Annexe A
Mesures concernant le marché du travail mises en place sous le programme PASCO.

Objectifs	Dispositions
réduction du coût unitaire du travail	- suppression de la contribution employeur sur la masse salariale des nationaux et maintien de la taxe d'apprentissage et de la taxe pour la formation professionnelle continue - réduction de la contribution employeur su la masse salariale des expatriés
augmentation de la flexibilité du marché du travail	- suppression du monopole de l'OMOCI (Office de la Main d'œuvre de Côte d'Ivoire) - suppression de restrictions sur l'embauche des employés occasionnels ou temporaires et les licenciements collectifs - suppression de l'obligation d'obtenir les autorisations préalables aux heures supplémentaires - révision du code du travail dans le sens de l'assouplissement

Annexe B
Stratification de la base de sondage (1995)

population totale en Côte d'Ivoire	base de sondage	taille de l'échantillon	signification
127	94	52	formelle - Abidjan, Bouaké, San-Pedro - agro-ind.
73	50	37	formelle - Abidjan, Bouaké, San-Pedro - bois
65	64	49	formelle - Abidjan, Bouaké, San-Pedro - métaux
21	21	17	formelle - Abidjan, Bouaké, San-Pedro - textile
14	14	6	semi-formelle - Abidjan, Bouaké, San-Pedro - agro-ind.
14	14	7	semi-formelle - Abidjan, Bouaké, San-Pedro - bois
14	14	5	semi-formelle - Abidjan, Bouaké, San-Pedro - métaux
14	14	8	semi-formelle - Abidjan, Bouaké, San-Pedro - textile
3187	92	16	informelle - Abidjan - bois
1159	16	3	informelle - Abidjan - métaux
5311	232	34	informelle - Abidjan - textile
9999	625	234	

source: Marchat (1997)

Annexe C
Définitions et caractéristiques des variables individuelles

nom	définition	moyenne	éc.-ty.	N
salaire	salaire mensuel net, primes et indemnités incluses, avant déduction des remboursements de prêts	122347	207301	1568
éducation	âge à la sortie de l'école moins 6 ans	12.76	4.91	1218
expérience	âge actuel - âge à la sortie de l'école - la durée d'une éventuelle période de chômage	16.45	8.52	1211
ancienneté	nombre d'années passées dans l'emploi actuel	7.69	6.78	1613
primaire	= 1 si le niveau scolaire le plus élevé acquis par l'employé est le primaire	0.32	0.46	1613
secondaire classique	= 1 si le niveau scolaire le plus élevé acquis est le niveau secondaire classique	0.10	0.31	1613
secondaire technique	= 1 si le niveau scolaire le plus élevé acquis est le niveau secondaire technique	0.16	0.37	1613
supérieur technique	= 1 si le niveau scolaire le plus élevé acquis est le niveau supérieur technique	0.04	0.20	1613
universités ivoiriennes	= 1 si le niveau scolaire le plus élevé acquis est le niveau universitaire en Côte d'Ivoire	0.01	0.13	1613
universités étrangères	= 1 si le niveau scolaire le plus élevé acquis est le niveau universitaire étranger	0.02	0.14	1613
apprenti	= 1 si l'employé a effectué un apprentissage dans l'entreprise ou dans le secteur	0.27	0.44	1610
africain de l'ouest	= 1 si l'employé est originaire d'un pays de l'Afrique de l'ouest, hors Côte d'Ivoire	0.19	0.39	1613
non africain	= 1 si l'employé n'est pas originaire d'Afrique	0.009	0.09	1613
sexe	= 1 si l'employé est un homme	0.90	0.29	1613
marié	= 1 si l'employé est marié	0.31	0.46	1613
hebdo	nombre d'heures de travail hebdomadaire	44.6	4.55	1534
permanent	= 1 si l'employé est un employé permanent	0.97	0.15	1605
licencié	= 1 si l'employé a déjà été licencié	0.24	0.42	1572
formation	= 1 si l'employé suit une formation dans le cadre de son travail	0.03	0.18	1592

Annexe D
Définitions et caractéristiques des variables d'entreprises

nom	définition	moyenne	éc.-ty.	N
Abidjan	= 1 si l'entreprise est située à Abidjan	0.87	0.32	464
taille	nombre total d'employés	171	707	441
âge	nombre d'années d'existence de l'entreprise	15.95	13.67	424
public	= 1 si l'entreprise est une société parapublique	0.015	0.12	464
étrangère	= 1 si le capital de l'entreprise est à plus de 50% étranger	0.57	0.49	464
ivoirienne	= 1 si l'entreprise est de propriété ivoirienne	0.40	0.49	464
multinationale	= 1 l'entreprise est filiale d'une multinationale	0.03	0.18	464
formel	= 1 si l'entreprise est de type formel	0.75	0.43	464
profit	profits bruts, avant dépréciation et impôts	550730676	0.22E+8	384
concurrence	nombre de concurrents de l'entreprise sur son 1er produit	3.69	1.40	426
syndicat	= 1 s'il existe un syndicat dans l'entreprise	0.43	0.49	387
permanent	nombre d'employés permanents	142	567	417
capital	valeur comptable des immobilisations remplacée par la valeur de remplacement si elle n'est pas disponible (en log)	17.64	4.23	230
valeur ajoutée	valeur de la production moins les coûts intermédiaires et la variation du stock (en log)	17.34	0.04	409
capital humain moyen	moyenne du niveau d'éducation des différentes catégories d'employés pondérée par l'effectif de chaque catégorie	2.00	1.00	348
turnover de l'emploi	rapport de la somme des employés embauchés, licenciés ou ayant démissionné dans la dernière année aux effectifs totaux	0.19	0.39	440
supervision	rapport du nombre de superviseurs à l'effectif attaché à la production	0.15	0.14	247
salaire relatif empl. de prod.	salaire moyen versé aux employés de production rapporté au salaire moyen du secteur versé à cette catégorie d'employés	0.99	0.88	329
salaire relatif empl. de superv.	salaire moyen versé aux employés de supervision rapporté au salaire moyen du secteur versé à cette catégorie d'employés	1.00	0.95	227
salaire moyen	salaire moyen versé dans l'entreprise	78624	25868	464
% prod. export.	pourcentage de la production qui est exportée	30	38	315
technologie	= 1 si l'entreprise détient des licences étrangères ou utilise une assistance technique	0.14	0.34	419

Annexe E
Estimateur à variables instrumentales (Hausman et Taylor, 1981, Amemiya et MaCurdy, 1986, Breush, Mizon et Schmidt, 1989)

Modèle introduisant un effet spécifique individuel

Le modèle à estimer s'écrit

$$\begin{aligned} y_{it} &= X_{it}\beta + Z_i\gamma + \epsilon_{it} \\ \epsilon_{it} &= \alpha_i + \eta_{it} \\ i &= 1,...,N; t = 1,...,T_i; \sum_{i=1}^{N} T_i = T \end{aligned}$$

où certaines des variables figurant dans X_{it} (matrices de variables variantes dans le temps, de dimension $TN \times k$) et dans Z_i (matrice de variables invariantes dans le temps, de dimension $TN \times g$) sont corrélées avec les effets spécifiques individuels α_i

$$E\left(\epsilon_{it}/X_{it}, Z_i\right) = E\left(\alpha_i/X_{it}, Z_i\right) \neq 0$$

Supposons que k_1 variables de X_{it} (notées X_{1it}) et g_1 variables de Z_i (notées Z_{1i}) ne sont pas corrélées avec les effets spécifiques. Nous avons

$$\begin{array}{ll} p\lim_{N\to\infty} \frac{1}{N} X'_{1it}\alpha_i = 0 & p\lim_{N\to\infty} \frac{1}{N} X'_{2it}\alpha_i = h_X \neq 0 \\ p\lim_{N\to\infty} \frac{1}{N} Z'_{1i}\alpha_i = 0 & p\lim_{N\to\infty} \frac{1}{N} Z'_{2i}\alpha_i = h_Z \neq 0 \end{array}$$

où
X_{1it} est une matrice de variables variantes dans le temps exogènes, de dimension $TN \times k_1$
X_{2it} est une matrice de variables variantes dans le temps endogènes, de dimension $TN \times k_2$
Z_{1i} est une matrice de variables invariantes dans le temps exogènes, de dimension $TN \times g_1$
Z_{2i} est une matrice de variables invariantes dans le temps endogènes, de dimension $TN \times g_2$

En raison de l'endogénéité des effets spécifiques individuels, l'estimateur MCQG de ce modèle n'est pas convergent. Il faut alors recourir à une procédure d'estimation à variables instrumentales. La difficulté de cette procédure consiste à trouver des instruments indépendants de l'effet α en nombre suffisants pour identifier les paramètres du modèle.

Si k_1 est supérieur à g_2, Hausman et Taylor (1981) proposent d'utiliser les moyennes individuelles des variables X_1 comme instruments pour les variables Z_2. En fait, cela revient à estimer le modèle en utilisant comme instruments $A_{HT} = [WX_1, WX_2, X_1, Z_1]$ qui est équivalent à $[WX_1, WX_2, BX_1, Z_1]$

(Breusch, Mizon et Schmidt, 1989), où WX_1 et WX_2 désignent les écarts à la moyenne individuelle des variable X_1 et X_2, BX_1 désignent les moyennes individuelles de X_1 $(WX_{1it} = X_{1it} - X_{1i.}, BX_1 = X_{1i.})$.

L'estimateur résultant de cette procédure est convergent mais il n'est pas efficace car il ignore la structure des erreurs liée à la présence de l'effet spécifique. Pour améliorer la précision des estimateurs, les auteurs proposent d'utiliser de nouvelles estimations des coefficients du modèle pour estimer la variance de l'effet individuelle (σ_α^2) et de la composante non spécifique de la perturbation (σ_η^2) et d'appliquer la méthode des variables instrumentales au modèle transformé

$$\begin{aligned} y_{it} - \left(1 - \sqrt{\hat{\theta}_i}\right) y_{i.} &= \left[X_{it} - \left(1 - \sqrt{\hat{\theta}_i}\right) X_{i.}\right] \beta \\ &\quad + \sqrt{\hat{\theta}_i} Z_i \gamma + \sqrt{\hat{\theta}_i} \alpha_i + \left[\eta_{it} - \left(1 - \sqrt{\hat{\theta}_i}\right) \eta_{i.}\right] \end{aligned}$$

avec

$$\hat{\theta}_i = \frac{\hat{\sigma}_\eta^2}{\hat{\sigma}_\eta^2 + T_i \hat{\sigma}_\alpha^2}$$

Les estimateurs de β et γ dans cette équation sont convergents et efficaces.

Nous pouvons résumer les différentes étapes nécessaires à l'obtention de l'estimateur de Hausman et Taylor (1981) dans le cas d'un échantillon non cylindré de la façon suivante

1. L'estimation intra-individuelle est appliquée au modèle

$$y_{it} = X_{it}\beta + Z_i\gamma + \alpha_i + \eta_{it}$$

l'estimateur $\widehat{\beta}_w$ est convergent

2. On calcule

$$\widehat{d_i} = y_{i.} - X_{i.}\widehat{\beta}_w = Z_i c + \alpha_i + \eta_{i.}$$

l'estimateur des moindres carrés ordinaires de c est non convergent $(E(\alpha_i / Z_{2i}) \neq 0)$. On applique alors la méthode des doubles moindres carrés, en instrumentant Z_2 avec les instruments $A = [X_1, Z_1]$

$$\widehat{d_i} = y_{i.} - X_{i.}\widehat{\beta}_w = Z_{1i}c_1 + \widehat{Z}_{2i}c_2 + \alpha_i + \eta_{i.}$$

L'estimateur de c obtenu, $\widehat{c}_{DMC} = \begin{pmatrix} \widehat{c}_1 \\ \widehat{c}_2 \end{pmatrix}$, est convergent.

3. En utilisant les paramètres estimés $\widehat{\beta}_w$ et $\widehat{c}_{DMC}$, on forme les résidus

$$\widehat{u}_w = (y_{it} - y_{i.}) - (X_{it} - X_{i.})\widehat{\beta}_w$$

et

$$\widehat{u}_{DMC} = \hat{d}_i - Z_i \widehat{c}_{DMC}$$

Ces deux vecteurs de résidus sont utilisés pour le calcul des composantes de la variance.

Un estimateur convergent de σ_η^2 est

$$s_\eta^2 = \frac{\widehat{u}'_w \widehat{u}_w}{\frac{N}{q_n} - N - k_w}$$

avec k_w le nombre de paramètres de la régression intra-individuelle $(k_1 + k_2)$ et $q_n = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T_i}$

Un estimateur convergent de σ_α^2 est

$$s_\alpha^2 = s^2 - s_\eta^2 q_n$$

avec

$$s^2 = \frac{\widehat{u}'_{DMC} \widehat{u}_{DMC}}{\frac{N}{q_n} - g}$$

avec g le nombre de paramètres à estimer $(g_1 + g_2)$.

4. On obtient finalement une estimation de θ_i

$$\widehat{\theta}_i = \frac{s_\eta^2}{s_\eta^2 + T_i s_\alpha^2}$$

5. On applique alors la méthode des variables instrumentales au modèle transformé

$$y_{it} - \left(1 - \sqrt{\widehat{\theta}_i}\right) y_{i.} = \left[X_{it} - \left(1 - \sqrt{\widehat{\theta}_i}\right) X_{i.}\right] \beta + \\ + \sqrt{\widehat{\theta}_i} Z_i \gamma + \left[\epsilon_{it} - \left(1 - \sqrt{\widehat{\theta}_i}\right) \epsilon_{i.}\right]$$

avec les instruments $A_{HT} = [WX_1, WX_2, BX_1, Z_1]$.

L'estimateur de Hausman et Taylor requiert des conditions minimales d'exogénéité sur les variables $(X_1, Z_1, WX_1$ et WX_2 ne sont pas corrélées à l'effet individuel), et en conséquence, il peut ne pas être le plus efficace si les conditions d'exogénéité sont rendues plus restrictives.

Amemiya et MaCurdy (1986) suggèrent un estimateur plus efficace en supposant que les réalisations de X_1 ne sont pas corrélées avec α à chaque période $(E(X_{1it}\alpha_i) = 0, \forall i, t)$. Par conséquent, ils n'utilisent pas uniquement X_{1it} comme instrument pour l'individu i à la période t, mais toute la série $X_{1i1}, X_{1i2}, \ldots, X_{1iT}$.

Ces auteurs définissent la matrice

$$X_1^* = \begin{bmatrix} X_{11} & X_{12} & \ldots & & X_{1T} \\ X_{11} & X_{12} & & & X_{1T} \\ \ldots & \ldots & & & \ldots \\ X_{21} & X_{22} & & & \\ X_{21} & \ldots & & & \\ \ldots & & & & \\ X_{N1} & X_{N2} & & & \\ X_{N1} & X_{N2} & & & \ldots \\ \ldots & \ldots & \ldots & \ldots & X_{NT} \end{bmatrix}$$

qui possède les propriétés

$$WX_1^* = 0, BX_1^* = X_1^*$$

Leur matrice d'instruments est

$$A_{AM} = [WX_1, WX_2, X_1^*, Z_1]$$

Un estimateur équivalent est obtenu en utilisant

$$A_{AM} = [WX_1, WX_2, BX_1, (WX_1)^*, Z_1]$$

Les auteurs suggèrent que leur estimateur est au moins aussi efficace que celui de Hausman et Taylor si les effets ne sont pas corrélés avec les régresseurs à chaque période.

Breush, Mizon et Schmidt (1989) décrivent un estimateur potentiellement plus performant, fondé sur une hypothèse d'exogénéité plus stricte encore. Dans le cadre d'Hausman et Taylor, WX_2 est un instrument légitime car la composante spécifique individuelle de X_2, corrélée avec α, disparaît dans la différenciation.

Breush, Mizon et Schmidt, en étendant l'estimateur de Amemiya et MaCurdy, supposent que WX_2 n'est pas corrélée avec l'effet individuel à chaque période. En d'autre termes, on n'utilise pas uniquement WX_{2it} comme instrument pour l'individu i à la période t, mais également $WX_{2i1}, WX_{2i2}, ...WX_{2iT}$. Leur estimateur est donc basé sur la matrice d'instruments

$$A_{BMS} = [WX_1, WX_2, (WX_1)^*, (WX_2)^*, Z_1]$$

Breush, Mizon et Schmidt supposent que le composantes de X_2 sont corrélées avec l'effet individuel via un terme indépendant du temps (l'endogénéité de X_2 vient uniquement de la corrélation entre l'effet individuel et une composante de X_2 qui est indépendante du temps). Si cette hypothèse n'est pas vérifiée, alors $(WX_2)^*$ est corrélée avec α et la matrice d'instruments de Breush, Mizon et Schmidt n'est pas légitime.

Annexe F
Le biais de sélection (Guillotin et Sevestre, 1994)

Le mode de constitution de l'échantillon en panel non cylindré, composé des observations des employés pour les deux années 1995 et 1996 n'est pas exogène: la probabilité pour qu'un individu appartienne à l'échantillon n'est pas indépendante des perturbations, et en particulier des caractéristiques individuelles non observées prises en compte dans les effets spécifiques. Il faut donc prendre en compte le processus de sélection lors de l'estimation du modèle.

Une première solution consiste à spécifier un modèle explicatif de ce processus et à appliquer la méthode en deux étapes de Heckman (1979). Cependant, cette méthode suppose que la sélection des individus appartenant à l'échantillon réponde à une seule et même logique que l'on puisse formaliser, alors que les processus qui conduisent à l'entrée ou la sortie dans l'échantillon sont vraisemblablement différents. De plus, nous ne disposons d'aucune information sur les employés pour l'année où ils sont absents de l'échantillon.

Nijman et Verbeek en 1992 ont proposé d'approcher la correction de la méthode d'Heckman en ajoutant au modèle des variables indicatrices liées au statut de la présence des individus dans l'échantillon. Cette technique consiste à estimer le modèle en incluant comme régresseur supplémentaire une variable indicatrice valant 1 si l'individu est présent sur toute la période, et/ou une autre variable indicatrice valant un si l'individu était présent à la période précédente ou encore le nombre de périodes de présence. La significativité des coefficients associés à ces variables est alors un indice de l'existence d'un biais de sélection et leur inclusion est censée permettre une correction de ce biais qui, si elle est imparfaite, est néanmoins facile à mettre en oeuvre, même en présence de multiples processus de sélection.

La solution consistant à inclure comme régresseur supplémentaire une variable indicatrice valant 1 si l'individu est présent sur toute la période, bien que ne constituant pas une solution rigoureuse aux biais de sélection éventuels, est retenue dans cette étude.

Annexe G
L'équation de sélection

Les estimations menées dans la deuxième section du chapitre deux ont pour but d'étudier la relation entre les différentes catégories d'entreprises et les salaires, en suivant la démarche en deux étapes proposée notamment par Lee (1978) et Heckman (1979).

Deux possibilités sont offertes pour chaque employé: il peut travailler soit dans une grande, soit dans une petite entreprise, cela dépend de son propre choix et de la décision de l'employeur. Chaque travailleur fait ainsi face à deux salaires, dépendant du type d'entreprise concernée.

Soient w_{ig} et w_{ip} les salaires de l'individu i dans une grande et dans une petite entreprise respectivement, soit ρ_i son salaire de réservation qui résume ses préférences spécifiques. L'individu i rejoint une grande entreprise si

$$\ln w_{ig} - \ln w_{ip} > \rho_i \qquad \text{(G.1)}$$

c'est-à-dire qu'il travaille dans une grande entreprise si le différentiel de salaire excède son salaire de réservation. Le salaire de réservation est une fonction de ses caractéristiques personnelles ainsi que des caractéristiques de l'entreprise dans laquelle il sera employé. Ainsi, ρ_i s'écrit

$$\rho_i = \alpha_0 + \alpha_1 X_i + \alpha_2 Z_j + \epsilon_i \qquad \text{(G.2)}$$

où X est un vecteur de caractéristiques individuelles, Z est un vecteur de caractéristiques de l'entreprise et ϵ est un terme d'erreur tel que $\epsilon_i \rightarrow N(0, \sigma_\epsilon^2)$.

L'employé travaillera donc dans une grande entreprise si

$$\ln w_{ig} - \ln w_{ip} \succ \alpha_0 + \alpha_1 X_i + \alpha_2 Z_j + \epsilon_i \qquad \text{(G.3)}$$

Ce critère de choix peut s'écrire sous la forme d'un modèle probit.

Soit la variable I_i^* définit par

$$I_i^* = \delta_0 + \delta_1 (\ln w_{ig} - \ln w_{ip}) + \delta_2 X_i + \delta_3 Z_j + \varepsilon_i \qquad \text{(G.4)}$$

si $I_i^* > 0$ l'employé i travaille dans une grande entreprise, sinon, il travaille dans une petite entreprise.

Si une variable muette pour la taille de l'entreprise est introduite dans l'équation de gains, il est implicitement supposé que les effets des caractéristiques individuelles sur les salaires sont les mêmes quelle que soit la taille de l'entreprise. Nous spécifions ici une équation de salaire pour chacun des deux types d'entreprise, supposant une interaction dans l'équation de gains. Nous avons donc

$$\begin{aligned} \ln w_{ig} &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \varepsilon_{ig} && \text{(G.5)} \\ \ln w_{ip} &= \theta_{0p} + \theta_{1p} X_{ip} + \theta_{2p} Z_{jp} + \varepsilon_{ip} && \text{(G.6)} \end{aligned}$$

où ε_{ig} et ε_{ip} sont deux termes aléatoires supposés suivre une distribution normale $N(0,\sigma_g^2)$ et $N(0,\sigma_p^2)$ respectivement.

Le modèle à estimer s'écrit

$$\begin{aligned} \ln w_{ig} &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \varepsilon_{ig} && \text{(G.7)} \\ \ln w_{ip} &= \theta_{0p} + \theta_{1p} X_{ip} + \theta_{2p} Z_{jp} + \varepsilon_{ip} && \text{(G.8)} \\ I_i^* &= \delta_0 + \delta_1 \left(\ln w_{ig} - \ln w_{ip}\right) + \delta_2 X_i + \delta_3 Z_j + \varepsilon_i && \text{(G.9)} \end{aligned}$$

Les variables observées sont les variables explicatives X et Z et le statut d'appartenance à une grande entreprise I_i. Le salaire observé dépend de l'appartenance à un type d'entreprise; nous observons w_{ig} si $I_i = 1$ et w_{ip} si $I_i = 0$, mais jamais les deux.

Dans ce modèle les salaires et ne peuvent pas être estimés de façon convergente par les moindres carrés ordinaires car

$$\begin{aligned} E\left(\ln w_{ig} \mid I=1\right) &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + E\left(\varepsilon_{ig} \mid I=1\right) && \text{(G.10)} \\ E\left(\ln w_{ip} \mid I=0\right) &= \theta_{0p} + \theta_{1p} X_{ip} + \theta_{2p} Z_{jp} + E\left(\varepsilon_{ip} \mid I=0\right) && \text{(G.11)} \end{aligned}$$

et

$$\begin{aligned} E\left(\varepsilon_g \mid I_i=1\right) &\neq 0 && \text{(G.12)} \\ E\left(\varepsilon_p \mid I_i=0\right) &\neq 0 && \text{(G.13)} \end{aligned}$$

Heckman (1979) et Lee (1978) proposent une autre méthode d'estimation consistant en une procédure des moindres carrés modifiés. L'idée de cette méthode est de trouer une expressions des moyennes $E\left(\varepsilon_g \mid I_i=1\right)$ et $E\left(\varepsilon_p \mid I_i=0\right)$ et d'ajuster les termes d'erreurs de façon à ce qu'ils aient des moyennes nulles.

En substituant $\ln w_{ig}$ et $\ln w_{ip}$ par leur expression dans (G.9), nous obtenons un modèle probit

$$I_i^* = \gamma_0 + \gamma_1 X_i' + \gamma_2 Z_j' + \varepsilon_i^* \qquad \text{(G.14)}$$

et

$$\begin{aligned} E\left(\ln w_{ig} \mid I=1\right) &= E\left(\ln w_{ig} \mid X_{ig}, I=1\right) && \text{(G.15)} \\ &= E\left(\ln w_{ig} \mid X_{ig}, \gamma W + \varepsilon_i^* > 0\right) && \text{(G.16)} \\ &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + E\left(\varepsilon_{ig} \mid \varepsilon_i^* > -\gamma W\right) && \text{(G.17)} \\ &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \rho \sigma_g \sigma_{\varepsilon_i^*} \frac{\phi\left(-\gamma W\right)}{1 - \Phi\left(-\gamma W\right)} && \text{(G.18)} \\ &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \rho \sigma_g \sigma_{\varepsilon_i^*} \frac{\phi\left(-\gamma W\right)}{\Phi\left(-\gamma W\right)} && \text{(G.19)} \end{aligned}$$

où $W = [X', Z']$, Φ est la fonction de répartition d'une variable aléatoire normale et ϕ est sa fonction de densité.

En normalisant $\sigma^2_{\varepsilon *} = 1$

$$\begin{aligned} E(\ln w_{ig} \mid I = 1) &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \rho \sigma_g \lambda && \text{(G.20)} \\ &= \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \beta_\lambda \lambda && \text{(G.21)} \end{aligned}$$

La première étape consiste à estimer le modèle probit et à obtenir les paramètres $\widehat{\gamma}_k$. Pour chaque observation l'inverse du ration de Mill est calculé (λ). La régression de $\ln w$ sur X et Z permet d'estimer les coefficients θ_k et β_λ.

Conditionnellement à la taille de l'entreprise, l'équation de salaires dans une grande entreprise est

$$\ln w_{ig} = \theta_{0g} + \theta_{1g} X_{ig} + \theta_{2g} Z_{jg} + \beta_\lambda \left(\frac{\phi(\widehat{\gamma} W)}{\Phi(\widehat{\gamma} W)} \right) + \eta_g \tag{G.22}$$

où $E\left(\eta_g \mid I_i = 1\right) = 0$.

Conditionnellement à la taille, l'équation de salaires dans une petite entreprise est

$$\ln w_{ip} = \theta_{0p} + \theta_{1p} X_{ip} + \theta_{2p} Z_{jp} + \beta_\lambda \left(\frac{-\phi\left(\widehat{\gamma}_0 + \widehat{\gamma}_1 X_i' + \widehat{\gamma}_2 Z_j'\right)}{1 - \Phi\left(\widehat{\gamma}_0 + \widehat{\gamma}_1 X_i' + \widehat{\gamma}_2 Z_j'\right)} \right) + \eta_p \tag{G.23}$$

où $E\left(\eta_p \mid I_i = 0\right) = 0$.

Estimation de la probabilité de travailler dans une grande entreprise

Variables	coefficients	$\mid t - Student \mid$
primaire	0.038	0.28
secondaire classique	0.277	1.62*
secondaire technique	0.094	0.59
université Côte d'Ivoire	0.787	2.02**
université étrangère	0.208	0.56
apprenti	0.079	0.62
marié	-0.148	-0.30
sexe	0.145	0.76
syndiqué	0.907	8.26***
ivoirien	-0.633	5.42***
constante	1.037	4.50***
N	797	
pseudo R^2	0.32	

Notes: (*): significatif au seuil de 10%, (**): significatif au seuil de 5%, (***): significatif au seuil de 1%; les observations sont pondérées par l'inverse de la probabilité de faire partie de l'échantillon; l'estimation inclue cinq variables de catégories socio-professionnelles.

Tableau de contingences

	prédites		
réelles	0	1	total
0	499	67	566
1	127	104	231
total	626	171	797

Annexe H
Estimation de frontières d'efficience (Cornwell et Schmidt, 1996)

La notion de frontière d'efficience a été développée par Farrell (1957) et a fait l'objet de nombreux travaux (parmi lesquels Battese et Coelli, 1988 et 1992, Jondrow, Lovell et Schmidt, 1982). On distingue trois formes d'inefficience: l'inefficience technique, l'inefficience allocative et l'inefficience d'échelle. C'est à la première que nous consacrons notre analyse; celle-ci se définit comme étant la relation technique qui permet d'obtenir un output maximal pour une combinaison de facteurs de production et une technologie donnés. Cette fonction est la frontière de production. L'écart pour chaque entreprise d'un échantillon entre l'output produit et l'output réalisable sur la frontière représente une mesure de son inefficience. L'inefficience technique peut être améliorée sans investissements supplémentaires puisqu'il suffit d'améliorer l'utilisation des facteurs utilisés pour atteindre la frontière.

Pour l'estimation de l'efficience technique, on distingue l'approche non paramétrique de l'approche paramétrique ou économétrique. La première a l'avantage de ne pas spécifier une fonction particulière de la frontière. Les deux méthodes les plus couramment utilisées sont la méthode Datta Envelopment Analysis et Free Disposal Hull. La frontière est non statistique et déterminée par un programme linéaire.

L'approche paramétrique de la frontière de production est basée sur une spécification particulière de la technologie dont il s'agit d'estimer les paramètres. On introduit une variable asymétrique d'erreur dans la fonction de production qui représente l'inefficience. Ces frontières peuvent être déterministes ou stochastiques. Une partie de la littérature consacrée aux frontières stochastiques utilise des données en coupe et ignore la dimension temporelle lors de l'estimation de l'efficience. L'identification du modèle y est rendue possible grâce à de fortes hypothèses quant aux distributions des deux termes d'erreurs. L'utilisation de données de panel a permis de lever ces hypothèses en supposant que le terme d'efficience est invariant dans le temps et devient ainsi un effet spécifique à l'entreprise. Ces effets peuvent être estimés sans hypothèses de distribution et peuvent ensuite être convertis en mesure d'inefficience. En fait, les développements de l'utilisation de données de panel permettent de lever également l'hypothèse d'efficience constante dans le temps.

Considérons une frontière de production, $f(X_i;\beta)$, associée à un niveau d'output, y_i, par un degré d'efficience technique, a_i :

$$y_i = a_i f(X_i;\beta), 0 < a_i \square 1 \tag{H.1}$$

où $i = 1, ..., N$ représente les entreprises.

Cette expression transformée en logarithme devient

$$\ln y_i = \ln f(X_i;\beta) - u_i \tag{H.2}$$

avec $u_i = -\ln a_i \geq 0$ représentant l'inefficience technique. L'efficience technique est donc donnée par $\exp(-u_i)$.

Pitt et Lee (1981) et Schmidt et Sickles (1984) ont établit le lien entre l'estimation de frontière de production et l'utilisation de données de panel. Le modèle considéré est

$$y_{it} = \alpha + X_{it}\beta + v_{it} - u_i, i = 1, ..., N, t = 1, ..., T \tag{H.3}$$

$u_i \geq 0$ représente de le terme d'inefficience technique.

En notant $\alpha_i = \alpha - u_i$, nous avons

$$y_{it} = \alpha_i + X_{it}\beta + v_{it} \tag{H.4}$$

Nous supposons que les v_{it} sont i.d.d. $(0, \sigma_v^2)$ et non corrélés avec les inputs.

Deux procédures de tests sont retenues. Elles se distinguent par la conception déterministe ou stochastique du terme résiduel.

Frontière déterministe

Nous considérons les niveaux d'inefficience u_i comme fixes. Nous ne formulons pas d'hypothèse quant à la distribution du terme d'inefficience, et ce terme peut être ou non corrélé avec les inputs et les v_{it}. Nous appliquons alors l'estimation within et obtenons l'estimateur within de β, $\widehat{\beta}_w$, qui est convergent pour N ou $T \to \infty$.

L'inefficience technique peut alors être calculée.

$\widehat{\alpha}_i$ est calculé à partir de la moyenne des résidus pour chaque entreprise[68]

$$\widehat{\alpha}_i = y_{i.} - X_{i.}\widehat{\beta}_w \tag{H.5}$$

où $y_{i.} = \sum_{t=1}^{T} y_{it}, X_{i.} = \sum_{t=1}^{T} X_{i.}$.

Les inefficiences techniques spécifiques à l'entreprise sont estimées par

$$\widehat{u}_i = \widehat{\alpha}_M - \widehat{\alpha}_i \text{ avec } \widehat{\alpha}_M = \max_i(\widehat{\alpha}_i) \tag{H.6}$$

[68] Les α_i peuvent également être estimés directement comme les coefficients des variables muettes spécifiques à l'entreprise.

L'entreprise la plus efficace est définit pour une efficience de 100% et les niveaux d'efficience technique des autres entreprises sont données par $\exp(-\widehat{u}_i)$.

Frontière stochastique

L'approche stochastique de la frontière de production (proposée simultanément par Aigner, Lovell et Schmidt (1977), Battese et Corra (1977) et Meeusen et Van den Broeck (1977)) est motivée par l'idée que les déviations de la frontière de production ne peuvent être entièrement contrôlées par l'individu étudié. Selon l'interprétation d'une frontière déterministe, un nombre inhabituel de pannes aléatoires de l'équipement, le mauvais temps, par exemple, apparaissent systématiquement comme inefficience. De même, toute erreur ou imperfection dans la spécification du modèle ou des mesures des variables peuvent également être traduites comme une croissance de la mesure de l'inefficience. Une interprétation plus réaliste est que chaque producteur fait face à sa propre frontière de production, et que cette frontière est placée de façon aléatoire par des éléments aléatoires qui peuvent entrer dans le modèle sans le contrôle du producteur.

Nous supposons maintenant que les u_i sont i.d.d., distribuées positivement et non corrélées avec les v_{it} et les X_{it}. Nous avons

$$y_{it} = \alpha + X_{it}\beta + \varepsilon_{it} \tag{H.7}$$

avec

$$\varepsilon_{it} = v_{it} - u_i \tag{H.8}$$

L'estimateur du modèle à erreur composée est l'estimateur des moindres carrés quasi généralisés (MCQG), qui est convergent pour $N \to \infty$. Cet estimateur est efficace si les conditions d'orthogonalité sont respectées. Les estimations de l'efficience

technique peuvent être construites comme précédemment,

$$\widehat{\alpha}_i = y_{i.} - X_{i.}\widehat{\beta}_{mcqg} \tag{H.9}$$

et

$$\widehat{u}_i = \widehat{\alpha}_M - \widehat{\alpha}_i, \text{où } \widehat{\alpha}_M = \max_i(\widehat{\alpha}_i) \tag{H.10}$$

La convergence des $\widehat{u}_i$ requiert que N et $T \to \infty$.

Cornwell et Schmidt (1996, p.860) notent que "l'estimationn de l'inefficience technique procède à partir des résidus EIV[69] exactement comme cela a été décrit pour les MCQG".

[69] *Efficient instrumental variables*, c'est-à-dire les estimateurs de Hausman et Taylor (1981), Amemiya et MaCurdy (1986) et Breush, Mizon et Schmidt (1989), présentés dans l'annexe E.

Nous avons souligné le fait que l'utilisation de données de panel nécessite moins d'hypothèses que l'estimation menée sur données transversales. Cependant, il est également possible d'adopter les mêmes hypothèses concernant la distribution du terme d'efficience. Pitt et Lee (1981) ont par exemple supposé que les v_{it} sont normalement distribués et que les u_i suivent une loi semi normale. Un des avantages de cette approche est de pouvoir estimer la constante de la frontière, α, directement sans recourir à l'opérateur 'max'. L'efficience de la meilleure entreprise de l'échantillon n'est plus normalisée à 100%. Pour estimer les efficiences spécifiques aux entreprises, Battesse et Coelli (1988) considèrent l'estimateur

$$u_i = E\left(u_i \mid v_{i1} - u_i, v_{i2} - u_i, ..., v_{iT} - u_i\right) \qquad \text{(H.11)}$$

Efficiences variables dans le temps

La restriction consistant à rendre les inefficiences constantes dans le temps est très forte, en particulier quand T est élevé. Cette hypothèse peut être levée en considérant que $u_{it} \geq 0$ est une mesure de l'inefficience technique de l'entreprise i à la date t, $\alpha_{it} = \alpha_t - u_{it}$. En suivant Cornwell, Schmidt et Sickles (1990), si on peut estimer α_{it} pour tout i et t, il est possible de calculer

$$\widehat{\alpha}_{Mt} = \max_i \left(\widehat{\alpha}_{it}\right), \widehat{u}_{it} = \widehat{\alpha}_{Mt} - \widehat{\alpha}_{it} \qquad \text{(H.12)}$$

Ce modèle suppose que l'inefficience technique évolue dans le temps selon une spécification particulière. Une paramétrisation supplémentaire est donc nécessaire. Différentes fonctions ont été proposées par Cornwell, Schmidt et Sickles (1990), Kumbhakar (1990), Battese et Coelli (1992).

Ces modèles supposent que toutes les entreprises ont la même évolution de l'inefficience dans le temps. Cette hypothèse peut cependant être relâchée en considérant une fonction spécifique à chaque entreprise; on émet alors une hypothèse forte au niveau de la spécification de l'évolution de l'inefficience dans le temps.

Annexe I
Différentiels de salaires selon les catégories d'emplois et les secteurs d'activité (après introduction de la coopération estimée)

	S1 *Bois*	*S2* *Métal*	*S3* *Textile*	*S4* *Agro-ali.*	**Equation (3a)**
CS1 Patrons cadres	1.23 (6.08)***	1.49 (2.48)**	0.80 (1.19)	1.30 (10.54)***	1.08 (12.2)***
CS2 Techniciens	0.73 (4.65)***	0.65 (2.96)***	0.65 (4.55)***	1.20 (8.48)***	0.68 (7.71)***
CS3 Cadres adm. moyen	0.63 (3.86)***	0.48 (2.08)*	0.79 (6.13)***	0.89 (5.76)***	0.57 (7.18)***
CS4 Bureau commerce	0.78 (5.38)***	1.00 (4.16)***	0.47 (1.79)	0.48 (2.81)***	0.50 (5.61)***
CS5 Ouvriers qualifiés	0.31 (2.61)***	0.24 (1.59)	0.34 (2.60)***	0.46 (2.69)***	0.13 (1.98)**
CS6 Ouvriers spécialisés	0.28 (1.61)**	0.21 (1.35)	0.20 (1.46)*	référence	référence
Equation (2a)	0.27 (4.24)***	0.33 (3.49)***	0.34 (5.21)***	référence	

Annexe J
Démonstration de la proposition 1

La démonstration repose sur le fait que les contraintes (7) et (8) sont saturées, indiquant que les deux engagements du principal doivent être crédibles.

Soient λ_1 et λ_2 les multiplicateurs de Lagrange du programme $[P_2']$. Le Lagrangien de ce programme s'écrit

$$\begin{aligned} L \;=\; & Q(\widehat{a})\left[P(\widehat{e})\,w_g + (1-P(\widehat{e}))\,w_b\right] + (1-Q(\widehat{a}))\,w_\emptyset + d(\widehat{a}) \\ & -\lambda_1\left[Q(\widehat{a})\left[P(\widehat{e})\,u(w_g) + (1-P(\widehat{e}))\,u(w_b)\right]\right] \\ & -\lambda_1\left[(1-Q(\widehat{a}))\,u(w_\emptyset) - v(\widehat{e}) - \underline{U}\right] \\ & -\lambda_2\;\, Q(\widehat{a})\left[u(w_g) + u(w_b)\right] - \frac{v'(\widehat{e})}{P'(\widehat{e})}\Bigg] \end{aligned}$$

Les conditions de premier ordre sont

$$(1-Q(\widehat{a})) - \lambda_1(1-Q(\widehat{a}))\,u'(w_\emptyset) = 0 \Longleftrightarrow \lambda_1 = \frac{1}{u'(w_\emptyset)} \tag{J.1}$$

$$Q(\widehat{a})\,P(\widehat{e}) - \lambda_1 Q(\widehat{a})\,P(\widehat{e})\,u'(w_g) - \lambda_2 Q(\widehat{a})\,u'(w_g) = 0 \tag{J.2}$$

$$\Longleftrightarrow \lambda_2 = P(\widehat{e})\;\; \frac{1}{u'(w_g)} - \frac{1}{u'(w_\emptyset)}\Bigg] \tag{J.3}$$

$$Q(\widehat{a})\,(1-P(\widehat{e})) - \lambda_1 Q(\widehat{a})\,(1-P(\widehat{e}))\,u'(w_b) + \lambda_2 Q(\widehat{a})\,u'(w_b) = 0 \tag{J.4}$$

c'est-à-dire, en combinant (16), (18) et (19)

$$\frac{1}{u'(w_\emptyset)} = \frac{P(\widehat{e})}{u'(w_g)} + \frac{(1-P(\widehat{e}))}{u'(w_b)} \tag{J.5}$$

Cette équation est indépendante du niveau de l'effort de contrôle $\widehat{a}$.
Nous avons donc

$$w_g > w_\emptyset > w_b \tag{J.6}$$

ce qui est incompatible avec la contrainte (4.75). Par conséquent, la solution $[P_2']$ ne peut être atteinte par $[P_1]$.
De plus, nous pouvons montrer que la contrainte (4.74) est également active.
De la condition nécessaire de premier ordre par rapport à la variable a dans le programme $[P_2']$, nous avons

$$\begin{aligned} & w_\emptyset - P(\widehat{e})\,w_g - (1-P(\widehat{e}))\,w_b + P(\widehat{e})\left[u(w_g) - u(w_b)\right]\frac{1}{u'(w_g)} \\ & \qquad - \left[u(w_\emptyset) - u(w_b)\right] = \frac{d'(\widehat{a})}{Q'(\widehat{a})} \end{aligned} \tag{J.7}$$

La contrainte (4.75) n'est vérifiée que si

$$P(\widehat{e})\left[u\left(w_g\right)-u\left(w_b\right)\right]\frac{1}{u'\left(w_g\right)}-\left[u\left(w_\emptyset\right)-u\left(w_b\right)\right]\frac{1}{u'\left(w_\emptyset\right)}=0 \qquad \text{(J.8)}$$

c'est-à-dire, si la solution de $[P_2']$ vérifie l'égalité

$$\frac{u\left(w_\emptyset\right)}{u'\left(w_\emptyset\right)}=\frac{P(\widehat{e})\,u\left(w_g\right)}{u'\left(w_g\right)}+\frac{\left(1-P(\widehat{e})\right)u\left(w_b\right)}{u'\left(w_b\right)} \qquad \text{(J.9)}$$

Le contrat optimal est défini par (4.70), (4.72) et $(J.5)$, il est donc clair que sur l'espace de fonction $u(.)$, $(J.9)$ n'est pas vérifiée.

Annexe K
Démonstration de la proposition 2

Le programme $[P_3]$ a pour Lagrangien

$$
\begin{aligned}
L &= Q(\widehat{a})[P(\widehat{e})w_g + (1-P(\widehat{e}))w_b] + (1-Q(\widehat{a}))w_\emptyset \\
&+Q(\widehat{a})[P(\widehat{e})s_g + (1-P(\widehat{e}))s_b] + (1-Q(\widehat{a}))s_\emptyset \\
&-\lambda_1\{Q(\widehat{a})[P(\widehat{e})u(w_\emptyset) + (1-P(\widehat{e}))u(w_b)]\} \\
&-\lambda_1\{(1-Q(\widehat{a}))u(w_\emptyset) - v(\widehat{e}) - \underline{U}\} \\
&-\lambda_2\left\{Q(\widehat{a})[u(w_\emptyset)+u(w_b)] - \frac{v'(\widehat{e})}{P'(\widehat{e})}\right\} \\
&-\lambda_3\{Q(\widehat{a})[P(\widehat{e})(s_g + k(w_g - w_\emptyset)) + (1-P(\widehat{e}))s_b]\} \\
&-\lambda_3\{(1-Q(\widehat{a}))s_\emptyset - d(\widehat{a})\} \\
&-\lambda_4\left\{P(\widehat{e})(s_g + k(w_g - w_\emptyset)) + (1-P(\widehat{e}))s_b - s_\emptyset - \frac{d'(\widehat{a})}{Q'(\widehat{a})}\right\} \\
&-\lambda_5\{s_b - s_\emptyset - k(w_\emptyset - w_b)\} \\
&-\lambda_6\{w_g - w_\emptyset\} - \lambda_7 w_b - \lambda_8 w_g - \lambda_9 w_\emptyset - \lambda_{10} s_b - \lambda_{11} s_g - \lambda_{12} s_\emptyset
\end{aligned}
$$

où les λ_i sont les multiplicateurs de Lagrange.

Les conditions de premier ordre sont

$$\frac{\partial L}{\partial w_g} = P(\widehat{e})[Q(\widehat{a})(1-\lambda_3 k) - \lambda_4 k] - \lambda_6 - \lambda_8 = 0 \quad \text{(K.1)}$$

$$\frac{\partial L}{\partial w_b} = (1-P(\widehat{e}))Q(\widehat{a})[1-\lambda_1 u'(w_b)] \quad \text{(K.2)}$$

$$-\lambda_2 Q(\widehat{a})u'(w_b) - \lambda_5 k - \lambda_7 \quad \text{(K.3)}$$

$$\frac{\partial L}{\partial w_\emptyset} = (1-Q(\widehat{a})) + u'(w_\emptyset)[-\lambda_2 Q(\widehat{a}) + \lambda_1 Q(\widehat{a})(1-P(\widehat{e})) + 1] \quad \text{(K.4)}$$

$$+\lambda_3 Q(\widehat{a})P(\widehat{e})k + \lambda_4 P(\widehat{e})k + \lambda_5 k + \lambda_6 - \lambda_9 \quad \text{(K.5)}$$

$$\frac{\partial L}{\partial s_b} = (1-P(\widehat{e}))[Q(\widehat{a})(1-\lambda_3) - \lambda_4] - \lambda_5 - \lambda_{10} = 0 \quad \text{(K.6)}$$

$$\frac{\partial L}{\partial s_g} = P(\widehat{e})Q(\widehat{a})(1-\lambda_3) - \lambda_4 P(\widehat{e}) - \lambda_{11} = 0 \quad \text{(K.7)}$$

$$\frac{\partial L}{\partial s_\emptyset} = (1-Q(\widehat{a}))(1-\lambda_3) + \lambda_4 + \lambda_5 - \lambda_{12} = 0 \quad \text{(K.8)}$$

De la contrainte d'incitation de l'agent et de la contrainte $w_g \geq w_\emptyset$,nous avons $w_g \succ 0$ et $w_\emptyset \succ 0$.

Par conséquent, $\lambda_8 = \lambda_9 = 0$.

Comme $Q(\widehat{a}) \geq Q(\widehat{a})\lambda_3 + \lambda_4$ de la quatrième condition de premier ordre, et $k \prec 1$, nous avons $\lambda_6 \succ 0$.
Ce qui implique, $w_\emptyset = w_g$, or l'optimalité requiert $w_\emptyset \prec w_g$.

Annexe L
Corrélation entre l'efficience et la politique de délégation par types d'entreprises

	salaire empl.	prod.	salaire empl.	superv.	capital	(log)	nombre	d'empl.
		corr.		corr.		corr.		corr.
quartile 1	35622	0.2725	229120	-0.0813	13.59	0.0917	7	0.2611
quartile 2	54928	-0.1565	311030	0.1627	17.52	0.1875	39	0.092
quartile 3	65891	0.2249	373080	0.1197	19.00	0.0763	81	0.1203
quartile 4	81075	0.0910	379090	0.3279	20.72	0.0658	362	-0.011
moyenne	73110		356361		19.08		318	
écart-type	23701		74187		2.22		797	