

JUIN 1999

Capital humain, productivité agricole, et travail féminin : variables latentes et séparabilité dans les modèles de ménage[#]

par

Claudio Araujo*, Catherine Araujo Bonjean* et Jean-Louis Arcand**

*CERDI-CNRS, Université d'Auvergne

** CERDI-CNRS, Université d'Auvergne
Instituto de Saude Coletiva (ISC), Universidade Federal da Bahia (UFBA)

Résumé : L'estimation d'une fonction de production agricole au niveau des parcelles incorporant comme variable explicative le capital humain moyen du ménage fait généralement apparaître un effet négatif de ce dernier sur la production. Nous expliquons ce paradoxe, qui apparaît souvent dans les travaux appliqués à l'agriculture sur l'impact du capital humain sur la production, dans le cadre d'un modèle de ménage non-séparable. Une nouvelle approche des tests de séparabilité, basée sur l'estimation d'un modèle de « switching regression », est proposée. Les résultats montrent que l'effet du capital humain sur la demande de main-d'œuvre féminine est positive dans le cas du régime non séparable tandis que, dans le cas du régime séparable, le capital humain n'apporte une contribution importante à la production que lorsque le ménage est soumis à une contrainte de crédit saturée.

[#] Nous remercions le projet PARADI financé par l'ACDI pour le financement qui a permis la création de la banque de données.

1. INTRODUCTION

Dans ce papier, nous présentons une nouvelle approche à la question de la séparabilité entre les décisions de production et de consommation des ménages paysans dans les PED. Nous fournissons également une explication pour un résultat paradoxal qui apparaît souvent dans les travaux empiriques sur l'agriculture en Afrique, à savoir, l'impact négatif ou nul du capital humain sur la productivité agricole ou la demande optimale des facteurs.

Par opposition à l'approche introduite dans la littérature par Benjamin (1992) (voir aussi Singh, Squire et Strauss, 1986, Lopez, 1986, Rosenzweig, 1980, 1988, Jacoby, 1993, Udry, 1996), dans laquelle on teste pour la séparabilité par le biais de restrictions d'exclusion sur l'ensemble d'un échantillon, et en contraste avec l'approche basée sur les modèles de structure de classe (Bardhan, 1982, Roemer, 1982, Pant, 1983, Eswaran et Kotwal, 1986, Swamy, 1994) où l'on teste ces mêmes restrictions d'exclusion en ayant préalablement séparé l'échantillon sur la base de caractéristiques observables des ménages (la séparabilité sélective, voir DeJanvry et Sadoulet, 1996, Arcand, Conning et Ethier, 1998), l'approche proposée dans ce papier permet à chaque observation de révéler son appartenance au régime séparable ou au régime non-séparable. Ainsi, la présente approche devrait fournir une voie alternative permettant d'identifier les sources de la non-séparabilité, une voie qui devrait éventuellement permettre de raffiner les modèles structurels liés à la séparabilité sélective.

Rappelons que la question de la séparabilité n'est pas purement une curiosité scientifique : l'absence de séparabilité implique que la productivité marginale des facteurs n'est pas égalisée et induit donc une perte d'efficacité, par rapport à l'optimum de premier rang, qui peut être importante. De plus, nos résultats, basés sur une « régression alternative » (switching regression) permettent d'opérer la quadrature du cercle : alors que l'impact du capital humain était statistiquement nul dans le cas de l'estimation conventionnelle de la demande de main-d'œuvre féminine, il devient positif et statistiquement significatif dans

l'estimation que nous proposons. On retrouve donc bien le signe attendu, et le rôle important du capital humain dans la production agricole se trouve ainsi confirmé.

Le plan du papier est comme suit. Dans la section 2 nous spécifions le modèle de ménage le plus simple qui permette de construire une charpente théorique pour les restrictions d'exclusion qui seront la base du test de séparabilité. Dans section 3 nous détaillons la procédure économétrique de la regression alternative avec équation de sélection exogène qui sera la base de notre travail empirique. Nous présentons également les résultats empiriques qui correspondent à notre modèle.

2. LE MODELE

Considérons le modèle de ménage agricole standard suivant. Soit

$$(1) \quad Q = F(A, H, T, X_M, X_F)$$

la fonction de production où Q représente l'output, $F(.)$ la fonction de production, A des facteurs affectant la productivité agricole (type de sol), H est le capital humain du chef du ménage paysan, T est la surface de la parcelle, X_M est l'intrant en main-d'œuvre masculine, X_F est l'intrant en main-d'œuvre féminine.

Les profits du ménage sont donnés par

$$(2) \quad \Pi = F(A, H, T, X_M, X_F) - w_M E_M - w_F E_F - K$$

où Π représente les profits du ménage, w_M est le salaire masculin, E_M est la main-d'œuvre embauchée masculine, w_F est le salaire féminin, E_F est la main-d'œuvre embauchée féminine, et K représente les coûts fixes.

Soit L_M la dotation en main-d'œuvre masculine du ménage et L_F sa dotation en main-d'œuvre féminine. En supposant une parfaite substituabilité entre la main-d'œuvre embauchée et la main-d'œuvre familiale, on écrira

$$(3) \quad \begin{aligned} E_M &= X_M - L_M \\ E_F &= X_F - L_F \end{aligned}$$

ce qui nous permet de réécrire les profits du ménage comme

$$(4) \quad \Pi = F(A, H, T, X_M, X_F) - w_M (X_M - L_M) - w_F (X_F - L_F) - K.$$

Comme dans les modèles de structure de classe à la Eswaran et Kotwal (1986), le ménage est potentiellement assujéti à une contrainte de crédit donnée par :

$$(5) \quad w_M (X_M - L_M) + w_F (X_F - L_F) + K \leq B(W, H),$$

où $B(W, H)$ est la disponibilité de crédit du ménage (qui pourra être une fonction du niveau de capital humain du chef du ménage) et soit W un vecteur de caractéristiques du ménage susceptibles d'influencer la disponibilité de crédit (propriété foncière, par exemple).

La fonction de coûts du ménage est donnée par

$$(6) \quad C(w_M, w_F, A, H, T, Q) = \left\{ \min_{\{X_M, X_F\}} (w_M X_M + w_F X_F) \text{ s.c. } Q = F(A, H, T, X_M, X_F) \right\}$$

.

Il est évident que la contrainte de crédit pourra donc être réécrite comme

$$(7) \quad C(w_M, w_F, A, H, T, Q) \leq B(W, H) + w_M L_M + w_F L_F - K.$$

Considérons maintenant le problème de maximisation des profits du ménage

$$(8) \quad Q^* = \arg \max_{\{Q\}} \Pi = \arg \max_{\{Q\}} (Q - C(w_M, w_F, A, H, T, Q)).$$

Nous dénoterons la solution à ce problème par

$$(9) \quad Q^* = Q^*(w_M, w_F, A, H, T).$$

La probabilité qu'un ménage fait face à une contrainte saturée est donnée par

$$(10) \quad \Pr(\text{contraint}) = \Pr(C(w_M, w_F, A, H, T, Q^*) \geq B(W, H) + w_M L_M + w_F L_F - K).$$

Notons que l'on aura généralement $Q_H^* > 0$ et que le côté gauche de l'inégalité sera donc une fonction croissante du capital humain du chef. Si la disponibilité en crédit du ménage est également une fonction croissante du capital humain du chef, il suit que le signe de la dérivée de la probabilité d'être contraint par rapport au capital humain du chef pourra être ambiguë.

Pour un ménage dont la contrainte de crédit est saturée, nous pouvons définir le niveau d'output par

$$(11) \quad \bar{Q} = \{Q \text{ tel que } C(w_M, w_F, A, H, T, \bar{Q}) - B(W, H) + w_M L_M + w_F L_F - K = 0\}.$$

Nous dénoterons la solution à ce problème par

$$(12) \quad \bar{Q} = \bar{Q}(w_M, w_F, A, H, T, B(W, H), L_M, L_F, K)$$

Considérons maintenant la demande totale de main-d'œuvre féminine du ménage sur une parcelle donnée. Par le Lemme de Shephard, nous pouvons écrire :

$$(13) \quad X_F(w_M, w_F, A, H, T, Q) = C_{w_F}(w_M, w_F, A, H, T, Q).$$

Si le ménage est non-contraint, nous aurons

$$(14) \quad X_F^* = C_{w_F}(w_M, w_F, A, H, T, Q^*(.)) = C_{w_F}(w_M, w_F, A, H, T).$$

Par contre, si le ménage est contraint, on écrira

$$(15) \quad \bar{X}_F = C_{w_F}(w_M, w_F, A, H, T, \bar{Q}(.)) = C_{w_F}(w_M, w_F, A, H, T, B(W, H), L_M, L_F, K).$$

Les équations (14) et (15) définissent deux régimes correspondant à (i) des ménages dont les décisions de production sont indépendantes des caractéristiques du ménage (à part le capital humain du chef, qui affecte directement la production agricole) et à (ii) des ménages dont les décisions de production seront fonction de leurs caractéristiques. La spécificité du modèle relève du fait que la probabilité qu'un ménage soit dans le régime contraint est donné explicitement par l'équation (10).

3. METHODOLOGIE EMPIRIQUE ET RESULTATS

Le modèle théorique indique que, dans un échantillon donné, la non-séparabilité tiendra pour les ménages contraints à produire en dessous du niveau optimal de l'output, tandis que le comportement des ménages non-contraints sera caractérisé par la séparabilité. Le problème empirique qui se pose est évidemment que l'appartenance d'une observation au régime contraint ou au régime non-contraint est inobservable, du point de vue de l'économètre. Il suit que, techniquement, l'appartenance à un régime ou à l'autre est une variable latente.

Si l'approche des modèles de structure de classe est une réponse possible à ce problème, une alternative est de "laisser parler les données" dans le sens d'endogénéiser, du point de vue empirique, l'appartenance d'une observation donnée à un régime ou à l'autre.¹ Econométriquement, il s'agit d'estimer un "switching regression" (régression alternative) où la différence entre les deux régimes sera assurée par les restrictions d'exclusion qui découlent du modèle théorique (ici, la différence entre l'équation pour la demande de main-d'œuvre donnée par (14) versus (15)). Le choix entre les deux régimes, quant à lui, sera déterminé par une équation de sélection qui découle naturellement de l'équation représentant la probabilité d'être assujéti à une contrainte saturée (l'équation (10) ci dessus).

Notons que dans l'implémentation empirique des équations (14) et (15), la productivité de la terre (A) sera représentée par une série de quatre variables muettes correspondant aux types de sols, H sera donné par le capital humain du chef du ménage, T sera donné par la surface de la parcelle, la disponibilité de crédit $B(.)$ sera fonction de la propriété foncière (W) du ménage (c'est la source la plus importante de richesse des paysans du village), la valeur de la dotation en main-d'œuvre du ménage sera représentée par la taille du ménage (nombre de personnes adultes) et que les salaires féminins et masculins, parce qu'ils sont constants à travers l'échantillon, seront absorbés par la constante. Nous rajouterons également une variable muette « année » permettant de contrôler pour des différences entre les deux années agricoles et communes à toutes les parcelles.

Plus formellement, soit x_i (l'indice $i = 1, \dots, N$ correspond aux observations) la matrice des variables explicatives qui sont communes aux deux régimes : $x_i = [A \ H \ T]$. Soit $x_{1i} = [x_i \ W \ L]$ la matrice de variables explicatives correspondant au régime 1 (non-séparable; ici L dénotera la taille du ménage) et $x_{2i} = [x_i]$ la matrice des variables explicatives correspondant au régime 2 (séparable). La spécification de base est donnée par le système d'équations suivant :

¹ Pour une application d'un modèle de structure de classe, que l'on dénote aussi parfois comme l'approche

Regime 1 : $y_{1i} = \mathbf{b}'_1 x_{1i} + u_{1i}$

Regime 2 : $y_{2i} = \mathbf{b}'_2 x_{2i} + u_{2i}$

où y_{ji} , $j=1,2$ est l'intrant en main-d'œuvre totale (masculine ou féminine) et où le choix entre les deux régimes est donné par la condition

$$y_i = \begin{cases} y_{1i} & \text{if } \mathbf{g}'z_i + v_i < 0 \\ y_{2i} & \text{if } \mathbf{g}'z_i + v_i \geq 0 \end{cases}$$

avec $v_i \sim N(0,1)$, et où $z_i = [W \quad L \quad H]$. Nous supposons que $(u_{1i}, u_{2i}) \sim N(0,0, \Sigma)$ avec matrice de covariance

$$\Sigma = \begin{bmatrix} \mathbf{s}_1^2 & \mathbf{s}_{12} \\ \mathbf{s}_{12} & \mathbf{s}_2^2 \end{bmatrix}.$$

On sait (Maddala (1983), Quandt (1988)) que le terme \mathbf{s}_{12} n'est pas identifiable. On procédera ainsi aux normalisations suivantes (Maddala et Nelson (1975), p. 424) : (i) $v_i \sim N(0,1)$ dans l'équation de sélection (plutôt que $N(\mathbf{m}, \mathbf{s}^2)$) (ii) le coefficient associé avec la constante dans l'équation de sélection sera égale à un (en d'autres termes, les coefficients de l'équation de sélection ne sont déterminés que jusqu'à une constante multiplicative près).

L'équation de sélection qui détermine si une parcelle de terre correspond à une situation contrainte ou non-contrainte est donnée par une fonction indicatrice latente que l'on définit comme suit :

$$I_i(z_i) = \begin{cases} 1 & \text{si } \mathbf{g}'z_i + v_i < 0 \\ 0 & \text{si } \mathbf{g}'z_i + v_i \geq 0 \end{cases}$$

Nous pouvons alors réécrire notre système d'équation sous la forme suivante :

$$y_i = I_i(z_i) \mathbf{b}'_1 x_{1i} + (1 - I_i(z_i)) \mathbf{b}'_2 x_{2i} + I_i(z_i) u_{1i} + (1 - I_i(z_i)) u_{2i},$$

où nous prenons une approximation de $I_i(z_i)$ sous forme continue par la fonction probit :

$$\hat{I}_i(z_i) = \int_{-\infty}^{g'z_i} \left(\frac{\exp\{-v^2/2\}}{\sqrt{2\mathbf{p}}} \right) dv.$$

On peut alors aisément construire la fonction de vraisemblance, et la maximiser par rapport aux paramètres $\mathbf{b}_1, \mathbf{b}_2, \mathbf{s}_1, \mathbf{s}_2, \mathbf{g}$. Remarquons qu'un coefficient négatif dans la l'équation de sélection implique une plus forte probabilité de se trouver dans le régime contraint (non-séparable).

La procédure du switching regression est appliquée au cas du village tunisien d'El Oulja, pour lequel nous disposons d'une banque de données au niveau des parcelles individuelles ainsi qu'au niveau des ménages pour les années agricoles 1993 et 1995. Pour une description détaillée des données, nous renvoyons le lecteur à Ai, Arcand et Ethier (1996, 1997), Arcand (1998). Les résultats sont présentés au Tableau 1. Dans ce qui suit, nous nous limiterons à la demande pour la main-d'œuvre féminine.

Dans la colonne de gauche nous présentons les résultats d'une estimation par MCO de l'équation de demande pour la main-d'œuvre féminine, où nous incluons les caractéristiques du ménage (propriété foncière, taille du ménage). Un test pour la restriction d'exclusion sur les caractéristiques du ménage correspond donc au test habituel de séparabilité « tout ou rien » (Benjamin, 1992). On remarquera que le coefficient sur la taille du ménage est statistiquement significatif. Dans le contexte d'une étude classique sur la séparabilité, on conclurait donc que la séparabilité est rejetée. Le résultat le plus surprenant concerne le coefficient sur le niveau de capital humain du chef du ménage qui,

dans le contexte des technologies de production habituelles où le capital humain augmente la productivité marginale des autres facteurs de production, devrait être positif. Ici, il est négatif, bien que statistiquement insignifiant.

Dans la colonne de droite, par contre, nous présentons les résultats d'une estimation, par maximum de vraisemblance, du modèle présenté ci-dessus. Plusieurs aspects de ces résultats sont remarquables. Premièrement, le coefficient associé avec le capital humain du chef du ménage est positif et fortement significatif ($t = 5,83$) dans l'équation correspondant au régime non-séparable, alors qu'il est statistiquement nul dans le régime séparable. Deuxièmement, le coefficient associé avec la surface de la parcelle est négatif et statistiquement significatif dans le régime non-séparable, alors qu'il est statistiquement nul dans le régime séparable. Troisièmement, dans l'équation de sélection, le capital humain du chef du ménage a un effet positif (c.a.d. son coefficient est négatif) sur la probabilité de se trouver dans le régime non-séparable, alors que la taille du ménage augmente la probabilité de se trouver dans le régime non-contraint.

Le premier résultat montre que l'effet du capital humain sur la demande de main-d'œuvre féminine est bien positive (dans le cas du régime non-séparable), comme le prédit la théorie de la production. Que le coefficient associé avec le capital humain du chef soit statistiquement nul dans le régime séparable indique que le capital humain n'apporte une contribution importante à la production agricole que lorsque le ménage est assujéti à une contrainte saturée.

Le deuxième résultat indique que les rendements d'échelle dans le régime non-séparable sont décroissants, alors qu'ils sont constants dans le régime séparable. Ensemble, ces résultats suggèrent que les technologies de production utilisées par les ménages contraints et non-contraints ne sont pas les mêmes : l'estimation traditionnelle présentée à la première colonne souffre donc d'un important biais d'agrégation.

Le troisième résultat, quant à lui, indique que l'impact du capital humain sur les coûts de production (il les augmente) est plus important que son impact sur la capacité du ménage à obtenir du crédit. De même, la taille de ménage, parce qu'elle augmente la valeur de sa dotation en main-d'œuvre, réduit, *ceteris paribus*, ses besoins en crédit.

CONCLUSIONS

Dans ce papier nous proposons une nouvelle approche aux tests pour la séparabilité basée sur l'estimation d'une régression alternative (switching regression). Nos résultats suggèrent que les tests traditionnels de restrictions d'exclusion basées sur l'ensemble d'un échantillon souffrent de biais d'agrégation. En effet, nous trouvons que l'impact du capital humain du chef de ménage sur la demande de main-d'œuvre féminine diffère selon que l'on est dans le cas séparable ou le cas non-séparable. Ce biais d'agrégation pourrait être en grande partie responsable des résultats surprenant sur données africaines, où l'on trouve souvent un impact négatif du capital humain sur la productivité agricole ou la demande de main-d'œuvre.

En outre, nos résultats nous ont permis d'étudier les causes de la non-séparabilité dans le sens que l'équation de sélection identifie les facteurs qui affectent la probabilité qu'un ménage soit assujéti à une contrainte de crédit et se trouve donc dans le cas non-séparable. Finalement, notre approche nous aura permis de cerner deux canaux par lesquels le capital humain affecte la productivité agricole : (i) il augmente la probabilité d'être assujéti à une contrainte de crédit parce qu'il augmente, conforme avec la théorie de production de base, les coûts de production; (ii) par contre, pour les ménages assujétis à une contrainte de crédit, le capital humain augmente la demande de main-d'œuvre et donc la productivité agricole.

REFERENCES

- AI, C., ARCAND, J.-L., AND ETHIER, F. (1996), "Moral Hazard and Marshallian Efficiency: Evidence from Tunisia," C.R.D.E. Working Paper No. 0896.
- AI, C., ARCAND, J.-L., AND ETHIER, F. (1997), "De l'efficacité allocative des contrats agricoles : Cheung avait-il raison?," *Revue d'Économie du Développement* 2:103-127
- ARCAND, J.-L., CONNING, J., AND ETHIER, F. (1998), "Class Position and Economic Behavior in a Tunisian Village: Selective Separability in a Multi-Factor Household Model", version antérieure présentée au American Economic Association / Middle East Economic Association, New Orleans, 1996.

- ARCAND (1998), "Inefficacité marshallienne, partage de coûts, et modèles contractuels avec marchés manquants: résultats empiriques tunisiens", *L'Actualité Économique*, à paraître.
- BARDHAN, P. (1982), "Agrarian Class Formation in India", *Journal of Peasant Studies* 10:73-94.
- BENJAMIN, D. (1994), "Can Unobserved Land Quality Explain the Inverse Productivity Relationship?" *Journal of Development Economics*.
- ESWARAN, M., AND KOTWAL, A. (1986), "Access to capital and agrarian production organization," *Economic Journal* 96:482-498.
- JACOBY, H. (1993), "Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra," *Review of Economic Studies* 60:903-921.
- LOPEZ, R. (1986), "Structural Models of the Farm Household that Allow for Interdependent Utility and Profit-Maximization Decisions," in SINGH, SQUIRE AND STRAUSS.
- MADDALA, G. S. (1983). *Limited Dependent and Qualitative Variables in Econometrics* (Cambridge : Cambridge University Press).
- MADDALA, G. S. AND NELSON, F. D. (1975), "Switching Regression Models with Exogenous and Endogenous Switching," *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 423-426.
- PANT, C. (1983), "Tenancy and family resources: a model and some empirical analysis ", *Journal of Development Economics* 12:27-39.
- ROEMER, J. (1982). *A General Theory of Exploitation and Class* (Cambridge, MA: Harvard University Press).
- ROSENZWEIG, M. (1980), "Neoclassical Theory and the Optimizing Peasant: An Econometric Analysis of Market Family Labor Supply in a Developing Country," *Quarterly Journal of Economics* 95:31-55.
- ROSENZWEIG, M. (1988), "Labor Markets in Low Income Countries," Chapter 15 in Chenery, H. and T. N. Srinivasan, eds., *Handbook of Development Economics*, Volume I (Amsterdam: North Holland), pp. 713-762.
- SINGH, I., SQUIRE, L., and STRAUSS, J. (1986). *Agricultural Household Models: Extensions, Applications, and Policy* (Baltimore: Johns Hopkins).
- SWAMY, A. (1994), "Access to capital, agrarian classes, and resource allocation: the Punjab, 1933-36", Center for Institutional Reform and the Informal Sector, University of Maryland at College Park, Working Paper No. 96.

QUANDT, R. (1988). *The Econometrics of Disequilibrium* (Oxford : Basil Blackwell).

UDRY, C. (1996), "Efficiency and Market Structure: Testing for Profit Maximization in African Agriculture," mimeo, Northwestern University, February.

Tableau 1
Résultats empiriques (statistiques t de Student entre parenthèses)

Rappel : coefficient négatif dans l'équation de sélection => régime 1 (non-séparable) plus probable
(variable dép. : log de la main-d'oeuvre féminine totale par hectare ; 256 observations)

Équation de sélection	le paradoxe	une explication
capital humain du chef de ménage (H)		-0.146 (-3.50)
taille du ménage (L)		0.170 (3.25)
Régime non-séparable		
constante	3.847 (13.18)	3.764 (4.85)
variables de contrôle (A) :		
type de sol 1	0.576 (1.70)	-0.983 (-1.59)
type de sol 2	0.479 (1.34)	-2.126 (-0.08)
type de sol 3	0.081 (0.26)	-1.420 (-2.12)
type de sol 4	0.682 (1.60)	-1.704 (-1.52)
année	0.138 (0.80)	0.087 (0.17)
surface de la parcelle (T)	-0.129 (-1.61)	-1.304(-4.27)
capital humain du chef de ménage (H) :	-0.021 (-1.19)	0.216 (5.83)
caractéristiques du ménage (W L) :		
propriété foncière du ménage	-0.0002 (-0.54)	0.014 (1.10)
taille du ménage	0.032 (2.05)	-0.349 (-2.66)
S_1		0.902 (17.17)
Régime séparable		
constante		4.345 (14.86)
variables de contrôle (A)		
type de sol 1		0.479 (1.44)
type de sol 2		0.347 (0.98)
type de sol 3		-0.018 (-0.05)
type de sol 4		0.632 (1.16)
année		0.087 (0.60)
surface de la parcelle (T)		-0.083 (-1.23)
capital humain du chef de ménage (H)		0.003 (0.156)
S_2		0.519 (4.25)
probabilité estimée moyenne du régime 1		0.907
log de la fonction de vraisemblance	n.a.	-363.5

Méthode d'estimation : (i) première colonne : MCO avec écarts-types robustes de White,
(ii) deuxième colonne : maximum de vraisemblance.