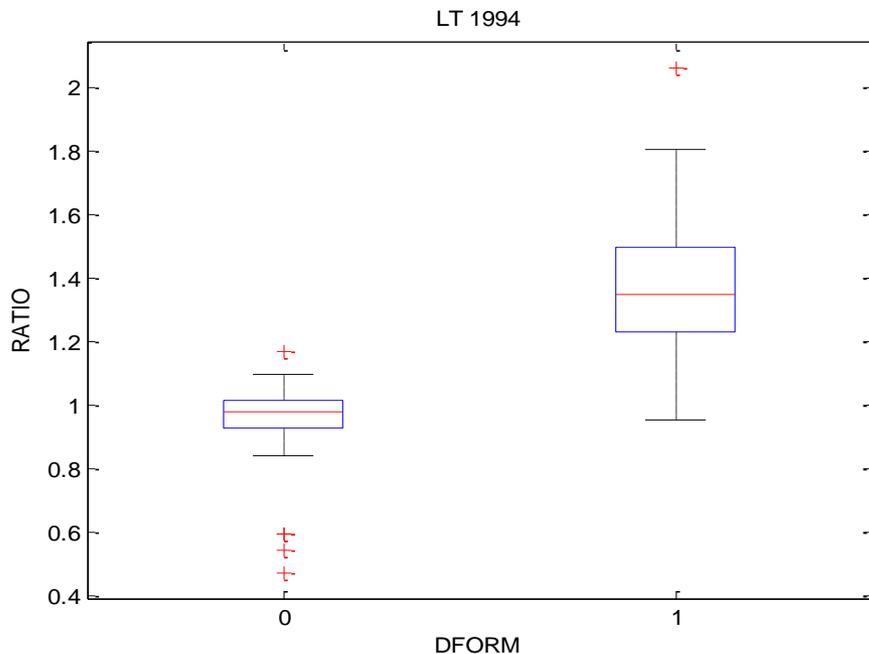
- Impact de **Z** sur l'efficacité : analyse du ratio **Q_Z** comme fonction de **Z**.

$$Q_Z = \frac{\hat{\lambda}_m(x, y|z)}{\hat{\lambda}_m(x, y)}$$

- Analyse indépendante de l'effet de chaque **Z**.
- Construction d'intervalles de confiance et tests par Bootstrap. (Florens Simar 2005)

Résultats-3

- Impact **formel/informel**:



Dform LT 1994

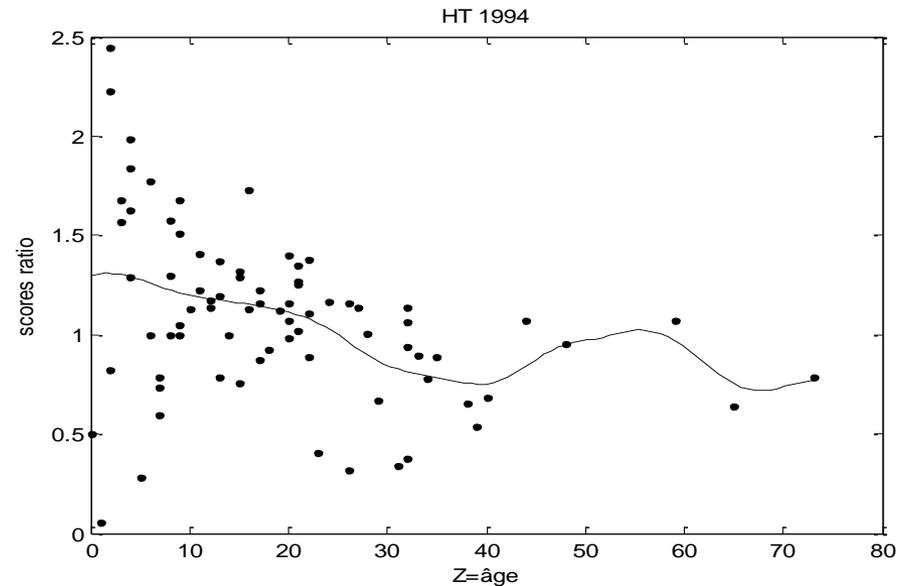
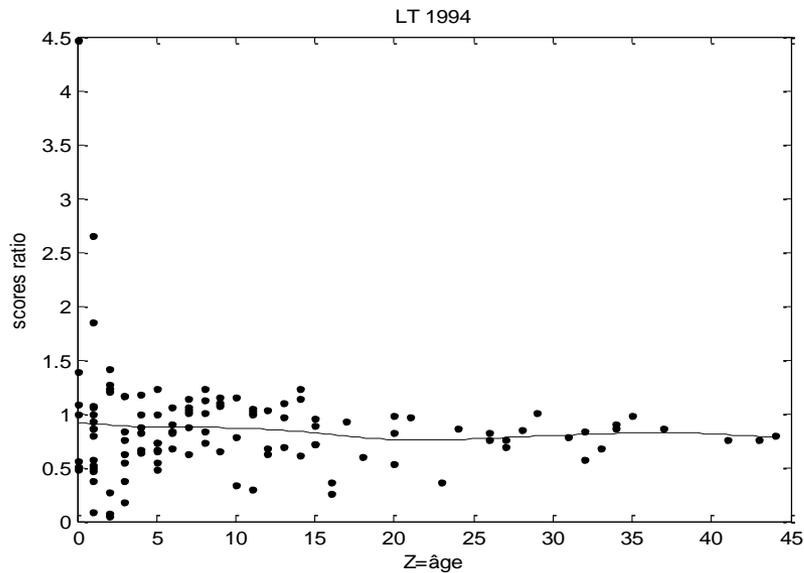
	1%	5%	10%	mean	10%	5%	1%
Z=0	0.4298	0.5740	0.6189	0.9690	1.1937	1.2579	1.3975
Z=1	0.9338	1.1503	1.2096	1.3771	1.7703	1.8452	1.8996

Remarque : Pas d'entreprises informelles dans le secteur HT.

- Avec un risque de 10%, le ratio Q_z est significativement plus élevé pour les entreprises formelles: effet positif de la variable sur l'efficacité.**
- Effet robuste sur les deux années.**

Résultats-4

- Impact **âge**:



- Pas d'effet de l'âge dans le secteur LT
- Effet décroissant dans le secteur HT: l'âge a un impact négatif sur l'efficacité
- Effet robuste sur les deux années

Conclusion

- Pour les variables *formel*, *exportateur*, *manager*, *syndicat*: effet positif sur l'efficacité dans le secteur LT, mais pas d'effet sur le secteur HT.
- Pour la variable *âge*: effet négatif dans le secteur HT, pas d'effet sur le secteur LT
- Effet de **Z** décroissant avec **m**