

Scolarisation primaire et pauvreté au Sénégal

par Abdoulaye Diagne
Consortium pour la recherche
économique et sociale (CRES)
Faculté de sciences économiques
et de gestion
Université Cheikh Anta Diop de Dakar

et Amen T. Dovoédo
Faculté de Sciences économiques et de gestion
Université Cheikh Anta Diop de Dakar

Scolarisation primaire et pauvreté au Sénégal

Résumé

Cette étude est consacrée à la description et à la modélisation de la demande d'éducation primaire au Sénégal. Les choix des ménages relatifs à la scolarisation de leurs enfants (scolarisation ou non de l'enfant, type d'école retenu pour cette scolarisation, poursuite ou non à son terme de sa scolarité par l'enfant), sont décrits et expliqués simultanément dans le cadre d'un modèle probit séquentiel. Les simulations effectuées avec ce modèle montrent entre autres résultats, la discrimination (en termes de probabilité d'être scolarisé et de terminer cette scolarité) subie par les enfants ruraux par rapport aux enfants des zones urbaines, quel que soit leur sexe, et celle subie par les filles par rapport aux garçons, quelle que soit leur zone de résidence.

Classification JEL :

Mots clés : éducation primaire ; pauvreté ; analyse en composantes principales ; quintiles ; modèle probit séquentiel ; Sénégal.

Economic analysis and econometric modeling of primary education demand in Senegal

Abstract

This study describes enrolment rates and constructs an econometric model for primary education demand in Senegal, relying on the results of the QUID survey realized in 2001.

A sequential probit model is constructed to explain the different choices made by the households relatively to their children enrolment, namely : decision to enroll or not, choice of the type of school (public or private, confessional or non confessional), decision to allow the children to attain the last grade or not. A dozen of explanatory variables relating to children and households' characteristics and availability of education supply, are retained in this model. Simulations based on the model estimation results show the discrimination supported by rural children relatively to urban ones whatever their gender, and by girls relatively to boys, in terms of the probability distribution of the dependant variable of the model.

Keywords : primary education; poverty; principal components analysis; quintiles; sequential probit model; Senegal.

I. Introduction

En dépit de l'avance considérable dont il bénéficiait au moment de son indépendance politique en 1960, le Sénégal a enregistré de faibles progrès en matière de scolarisation universelle. Avec un taux brut de scolarisation de 69,4% en 2001 dans l'enseignement primaire, il se situe nettement en deçà de la moyenne de l'Afrique subsaharienne (85% en 2000). Pourtant, l'État consacre près de 28% de son budget de fonctionnement au secteur de l'éducation alors que la moyenne subsaharienne se situe à 15% au cours des années 1990. L'insuffisance de l'offre publique d'éducation est ainsi étroitement liée à l'inefficience des ressources publiques d'éducation, inefficience qui est elle-même le résultat en grande partie d'une allocation intra-sectorielle défavorisant l'enseignement primaire au profit des autres niveaux. C'est ainsi que la part de l'enseignement primaire dans les dépenses publiques de fonctionnement représente seulement 42% en 2001. La plus grande part de celles-ci est ainsi allouée aux autres niveaux qui ont pourtant des taux de rendement sociaux nettement plus faibles que ceux de l'enseignement élémentaire (D. Boccanfuso, A. Diagne 2003). L'incapacité apparente du gouvernement à opérer des transferts de ressources de l'enseignement supérieur à l'enseignement primaire a accentué les déséquilibres dans la composition des dépenses publiques dans ce dernier niveau d'enseignement, où près de 97% des dépenses récurrentes sont consacrées au paiement des salaires. Cette répartition du financement public au détriment des coûts non salariaux tend à être corrigée en partie par la contribution des ménages à la couverture des autres catégories de dépenses tels que l'achat de manuels et l'entretien.

Dans le schéma actuel de financement de l'éducation, tout se passe en définitive comme si, face aux contraintes budgétaires, les autorités gouvernementales ont choisi d'accorder la priorité au recrutement d'enseignants et au paiement des salaires, tout en pariant sur la capacité des parents d'élèves à prendre en charge les fournitures et les autres intrants scolaires. Un tel choix introduit un facteur supplémentaire de discrimination à l'égard des élèves dont les parents ne peuvent supporter ces dépenses faute de revenus suffisants.

Ainsi si le système éducatif sénégalais est confronté à d'importants problèmes d'offre, ils ne devraient pas faire perdre de vue cependant ceux liés à la demande d'éducation. Même si le système parvenait à accroître rapidement la production de services éducatifs, notamment dans l'enseignement primaire, il n'est pas garanti que la scolarisation universelle serait assurée. Les comportements des ménages vis-à-vis de l'école, et donc, les choix qu'ils opèrent à propos de la scolarisation de leurs enfants, peuvent être une contrainte majeure à la réalisation d'un tel objectif. C'est pourquoi, il est essentiel de connaître les facteurs principaux qui influent sur la décision des ménages en matière de demande d'éducation si l'on veut lever les obstacles qui empêchent ces derniers d'envoyer à l'école la totalité de leurs enfants en âge de la fréquenter.

Les ménages font face à une quadruple décision en matière de scolarisation. En premier lieu, ils doivent décider s'ils acceptent ou non de scolariser chacun de leurs enfants. En deuxième lieu, ils doivent choisir le type d'école public ou privé que fréquenteront les enfants à scolariser. En troisième lieu, s'ils préfèrent l'école privée, ils ont à choisir entre l'école confessionnelle et l'école laïque. En quatrième lieu, ils doivent décider si leurs enfants qui vont à l'école termineront ou non le cycle primaire. L'objet de cette étude est la modélisation de la séquence de ces décisions que doit prendre un ménage sénégalais relativement à

l'éducation élémentaire de ses enfants. On pourra alors identifier et évaluer l'importance relative des principaux facteurs qui agissent sur chacun de ces quatre niveaux de décision des parents.

Les travaux consacrés à la demande d'éducation primaire dans les pays en développement se sont multipliés ces dernières années, en raison du rôle décisif reconnu au capital humain dans les travaux récents sur la croissance et le développement économique. Parmi les travaux les plus récents sur ce thème, nous pouvons citer ceux de S. Handa (2002) sur le Mozambique, H. Skyt Nielsen (2001) sur la Zambie, et de Filmer et Pritchett (1998) sur l'Inde. Handa étudie, à l'aide d'un modèle probit, l'impact des facteurs relevant de l'offre et de ceux relevant de la demande, sur la décision de scolarisation de leurs enfants par les ménages ruraux du Mozambique. Les variables retenues par cet auteur du côté de l'offre d'éducation sont la durée du trajet à l'école primaire la plus proche, le nombre d'écoles primaires par poste administratif en 1996 et la variation de ce nombre entre 1993 et 1996, et le nombre de classes en cours de construction dans les postes administratifs d'appartenance des ménages. Du côté de la demande, l'auteur privilégie le revenu par tête, estimé par la consommation journalière par tête, le fait que le chef de ménage soit alphabétisé ou non, le fait qu'un membre adulte du ménage ait terminé ou non la 7^{ème} année de l'école primaire et le fait qu'une femme adulte du ménage ait terminé ou non la 5^{ème} année du primaire. Après avoir testé l'endogénéité des dépenses de consommation par tête, suite aux remarques et suggestions pertinentes de Rivers et Vuong (1988) sur cette question, l'auteur est amené à contourner cette difficulté économétrique en remplaçant cette variable par une variable instrumentale, à savoir la consommation médiane par tête dans chaque grappe de l'échantillon.

Skyt Nielsen, quant à elle, utilise une démarche analogue pour étudier l'impact des facteurs économiques sur la décision des ménages zambiens de scolariser leurs enfants. Cet auteur estime un modèle logit de la demande de scolarisation des enfants par les ménages. Mais, il retient du côté de la demande une liste plus complète de variables économiques, à savoir : la dépense par équivalent d'adulte du ménage (qui reflète le niveau de pauvreté de celui-ci), l'accès du ménage au crédit, les actifs détenus par le ménage, les coûts de transports et des fournitures scolaires. Le modèle estimé par cet auteur prend aussi en compte un effet spécifique aléatoire reflétant l'influence des facteurs communautaires inobservables d'ordre sociologique ou anthropologique. Une fois le modèle estimé, l'auteur procède au test de la pertinence d'un sous-échantillonnage selon la zone de résidence (urbaine et rurale) et selon le sexe des enfants, et au test d'exogénéité des dépenses de consommation par équivalent-d'adulte. Le premier test conduit l'auteur à estimer des modèles distincts pour les ménages ruraux et urbains, et pour les ménages urbains, à procéder à des estimations séparées pour les filles et les garçons. Quant au second test, il révèle que l'hypothèse d'exogénéité des dépenses par équivalent-adulte du ménage n'est pas valide pour le sous-échantillon des garçons résidant en zone urbaine. L'auteur est conduit ainsi à recourir, pour l'estimation du modèle relatif à ce sous-échantillon, à des variables instrumentales identifiantes comme l'âge et le carré de l'âge du chef de ménage et son secteur d'activité économique.

La dernière étude récente relative à la demande d'éducation primaire sur laquelle nous insisterons, à savoir celle de Filmer et Pritchett, porte sur l'Inde. Elle se différencie des deux précédentes par une analyse descriptive préliminaire plus fouillée, avant l'estimation de modèles logit de la demande d'éducation primaire pour l'Inde entière et pour ses différentes régions. La partie descriptive procède notamment à l'élaboration d'un indicateur synthétique de pauvreté à partir des résultats d'une analyse en

composantes principales (ACP) portant sur les avoirs des ménages. Cet indicateur de pauvreté sera ensuite utilisé par ces auteurs dans leurs différents modèles. Il leur permet en effet, de contourner le problème de l'endogénéité éventuelle du revenu ou de la consommation des ménages. C'est cette démarche qui sera retenue dans notre étude sur le Sénégal. Notre travail se distingue cependant des études présentées ci-dessus qui portent exclusivement sur la décision des ménages de scolariser ou non leurs enfants. Nous nous intéressons à l'ensemble des décisions qu'ils prennent dans le cadre de la scolarisation primaire de leurs enfants, c'est-à-dire également au choix du type d'école et à celui entre l'achèvement de l'école primaire ou une scolarité incomplète pour l'enfant.

II Niveau de vie des ménages et scolarisation des enfants au Sénégal : analyse descriptive

2.1 Construction d'un indicateur de niveau de vie

Les données utilisées proviennent de l'enquête QUID (Questionnaire unifié sur les indicateurs de développement), réalisée en 2001 par la Direction de la prévision et de la statistique du Sénégal. Le fichier ménages de l'enquête contient, entre autres, un volet important relatif aux avoirs des ménages et aux caractéristiques de leur logement. Ces variables permettent de construire un indicateur synthétique de leur niveau de vie. Cet indicateur donne une mesure du statut économique des différents ménages, plus pertinente ou moins conjoncturelle que leur revenu (assimilé à leurs dépenses annuelles totales).

L'indicateur est construit en utilisant l'ACP qui est une technique statistique classique d'extraction de l'information contenue dans un grand tableau de données décrivant les n individus d'une population à l'aide d'un nombre p élevé de caractères quantitatifs ou assimilés. Elle consiste à construire un ou plusieurs nouveaux caractères, combinaisons linéaires des caractères initiaux, de variabilité maximale et décroissante, c'est-à-dire restituant de façon décroissante le maximum d'information contenue dans le tableau de données, et orthogonaux entre eux. Dans cette étude, il semble naturel de s'en tenir à la première composante principale, si le pourcentage de la variance totale du nuage qu'elle restitue est relativement élevé, et de l'interpréter essentiellement comme un résumé des avoirs, c'est-à-dire de leur richesse ou de leur niveau de vie. L'indicateur synthétique de la richesse (ou de la pauvreté) des ménages se présente donc comme une combinaison linéaire des caractères initiaux centrés et réduits, leurs coefficients mesurant leurs poids appelés aussi scores.

Nous avons retenu 19 variables. La réalisation de l'ACP révèle que la première composante principale restitue 20,7% de l'inertie totale du nuage de points, et que les cinq premières composantes principales restituent 51,8% de l'inertie totale. La corrélation significative de la première composante principale avec la majorité des variables initiales (corrélations positives avec les avoirs de type électroménager et le standing du logement, corrélations négatives avec les variables reflétant la résidence en zones rurales et /ou un certain dénuement), justifie son assimilation à un indicateur de niveau de vie ou de pauvreté des ménages. Les scores associés aux différentes variables dans la première composante principale sont reproduits dans le tableau 1. Celui-ci donne également les moyennes et écarts-types des variables initiales à l'échelle nationale et pour trois groupes économiques construits à partir des quintiles associés à l'indicateur de niveau de vie.

L'examen de ce tableau confirme la pertinence de notre interprétation de la première composante principale comme une mesure du niveau de pauvreté des ménages. Les trois dernières colonnes

montrent en effet que l'intervalle inter-quintile supérieur se caractérise par les taux de possession les plus élevés pour les actifs non fonciers et non ruraux, et pour le standing de leur logement. Les intervalles inter-quintiles 1 à 4, outre leurs faibles taux de possession des appareils synonymes de « modernité », se signalent par un taux élevé (mais trompeur) de ménages propriétaires de leur logement. Ce taux est trompeur, car les logements en question sont de mauvaise qualité puisqu'ils n'ont généralement pas d'eau courante (66% des ménages appartenant aux quintiles 1 et 2 doivent recourir à des puits non protégés et autres sources pour leur approvisionnement en eau). L'interprétation retenue pour la première composante principale est donc pleinement justifiée.

Tableau 1 : Scores des variables de l'ACP pour la première composante principale, et résumés statistiques de ces variables pour l'ensemble des ménages et pour trois catégories de ménages de niveau de vie croissant.

Variabiles de l' analyse en composantes principales (ACP)	Poids des variables (scores)	Scores réduits (poids divisé par l'écart-type)	Moyenne ensemble des ménages	Ecart type. ensemble des ménages	Moyenne (inter-quintiles 1 et 2)*	Moyenne (inter-quintiles 3 et 4)**	Moyenne (inter-quintile 5)***
Nombre de pièces séparées du logement	0,019	0,008	4,05	2,391	3,99	3,83	4,59
Superficie de terre possédée	-0,089	-0,002	24,91	54,163	48,63	12,55	2,18
Nombre de têtes de bétail et gros animaux	-0,055	-0,004	3,25	13,759	6,58	1,34	0,40
Nombre de moutons, chèvres et animaux de taille moyenne	-0,052	-0,002	5,98	24,300	11,45	2,71	1,58
1=propriétaire du logement	-0,079	-0,182	0,75	0,434	0,92	0,66	0,60
1= le ménage a un fer à repasser électrique	0,121	0,630	0,04	0,192	0,00	0,00	0,18
1=le ménage a un réfrigérateur/congélateur	0,192	0,490	0,19	0,392	0,00	0,07	0,80
1=le ménage a un téléviseur	0,195	0,428	0,29	0,456	0,00	0,26	0,94
1=le ménage a un matelas/lit	0,053	0,243	0,95	0,218	0,89	0,98	1,00
1=le ménage a une radio/ cassette	0,079	0,188	0,77	0,421	0,60	0,85	0,95
1=le ménage a une montre/réveil	0,097	0,198	0,60	0,490	0,37	0,69	0,88
1=le ménage a une cuisinière moderne	0,124	0,561	0,05	0,221	0,00	0,02	0,22
1=le ménage a une voiture/camion	0,129	0,538	0,06	0,240	0,00	0,02	0,26
1=le ménage a une mobylette	0,035	0,199	0,03	0,176	0,01	0,03	0,07
1=logement de « bonne qualité »	0,171	0,414	0,22	0,413	0,00	0,19	0,71
1=robinet intérieur	0,199	0,412	0,37	0,483	0,00	0,45	0,97
1=robinet public ou du voisin	-0,057	-0,132	0,25	0,431	0,25	0,35	0,03
1=puits protégé	-0,037	-0,139	0,06	0,245	0,08	0,08	0,00
1=puits non protégé et autres sources	-0,137	-0,294	0,32	0,466	0,66	0,13	0,01
Ensemble de l'échantillon			0,00	1,000	-0,84	0,02	1,65

Notes :

* Il s'agit plus précisément des ménages ayant une valeur de l'indicateur de niveau de vie inférieure à celle du 2^{ème} quintile (les 40% les plus pauvres).

** Il s'agit des ménages ayant une valeur de l'indicateur de niveau de vie comprise entre celles du deuxième et du quatrième quintiles de cet indicateur.

*** Il s'agit des ménages ayant une valeur de l'indicateur de niveau de vie supérieure à celle du début du 4^{ème} quintile (les 20% les plus riches).

Source : QUID 2001, DPS.

Nous pouvons apprécier et comparer les distributions intra-régionales des ménages selon le niveau de vie dans le tableau 2.

Tableau 2 : Répartition des ménages selon la région et le niveau de vie

Quintiles	1	2	3	4	5	Total
Région	Effectif %	Effectif %	Effectif %	Effectif %	Effectif %	
Dakar	12 0,7%	86 4,8%	253 14,1%	589 32,9%	848 47,4%	1788 100,0%
Ziguinchor	95 36,1%	49 18,6%	65 24,7%	41 15,6%	13 4,9%	263 100,0%
Diourbel	69 14,5%	95 20,0%	114 24,0%	114 24,0%	83 17,5%	475 100,0%
St-Louis	115 25,1%	112 24,4%	109 23,7%	83 18,1%	40 8,7%	459 100,0%
Tambacounda	125 43,6%	81 28,2%	57 19,9%	14 4,9%	10 3,5%	287 100,0%
Kaolack	161 26,4%	209 34,3%	120 19,7%	78 12,8%	42 6,9%	610 100,0%
Thiès	116 17,3%	147 21,9%	196 29,3%	142 21,2%	69 10,3%	670 100,0%
Louga	198 36,9%	149 27,8%	105 19,6%	51 9,6%	33 6,2%	536 100,0%
Fatick	110 31,0%	124 34,9%	87 24,5%	29 8,2%	5 1,4%	355 100,0%
Kolda	145 50,0%	95 32,8%	41 14,1%	6 2,1%	3 1,0%	290 100,0%
Total	1146 20,0%	1147 20,0%	1147 20,0%	1147 20,0%	1146 20,0%	5733 100,0%

Ce dernier fait apparaître une nette différence entre Dakar, où les ménages des quintiles 4 et 5 sont majoritaires, et les autres régions, caractérisées par la prépondérance des 40 % des ménages les plus pauvres (les deux premiers quintiles).

Cette opposition flagrante entre Dakar et les autres régions s'observe également entre les zones urbaines et rurales (le tableau prouvant cette remarque n'a pas été reproduit par manque de place).

2.2 Scolarisation, localisation, genre et niveau de vie

Le paragraphe précédent a mis en évidence les grandes disparités du niveau de vie des ménages entre les régions, entre les zones urbaines et rurales. Il importe donc de s'interroger maintenant sur l'impact de cette situation sur la scolarisation des enfants et adolescents (individus âgés de 7 à 19 ans).

Le tableau 3 donne les taux de scolarisation dans l'enseignement primaire et le cycle inférieur de l'enseignement secondaire pour les différentes régions du pays et pour deux groupes économiques : les ménages les plus pauvres (1^{ère} et 2^{ème} quintiles) et les ménages les plus riches (5^{ème} quintile).

Ce tableau met en relief l'impact du niveau de pauvreté sur le taux de fréquentation des enfants scolarisables dans l'enseignement primaire (7 à 12 ans). Dans toutes les régions, ce taux est plus élevé pour les enfants du groupe économique le plus riche. C'est dans les régions de Tambacounda, Louga, et Thiès qu'on observe les écarts les plus élevés (supérieurs à 60 %).

Le tableau 3 montre aussi une différence importante entre les deux groupes économiques pour le cycle inférieur de l'enseignement secondaire. Les écarts les plus importants (plus de 44%) sont enregistrés à Louga, Fatick, Dakar et Thiès.

La tendance générale qui se dégage de la comparaison des tableaux 3 et 4, est que l'écart des taux de scolarisation entre les enfants des 20% des ménages les plus riches et ceux des 40% les plus pauvres n'augmente pas avec le niveau de la scolarité sauf pour la région de Fatick.

Tableau 3 : Taux de scolarisation relatifs aux enseignements primaire et moyen, par région et groupe économique

Région	Proportion des 7-12 ans inscrits à l'école primaire				Proportion des 13-16 ans inscrits dans l'enseignement moyen			
	Ensemble	40% les plus pauvres (1)	20% les plus riches (2)	Ecart (2)-(1)	Ensemble	40% les plus pauvres (3)	20% les plus riches (4)	Ecart (4)-(3)
Dakar	0,872	0,450	1,026	0,576	0,396	0,114	0,571	0,457
Ziguinchor	1,115	1,077	1,230	0,153	0,550	0,458	0,556	0,098
Diourbel	0,447	0,346	0,614	0,268	0,139	0,041	0,248	0,207
St-Louis	0,496	0,378	0,862	0,484	0,142	0,034	0,341	0,307
Tambacoun-Da	0,554	0,481	1,142	0,661	0,058	0,015	0,250	0,235
Kaolack	0,534	0,386	0,984	0,598	0,153	0,085	0,357	0,272
Thiès	0,681	0,492	1,102	0,610	0,188	0,043	0,484	0,441
Louga	0,394	0,287	0,914	0,627	0,091	0,027	0,500	0,473
Fatick	0,720	0,612	1,000	0,388	0,249	0,134	0,600	0,466
Kolda	0,731	0,696	1,000	0,304	0,130	0,095	-	-
Ensemble du Sénégal	0,651	0,487	0,977	0,490	0,227	0,078	0,505	0,427

Il nous faut examiner à présent les différences entre les taux de scolarisation des garçons et des filles. Il apparaît d'une part que l'écart de scolarisation est positif dans toutes les régions dans le groupe des 40% les plus pauvres. Les différentiels de taux les plus élevés en faveur des garçons pour ce groupe concernent les régions de Ziguinchor et de Kolda (plus de 26%), et les différentiels de taux nuls ou très faibles se trouvent à Louga, Dakar et dans une moindre mesure Thiès.

Tableau 4 : Taux de fréquentation actuelle des 7-12 ans selon le sexe, la région et le groupe économique

Sexe et groupe économique	Filles Groupe des 40% les plus pauvres (1)	Garçons Groupe des 40% les plus pauvres (2)	Ecart (2)-(1)	Filles Groupe des 20% les plus riches (3)	Garçons Groupe des 20% les plus riches (4)	Ecart (4)-(3)	Filles Ensemble. (5)	Garçons Ensemble. (6).	Ecart (6)-(5)
Région									
Dakar	0,448	0,451	0,003	0,989	1,063	0,074	0,830	0,915	0,085
Ziguinchor	0,941	1,204	0,263	1,375	1,000	-0,375	0,995	1,255	0,260
Diourbel	0,280	0,409	0,129	0,522	0,746	0,224	0,383	0,520	0,137
St-Louis	0,411	0,345	0,066	1,032	0,706	-0,326	0,521	0,472	-0,049
Tambacounda	0,432	0,536	0,104	1,000	1,333	0,333	0,486	0,634	0,148
Kaolack	0,355	0,419	0,064	0,882	1,103	0,221	0,503	0,564	0,061
Thiès	0,482	0,502	0,020	1,014	1,222	0,208	0,653	0,709	0,056
Louga	0,287	0,287	0,000	0,949	0,881	-0,068	0,417	0,373	-0,044
Fatick	0,588	0,633	0,045	1,000	1,000	0,000	0,655	0,778	0,123
Kolda	0,561	0,836	0,275	1,000	1,000	0,000	0,606	0,863	0,257
Ensemble du Sénégal	0,450	0,523	0,073	0,934	0,943	0,009	0,613	0,689	0,073

En ce qui concerne le groupe économique le plus aisé, on note une grande disparité entre les régions. On peut en effet distinguer trois types de régions : celles caractérisées par l'absence de discrimination fondée sur le genre (Fatick et Kolda), celles où les filles sont plus scolarisées que les garçons (Ziguinchor, St-Louis et Louga), et celles pratiquant une discrimination à l'encontre des filles (il s'agit de cinq autres régions, avec toutefois une grande différence pour Dakar où la discrimination pro-garçons est de loin la moins prononcée. La comparaison interrégionale des taux de scolarisation, tous groupes économiques confondus,

montre que la discrimination à l'encontre des filles est la règle générale, sauf à Louga et à St-Louis.

Il serait utile enfin d'examiner le lien entre le niveau de vie des ménages, la zone de résidence d'une part, leurs choix entre les différents types d'école pour la scolarisation primaire de leurs enfants, d'autre part.

Tableau 5 : Répartition des enfants âgés de 7 à 12 ans inscrits à l'école primaire et résidant en zones urbaines, selon le quintile d'appartenance et le type d'école fréquenté

Quintile	1	2	3	4	5	Total
Type d'école						
Ecole publique Effectif % dans le quintile	29 78,4%	118 87,4%	332 91,7%	670 87,1%	693 68,0%	1842 79,3%
Ecole privée confessionnelle Effectif % dans le quintile	1 2,7%	2 1,5%	6 1,7%	36 4,7%	132 13,0%	177 7,6%
Ecole privée laïque Effectif % dans le quintile	4 10,8%	12 8,9%	17 4,7%	54 7,0%	182 17,9%	269 11,6%
Autres types d'école Effectif % dans le quintile	3 9,1%	3 2,2%	7 1,9%	9 1,1%	12 1,2%	34 1,5%
Total Effectif % dans le quintile	37 100,0%	135 100,0%	362 100,0%	769 100,0%	1019 100,0%	2322 100,0%

Tableau 6 : Répartition des enfants âgés de 7 à 12 ans inscrits à l'école primaire et résidant en zone rurale, selon l'intervalle inter-quintile d'appartenance et le type d'école fréquenté

Inter-quintile	1	2	3	4	5	Total
Type d'école						
Ecole publique Effectif % dans l'inter-quintile	622 96,4%	557 93,1%	432 93,7%	177 95,2%	63 88,7%	1851 94,4%
Ecole privée confessionnelle Effectif % dans l'inter-quintile	11 1,7%	11 1,8%	12 2,6%	4 2,2%	6 8,5%	44 2,2%
Ecole privée laïque Effectif % dans l'inter-quintile	1 0,2%	5 0,8%	4 0,9%	3 1,6%	1 1,4%	14 0,7%
Autres types d'école Effectif % dans l'inter-quintile	11 1,7%	25 4,2%	13 2,8%	2 1,0%	1 1,4%	52 2,6%
Total Effectif % dans l'inter-quintile	645 100,0%	598 100,0%	461 100,0%	186 100,0%	71 100,0%	1961 100,0%

Les tableaux 5 et 6 montrent la nette préférence des ménages pour l'école publique, quel que soit leur niveau de vie et quelle que soit leur zone de résidence. Néanmoins, il apparaît également que cette préférence est plus forte en zones rurales, et moins prononcée pour les 20% des ménages les plus riches, notamment en zones urbaines. Le deuxième type d'école préféré par les ménages, loin derrière l'école publique, est l'école privée laïque en zones urbaines et l'école privée confessionnelle en zones rurales. Il s'avère que ce sont les ménages aisés (les 20% les plus riches) qui accordent plus de crédit à ces types d'école. Le troisième type d'école préféré par les ménages urbains aisés est l'école privée confessionnelle, contrairement aux plus pauvres qui se tournent en troisième position vers les écoles communautaires et autres. Chez les ménages ruraux, les écoles communautaires sont légèrement

préférées en troisième position parmi les ménages les plus pauvres, alors que les plus riches sont indifférents entre les écoles communautaires et autres et les écoles privées laïques.

III. Modélisation du processus de décisions séquentielles des ménages dans le cadre de la scolarisation de leurs enfants

Le modèle de la demande d'éducation utilisé dans cette étude s'inspire du cadre développé par Becker (1975), Gertler et Glewwe (1990), Alderman et alii (1996). Il est cependant modifié pour tenir compte de la séquence de décisions que le ménage prend dans le cadre de la scolarisation de ses enfants. Ce dernier est considéré comme étant à la fois une unité de consommation et une unité de production. Le capital humain, défini ici comme l'éducation reçue par les enfants du ménage, est le résultat de l'activité de production de celui-ci. On suppose aussi que le ménage tire une utilité U de sa consommation de biens C et du capital humain H de sa progéniture : $U=U(C, H)$.

Le premier argument de U se justifie par le fait que la scolarisation d'une partie ou de tous ses enfants contraint le ménage à réduire sa consommation courante pour investir en frais de scolarité, fournitures pédagogiques, transport, etc. Le deuxième argument de la fonction d'utilité du ménage, à savoir le stock de capital humain qu'il souhaite accumuler, dépend de la séquence des décisions importantes qu'il doit prendre relativement à l'éducation élémentaire de ses enfants. Ces décisions, qui sont supposées relever du chef de ménage, sont prises dans l'ordre logique suivant, pour chacun des enfants du ménage :

1. Choisir de scolariser ou de ne pas scolariser l'enfant.
2. Choisir entre l'école publique et l'école privée, lorsque l'option de scolarisation de l'enfant a été retenue.
3. Choisir entre la poursuite ou la non poursuite à son terme par l'enfant de la scolarité primaire dans une école publique, lorsque ce type d'école a été préféré.
4. Choisir entre l'école confessionnelle et l'école privée, lorsque la préférence du ménage va plutôt à la scolarisation dans une institution privée.
5. Permettre ou non à l'enfant de terminer son cycle primaire dans une école confessionnelle, lorsque c'est ce type d'école qui a été retenu pour l'enfant.
6. Permettre ou non à l'enfant de terminer son cycle élémentaire dans une école laïque, lorsqu'il a été inscrit dans ce type d'école.

Chacune de ces six décisions binaires α relatives à un enfant quelconque i , peut être représentée formellement par une variable indicatrice $\delta_{i\alpha}$ qui prend la valeur 1 lorsque le choix est positif, et la valeur 0 lorsque le choix est négatif. De plus, chaque décision est censée prise de façon rationnelle, c'est-à-dire qu'elle résulte de la maximisation de la fonction d'utilité U introduite plus haut. La valeur optimale $U^*_{i\alpha}$ de U relative à la décision α et à l'enfant i , peut être approximée par une fonction linéaire de la consommation du ménage et du capital humain, ce dernier étant relié de façon linéaire aux caractéristiques du ménage, de l'enfant et du système éducatif regroupées dans le vecteur $X_{i\alpha}$. On a donc :

$$U^*_{i\alpha} = C_{i\alpha}k_{\alpha} + H_{i\alpha}\beta_{\alpha} = C_{i\alpha}k_{\alpha} + X_{i\alpha}\beta_{\alpha} + u_{i\alpha} \quad (1)$$

Le terme β_α désigne le vecteur des coefficients des différentes variables expliquant le capital humain, le terme k_α représente le coefficient associé à la consommation, et le terme $u_{i\alpha}$ représente l'erreur aléatoire dont est entachée la relation supposée entre le niveau d'utilité et les variables explicatives retenues.

Selon un raisonnement bien établi dans la spécification des modèles à réponses qualitatives, la connexion entre la variable indicatrice décisionnelle δ_α et le niveau optimal (inobservable) de l'utilité $U^*_{i\alpha}$ s'effectue de la façon suivante :

$$(\delta_\alpha=1 \text{ si } U^*_{i\alpha}>0) \text{ et } (\delta_\alpha=0 \text{ si } U^*_{i\alpha}\leq 0). \quad (2)$$

La probabilité d'une décision α négative relative à l'enfant i s'écrit donc :

$$\Pr(\delta_\alpha=0)=\Pr(U^*_{i\alpha}\leq 0)=\Pr(u_{i\alpha}\leq -X_{i\alpha}\beta_\alpha)=F(-X_{i\alpha}\beta_\alpha), \quad (3)$$

F étant la fonction de répartition de la loi suivie par la variable aléatoire $u_{i\alpha}$. Par conséquent, la probabilité d'une décision positive est :

$$\Pr(\delta_\alpha=1)=1-F(-X_{i\alpha}\beta_\alpha). \quad (4)$$

Si nous retenons l'hypothèse commode de symétrie pour la loi de l'erreur $u_{i\alpha}$ (cas usuel de la loi normale ou logistique), les probabilités des décisions du ménage peuvent s'écrire aussi comme suit :

$$\Pr(\delta_\alpha=1)=F(X_{i\alpha}\beta_\alpha) \text{ et } \Pr(\delta_\alpha=0)=1-F(X_{i\alpha}\beta_\alpha) \quad (5)$$

Le modèle de décisions séquentielles complet sera spécifié lorsqu'une hypothèse relative à la dépendance ou à l'indépendance des choix successifs du ménage aura été retenue. Il semble raisonnable ici, en l'absence d'informations objectives sur la qualité respective des différents types d'école, de supposer qu'une fois la décision de scolarisation prise, le choix du type d'école et la poursuite ou non à son terme par l'enfant de sa scolarité sont des choix effectués de façon indépendante par le chef de ménage. Cette hypothèse permet de calculer facilement la probabilité associée à chacune des sept branches de l'arbre de décisions du ménage. Ces branches seront codifiées à l'aide d'une variable discrète Y à valeurs dans l'ensemble des 7 premiers entiers naturels, de la façon suivante pour un enfant quelconque i du ménage :

$$Y_i=1 \text{ si } (\delta_1=1, \delta_2=1, \delta_3=1), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}>0, U^*_{i3}>0). \quad (6)$$

$$Y_i=2 \text{ si } (\delta_1=1, \delta_2=1, \delta_3=0), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}>0, U^*_{i3}\leq 0). \quad (7)$$

$$Y_i=3 \text{ si } (\delta_1=1, \delta_2=0, \delta_4=1, \delta_5=1), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}\leq 0, U^*_{i4}>0, U^*_{i5}>0). \quad (8)$$

$$Y_i=4 \text{ si } (\delta_1=1, \delta_2=0, \delta_4=1, \delta_5=0), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}\leq 0, U^*_{i4}>0, U^*_{i5}\leq 0). \quad (9)$$

$$Y_i=5 \text{ si } (\delta_1=1, \delta_2=0, \delta_4=0, \delta_6=1), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}\leq 0, U^*_{i4}\leq 0, U^*_{i6}>0). \quad (10)$$

$$Y_i=6 \text{ si } (\delta_1=1, \delta_2=0, \delta_4=1, \delta_6=0), \text{ c'est-à-dire si } (U^*_{i1}>0, U^*_{i2}\leq 0, U^*_{i4}\leq 0, U^*_{i6}\leq 0). \quad (11)$$

$$Y_i=7 \text{ si } (\delta_1=0), \text{ c'est à dire si } (U^*_{i1}\leq 0). \quad (12)$$

Les probabilités associées à chacun de ces 7 événements qui définissent les 7 branches de l'arbre de décisions du ménage s'écrivent donc, en vertu de l'hypothèse d'indépendance des décisions et lorsqu'on suppose de plus que la loi des erreurs $u_{i\alpha}$ est identique pour toutes les décisions :

Probabilité qu'un enfant scolarisé dans une école publique termine sa scolarité :

$$\Pr(Y_i=1)=F(X_{i1}\beta_1)\times F(X_{i2}\beta_2)\times F(X_{i3}\beta_3) \quad (13)$$

Probabilité qu'un enfant scolarisé dans une école publique ne termine pas sa scolarité :

$$\Pr(Y_i=2) = F(X_{i1}\beta_1) \times F(X_{i2}\beta_2) \times (1-F(X_{i3}\beta_3)) \quad (14)$$

Probabilité qu'un enfant scolarisé dans une école privée confessionnelle termine sa scolarité :

$$\Pr(Y_i=3) = F(X_{i1}\beta_1) \times (1-F(X_{i2}\beta_2)) \times F(X_{i4}\beta_4) \times F(X_{i5}\beta_5) \quad (15)$$

Probabilité qu'un enfant scolarisé dans une école privée confessionnelle ne termine pas sa scolarité :

$$\Pr(Y_i=4) = F(X_{i1}\beta_1) \times (1-F(X_{i2}\beta_2)) \times F(X_{i4}\beta_4) \times (1-F(X_{i5}\beta_5)) \quad (16)$$

Probabilité qu'un enfant scolarisé dans une école privée laïque termine sa scolarité :

$$\Pr(Y_i=5) = F(X_{i1}\beta_1) \times (1-F(X_{i2}\beta_2)) \times (1-F(X_{i4}\beta_4)) \times F(X_{i6}\beta_6) \quad (17)$$

Probabilité qu'un enfant scolarisé dans une école privée laïque ne termine pas sa scolarité :

$$\Pr(Y_i=6) = F(X_{i1}\beta_1) \times (1-F(X_{i2}\beta_2)) \times (1-F(X_{i4}\beta_4)) \times (1-F(X_{i6}\beta_6)) \quad (18)$$

Probabilité que l'enfant ne soit pas scolarisé :

$$\Pr(Y_i=7) = 1-F(X_{i1}\beta_1) \quad (19)$$

Si, comme nous l'avons supposé plus haut, les erreurs $u_{i\alpha}$ suivent toutes la loi normale de fonction de répartition $F=\Phi$, le modèle ainsi spécifié est un modèle probit séquentiel, qui généralise à n choix binaires successifs le modèle probit ordinaire caractérisé par un choix binaire unique.

Les sept branches de l'arbre de décision des ménages sont modélisées avec les mêmes variables explicatives dont les coefficients sont estimées en maximisant le logarithme de la vraisemblance du modèle (log-vraisemblance). Cette log-vraisemblance a pour expression :

$$\begin{aligned} \ln L(y, \beta) = & 6 \sum_{i=1}^{n_{11}} \ln[\Phi(X_{i1}\beta)] + \sum_{i=1}^{n_{10}} \ln[1 - \Phi(X_{i2}\beta)] + 2 \sum_{i=1}^{n_{21}} \ln[\Phi(X_{i2}\beta)] + 4 \sum_{i=1}^{n_{20}} \ln[1 - \Phi(X_{i2}\beta)] \\ & + \sum_{i=1}^{n_{31}} \ln[\Phi(X_{i3}\beta)] + \sum_{i=1}^{n_{30}} \ln[1 - \Phi(X_{i3}\beta)] + 2 \sum_{i=1}^{n_{41}} \ln[\Phi(X_{i4}\beta)] + 2 \sum_{i=1}^{n_{40}} \ln[1 - \Phi(X_{i4}\beta)] \\ & + \sum_{i=1}^{n_{51}} \ln[\Phi(X_{i5}\beta)] + \sum_{i=1}^{n_{50}} \ln[1 - \Phi(X_{i5}\beta)] + \sum_{i=1}^{n_{61}} \ln[\Phi(X_{i6}\beta)] + \sum_{i=1}^{n_{60}} \ln[1 - \Phi(X_{i6}\beta)] \end{aligned}$$

Dans cette formule, Φ désigne la fonction de répartition de la loi normale centrée réduite, et les n_{hk} , ($h=1$ à 7 et $k=0, 1$) représentent les effectifs des individus concernés par les 7 choix binaires successifs des ménages relativement à leurs enfants et vérifient : $n_{h0} + n_{h1} = n_h$, pour $h=1$ à 7.

La maximisation de $\text{Log}L(y, \beta)$ équivaut à la maximisation séparée des six composantes de cette fonction correspondant aux six choix binaires successifs du ménage, en raison de l'hypothèse d'indépendance retenue pour ces choix. Les équations de vraisemblance relatives à chacune de ces composantes sont obtenues en annulant les dérivées partielles de ces composantes par rapport aux variables explicatives, et sont non linéaires par rapport aux coefficients β .

IV. Variables, données et résultats d'estimation du modèle

4.1. Les facteurs explicatifs des décisions des ménages

Le modèle théorique décrivant les choix successifs des ménages dans le cadre de la scolarisation de leurs enfants retient comme arguments de la fonction d'utilité, la consommation et le capital humain, ce dernier étant défini comme une fonction linéaire des caractéristiques du ménage, de l'enfant, du système éducatif, etc. Il nous faut retenir de façon concrète les variables explicatives potentielles du capital humain, sur la base des résultats de l'enquête QUID.

Compte tenu de la continuité, voire de la concomitance des décisions prises par les ménages relativement à la scolarisation de leurs enfants, nous retiendrons pour simplifier les mêmes groupes de variables explicatives de base pour le capital humain. Ces groupes ne sont pas retenus sur une base purement théorique, mais en tenant compte des variables disponibles dans les résultats de l'enquête qui servira de base à l'estimation du modèle, à savoir l'enquête QUID 2001.

Indicateur de niveau de vie des ménages :

Dans la section 2, nous avons construit un indicateur de niveau de vie qui n'est autre que la première composante principale de l'ACP. Il semble donc naturel de retenir cette variable synthétique dont la pertinence pour expliquer leurs décisions dans le cadre de la scolarisation de leurs enfants a été prouvée. En outre, elle permet de résoudre le problème d'endogénéité que poserait l'utilisation de la consommation comme indicateur de niveau de vie.

Caractéristiques démographiques du ménage :

- Le sexe du chef de ménage : l'importance de la scolarisation des enfants n'est peut-être pas perçue avec la même acuité par un chef de ménage selon qu'il est de sexe masculin ou féminin. Le poids de la tradition n'a peut-être pas le même impact sur les choix effectués par un homme ou une femme en matière de scolarisation des enfants.
- L'âge du chef de ménage : un chef de ménage jeune est certainement plus ouvert à l'école qu'un chef de ménage âgé.
- La taille du ménage : le nombre d'individus du ménage peut être un facteur défavorable à la scolarisation des enfants, notamment dans les ménages de condition modeste qui ont souvent le plus d'enfants. Ce facteur peut s'avérer décisif en zones rurales où le coût d'opportunité de la scolarisation d'un enfant peut être élevé, en raison de l'importance du travail des enfants.

Le niveau d'éducation du chef de ménage

Il est appréhendé à l'aide de deux variables : le fait que le chef de ménage soit alphabétisé ou non et son niveau d'éducation (classe achevée). Ces variables explicatives semblent pertinentes a priori, puisqu'un chef de ménage alphabétisé ou éduqué admet plus facilement l'utilité de l'éducation qu'un chef de ménage analphabète.

Disponibilité de l'offre d'éducation

Elle est mesurée par une variable indicatrice de la proximité d'une école primaire, qui prend la valeur 1 lorsque l'école la plus proche de la résidence de l'enfant est à moins de 30 minutes de trajet.

Sexe, zone, région de résidence et opinion de l'enfant sur le système éducatif

Chacune de ces trois premières variables semble pertinente pour fonder les choix des ménages, comme l'a montré la section 2. Il en est de même de la quatrième variable qui explique a priori le degré de motivation de l'enfant à poursuivre sa scolarité à terme.

4.2. Etude de l'échantillon jugé pertinent et tirage d'un sous échantillon approprié :

Les données utilisées dans cette étude sont tirées des fichiers de l'enquête nationale QUID qui comporte des volets relatifs à l'éducation, la santé, l'emploi, aux avoirs des ménages et aux enfants. Cette enquête porte sur 60193 individus, membres de 6613 ménages regroupés en 994 grappes de 12 ménages environ chacune.

De ce fichier sont extraites les informations pertinentes pour l'estimation du modèle. Nous avons décidé de ne considérer que les enfants âgés de 6 à 19 ans. L'effectif de cette sous population est de 16672 individus. Ce sous fichier a été tiré du fichier global obtenu en appariant le fichier des ménages et celui des personnes, de façon à affecter à ces dernières les caractéristiques des ménages auxquels elles appartiennent. Une fois obtenu l'échantillon des 16672 enfants, nous avons extrait les variables qui nous intéressent et construit celles qui ne sont pas immédiatement disponibles.

Il est indispensable de décrire cet échantillon à l'aide de la variable dépendante Y qui codifie la séquence des décisions prises par le chef de ménage et des variables explicatives retenues dans notre modèle. La distribution de la variable Y dans la population des enfants scolarisables est donnée dans le tableau 7.

Tableau 7: Distribution des enfants scolarisables selon leur situation scolaire

Situation de l'enfant	Effectif	Fréquence en %
A terminé l'école publique (Y=1)	1115	6,7
N'a pas terminé l'école publique (Y=2)	5109	30,7
A terminé l'école privée confessionnelle (Y=3)	103	0,6
N'a pas terminé l'école privée confessionnelle (Y=4)	289	1,7
A terminé l'école privée laïque (Y=5)	283	1,7
N'a pas terminé l'école privée laïque (Y=6)	425	2,5
N'est pas scolarisé (Y=7)	9348	56,1
Total	16672	100,0

La première remarque concerne évidemment le pourcentage élevé d'enfants non scolarisés (56,12%). La seconde est liée à l'importance relative de la proportion d'enfants inscrits dans des écoles publiques et de ceux qui ne l'ont pas terminé. Il convient à ce propos de signaler que toutes les modalités paires de Y, qui codifient la situation des enfants n'ayant pas terminé leur scolarité, se réfèrent à des individus toujours en cours de scolarisation. Compte tenu de la grande taille de l'échantillon des individus pertinents pour l'estimation du modèle, il semble opportun d'en extraire un sous-échantillon de taille raisonnable. Nous avons opté pour le tirage d'un sous-échantillon de 1500 enfants environ, soit 9% de l'échantillon initial.

Toutefois, le faible effectif de certaines modalités de la variable expliquée (tableau 7) nous oblige à imposer des quotas pour garantir la présence en nombre suffisant des individus présentant ces modalités.

La représentativité de ce sous-échantillon doit être étudiée en comparant les résumés statistiques des variables explicatives du modèle dans l'échantillon parent et le sous-échantillon. Ces résumés sont consignés dans les tableaux 8 et 9.

Tableau 8 : Résumés statistiques des variables retenues pour l'échantillon initial d'enfants.

VARIABLES	Effectif	Minimum	Maximum	Moyenne	Ecart type
Sexe de l'enfant(1=féminin)	16672	0	1	0,51	0,500
Age de l'enfant	16672	6	19	11,59	3,894
Taille du ménage	16669	2	35	12,33	5,766
Sexe du chef de ménage (cm) (1=féminin)	16659	0	1	0,16	0,363
Age du cm	16659	16	99	51,698	13,082
* Enfant satisfait de l'enseignement (1=oui)	16672	0	1	0,24	0,426
* Alphabétisation du cm	16654	0	1	0,39	0,488
* Classe achevée par le cm	16672	0	18	2,464	4,812
* Région de résidence (1=Dakar)	16672	0	1	0,23	0,422
Indicateur de niveau de vie	16668	-3,395	3,620	-0,093	0,941
* Zone de résidence (1=urbaine)	16672	0	1	0,42	0,493
Proximité d'une école primaire (1= école primaire à moins de 30 mn)	16560	0	1	0,79	0,406

Tableau 9 : Résumés statistiques des variables retenues pour le sous-échantillon de 1500 enfants.

VARIABLES	Effectif	Minimum	Maximum	Moyenne	Ecart type
Sexe de l'enfant(1=féminin, 0=masculin)	1500	0	1	0,51	0,500
Age de l'enfant	1500	6	19	12,06	3,948
Taille du ménage	1499	2	32	11,48	5,418
Sexe du chef de ménage (cm) (1=féminin)	1498	0	1	0,20	0,399
Age du cm	1498	18	98	51,18	12,261
Enfant satisfait de l'enseignement (1=oui)	1500	0	1	0,48	0,500
Alphabétisation du cm (1=oui)	1498	0	1	0,54	0,498
Classe achevée par le cm	1500	0	18	5,09	6,477
Région de résidence (1=Dakar)	1500	0	1	0,45	0,497
Indicateur de niveau de vie	1499	-2,467	3,620	0,451	1,199
Zone de résidence (1=urbaine)	1500	0	1	0,62	0,485
Proximité d'une école primaire (1 si école primaire la plus proche à moins de 30 mn)	1488	0	1	0,84	0,370

Ces tableaux révèlent une grande similitude entre les deux échantillons pour les résumés de certaines variables (moyenne et écart-type de la taille du ménage, de l'âge du chef de ménage, de l'âge de l'enfant, proportions de garçons et de filles, proportion d'enfants résidant à proximité d'une école), mais aussi une certaine différence relativement à d'autres variables (alphabétisation du chef de ménage, classe achevée par le chef de ménage, région et zone de résidence, indicateur de richesse). Mais les différences constatées, inévitables dans tout problème d'échantillonnage, sont d'un ordre de grandeur suffisamment faible pour autoriser l'utilisation du sous-échantillon de 1500 individus pour estimer le modèle de décisions séquentielles des ménages.

V. Résultats d'estimation du modèle

5.1 Procédure d'estimation, significativité et signe des coefficients estimés

L'estimation du modèle peut se faire à l'aide de divers algorithmes. Nous avons utilisé l'algorithme de Broyden-Fletcher-Goldfarb-Shanno programmée dans le logiciel économétrique RATS.

Une fois le modèle estimé, sa pertinence globale peut être évaluée en le comparant au plus simple modèle concevable n'ayant comme variables « explicatives » que les termes constants. Ce test de la significativité globale des véritables variables explicatives repose sur le calcul de la statistique du rapport entre vraisemblance du modèle sans variables explicatives autres que les constantes et celui du modèle incorporant les variables explicatives, qui s'écrit :

$LR = -2[\text{Log}L(y, \beta_0) - \text{Log}L(y, \beta)]$, où $L(y, \beta_0)$ est la vraisemblance du modèle n'ayant que les constantes β_0 comme variables explicatives, et $L(y, \beta)$ est la vraisemblance du modèle incorporant les variables explicatives et les constantes. Sous l'hypothèse nulle que les variables explicatives retenues sont globalement non pertinentes, LR suit une loi du KHI-DEUX dont le degré de liberté est égal au nombre total de coefficients estimés du modèle moins le nombre de constantes. Les variables explicatives retenues seront globalement significatives si la statistique LR du rapport des vraisemblances est supérieure à la valeur critique du KHI-DEUX, pour un risque α donné.

En raison du caractère non linéaire, l'interprétation d'un modèle probit séquentiel, l'interprétation de ses coefficients, comme celle des coefficients d'un modèle probit ordinaire, est loin d'être aisée. Les valeurs des coefficients estimés ne mesurent pas les impacts marginaux directs des variables explicatives sur la variable dépendante (ou plus exactement sur son espérance mathématique), comme dans la régression linéaire classique. On se contente donc de noter le signe et la significativité des coefficients estimés qui mesurent ici l'impact marginal des variables explicatives sur une fonction de l'espérance mathématique de la variable dépendante, fonction n'ayant aucune signification intuitive. L'interprétation approfondie des résultats d'estimation s'appuie plutôt sur le calcul de la distribution de probabilités de Y pour un individu de référence, par exemple l'individu moyen de l'échantillon. On peut aussi déterminer l'impact marginal des différentes variables explicatives sur cette distribution au point moyen. On fait varier alors d'une unité la variable concernée, les autres restant constantes, et on prend la différence des probabilités obtenues. Lorsque la variable explicative concernée est continue, cet impact marginal est tout simplement la dérivée partielle de la fonction de vraisemblance par rapport à cette variable.

Les résultats d'estimation de notre modèle séquentiel sont donnés dans le tableau 10. Il importe de noter la significativité globale des variables explicatives retenues dans le modèle, en raison de la forte valeur de la statistique LR du rapport des vraisemblances (2380), largement supérieure à la valeur critique, au seuil de 1%, d'une loi du Khi-deux à 72 degrés de liberté.

Tableau 10 : Résultats d'estimation du modèle de décisions séquentielles de solarisation des enfants

	Décision de scolarisation	Ecole publique ou privée	Ecole publique terminée ou non	Ecole privée confessionnelle ou laïque.	Ecole confessionnelle terminée ou non	Ecole laïque terminée ou non
	Coefficient (écart type)	Coefficient (écart type)	Coefficient (écart type)	Coefficient (écart type)	Coefficient (écart type)	Coefficient (écart type)
Taille (du ménage)	-0,0234 (0,0066)***	0,0427 (0,0071)***	-0,0004 (0,0114)	-0,0446 (0,0108)***	-0,0108 (0,021)	-0,0244 (0,0186)
Sexe du chef de ménage (cm) : 1 si féminin	0,4711 (0,1046)***	0,0067 (0,0891)	-0,1117 (0,1722)	0,3044 (0,1206)**	0,3320 (0,2319)	0,4332 (0,2097)**
Age du cm	0,0065 (0,0030)**	-0,0001 (0,0020)	-0,0052 (0,0067)	0,0108 (0,0025)***	0,0190 (0,0083)**	0,0256 (0,0085)***
Le cm est alphabétisé : 1 si oui.	-0,2585 (0,0752)***	-0,0941 (0,0830)	0,8280 (0,1449)***	-0,0333 (0,1010)	0,0311 (0,2768)	1,1262 (0,2913)***
Classe achevée par le cm	0,0962 (0,0085)***	-0,0251 (0,0053)***	-0,0015 (0,0146)	0,0371 (0,0072)***	0,0569 (0,0114)***	0,0200 (0,0186)
Indicateur de niveau de vie (ACP)	0,1433 (0,0468)***	-0,1839 (0,0464)***	0,1861 (0,1092)*	0,3176 (0,0538)***	0,1790 (0,0842)**	-0,0634 (0,0818)
Région de résidence : 1 si Dakar	0,2440 (0,1084)**	-0,7461 (0,0894)***	-0,3555 (0,1628)**	-0,9759 (0,1294)***	-0,4704 (0,2246)**	-0,4192 (0,2492)*
Proximité de l'école : 1 si école à moins de 30 mn	0,5919 (0,0984)***	0,3086 (0,0899)***	-0,0250 (0,1969)	-0,0828 (0,1363)	-0,6715 (0,3268)**	0,8801 (0,3254)***
Zone de résidence : 1 si urbaine	0,1710 (0,1156)	-0,3062 (0,0860)***	0,7277 (0,1474)***	-1,1847 (0,1603)***	1,0407 (0,2356)***	-1,0562 (0,5236)**
Sexe de l'enfant (1 si féminin)	-0,2557 (0,0788)***	0,1863 (0,0634)***	-0,4997 (0,1376)***	0,0453 (0,09010)	0,5455 (0,2335)**	-0,2037 (0,1462)
Age de l'enfant	0,0116 (0,0065)*	0,0413 (0,0063)***	0,5290 (0,0392)***	-0,0006 (0,0092)	0,5879 (0,0591)***	0,5319 (0,0433)***
Enfant satisfait de l'enseignement: 1 si oui	6,4499 (0,2398)***	-0,1967 (0,0841)**	-0,2251 (0,1417)	0,2114 (0,1147)*	0,0068 (0,1732)	0,1661 (0,1350)
Constante	-1,0301 (0,2710)***	-0,3739 (0,1430)***	-7,7607 (0,8228)***	0,5806 (0,1365)***	-9,6308 (1,3745)***	-8,8247 (0,9981)***

Notes :

* = significatif au seuil de 10% ; ** = significatif au seuil de 5% ; *** = significatif au seuil de 1%.

Nombre total d'observations = 1500 ; nombre d'observations utiles = 1487 ; degré de liberté = 1409

Valeur maximale de la log-vraisemblance = -1530,9

Valeur de la statistique LR du rapport des vraisemblances = 2380 ; degrés de liberté du Khi-deux = 72.

La majorité des variables proposées expliquent significativement au seuil de 1 ou 5 % le choix des ménages entre la scolarisation et la non scolarisation de leurs enfants. La seule variable non pertinente est la zone de résidence de l'enfant. Quant à l'âge de l'enfant, il n'est significatif qu'au seuil de 10%. En ce qui concerne le signe de l'impact des différentes variables significatives sur la décision de scolarisation ou non des ménages, on observe qu'il est négatif pour la taille du ménage, l'alphabétisation du chef de ménage et le sexe de l'enfant lorsqu'il est une fille, alors qu'il est positif pour le sexe du chef de ménage, la classe achevée par celui-ci, l'indicateur de niveau de vie, la région de résidence, la disponibilité de l'offre (proximité de l'école), l'âge du chef de ménage et celui de l'enfant.

Neuf variables sur les 12 proposées expliquent significativement le choix des ménages entre l'école privée et l'école publique. Les variables affectant positivement le choix de l'école publique plutôt que l'école privée sont la taille du ménage, la proximité de l'école, le sexe et l'âge de l'enfant. Interviennent négativement la classe achevée par le chef de ménage, l'indicateur de niveau de vie, la région ainsi que la zone de résidence, et l'opinion de l'enfant sur l'école.

Cinq des variables retenues expliquent très significativement la situation de l'enfant en matière de poursuite à son terme ou non de la scolarisation d'un enfant dans une école publique. Le sexe de l'enfant, l'âge du chef de ménage et la région de résidence exercent une influence négative sur la décision du ménage, tandis que l'alphabétisation du chef de ménage et la zone de résidence déterminent positivement et très significativement cette décision. L'indicateur de niveau de vie influence moyennement (seuil de 10%) et positivement cette décision.

Les variables suivantes exercent une influence négative (significative à 5% au moins) sur le choix entre école privée confessionnelle et laïque : la taille du ménage, la région et la zone de résidence. Celles qui interviennent positivement et très significativement dans ce choix sont l'âge du chef de ménage, la classe achevée par lui, et l'indicateur de niveau de vie. Quant à l'opinion de l'enfant sur l'école, elle a un impact positif au seuil de 10% seulement sur la décision du ménage.

Six variables interviennent positivement et significativement au seuil de 5% au moins dans l'achèvement de la scolarisation dans l'école confessionnelle. Il s'agit de l'âge du chef de ménage et de l'enfant, de l'alphabétisation et de la classe achevée par le chef de ménage, de l'indicateur de niveau de vie, de la zone de résidence, et du sexe de l'enfant. En revanche, seules la région de résidence et la proximité de l'école exercent un impact négatif significatif sur ce résultat scolaire.

Enfin, l'achèvement ou non de la scolarisation primaire dans une école laïque dépend positivement et significativement de son âge et de celui du chef de ménage, du sexe de ce dernier, du fait qu'il est alphabétisé ou non, et de la proximité de l'école, et négativement de la région et de la zone de résidence et du sexe de l'enfant.

5.2 Distribution de probabilité de la variable expliquée pour quelques individus de référence et détermination de l'impact marginal des variables explicatives

L'individu-type du sous-échantillon de 1500 enfants retenu est celui dont les coordonnées sont les valeurs moyennes pour les variables explicatives quantitatives, et les codes les plus fréquents pour les variables explicatives qualitatives. En nous reportant au tableau 9, nous remarquons que cet individu de référence est une fille de 12 ans habitant une zone urbaine, en dehors de la région de Dakar, à proximité d'une école, dans un ménage de 11 personnes, dont le chef est un homme âgé de 51 ans, alphabétisé, ayant le niveau de la 4^{ème} année d'éducation primaire (CE2), dont l'indicateur de niveau de vie est égal à 0,451 (4^{ème} intervalle inter-quintile). Cet individu-type est par ailleurs non satisfait de l'école. Nous commençons par la détermination de la distribution de probabilité de Y pour l'individu type et pour les individus qui ne diffèrent de lui que par le sexe et la zone de résidence. On peut évaluer ainsi l'impact de ces deux variables sur les probabilités des modalités de Y. Les résultats de ces calculs, obtenus à partir des formules 14 à 20, sont donnés dans le tableau 11.

Tableau 11 : Distribution de probabilités de Y pour l'individu-type et les individus qui ne diffèrent de lui que par le sexe et la zone de résidence

Probabilités de Y	Individu-type (fille d'une zone urbaine)	Fille d'une zone rurale	Garçon d'une zone urbaine	Garçon d'une zone rurale
Pr(Y=1)	0,1285	0,0453	0,2040	0,0990
Pr(Y=2)	0,4797	0,6740	0,3312	0,5544
Pr(Y=3)	0,0624	0,0178	0,0354	0,0068
Pr(Y=4)	0,1208	0,2250	0,1736	0,2895
Pr(Y=5)	0,0702	0,0279	0,1058	0,0401
Pr(Y=6)	0,1383	0,0099	0,1499	0,0101
Pr(Y=7)	0,0001	0,0001	0,0001	0,0001

Ce tableau montre que les perspectives, en termes de probabilité de terminer sa scolarité ou d'être encore à l'école primaire, sont meilleures en zones urbaines qu'en zones rurales, quel que soit le sexe de l'enfant. En effet, les probabilités d'avoir terminé sa scolarité sont plus élevées pour les garçons et les filles des zones urbaines que pour ceux et celles de même sexe résidant en zones rurales, tandis que les probabilités d'être encore à l'école primaire sont moins fortes pour les enfants des zones urbaines pour tous les types d'école, exceptée l'école privée laïque. Ceci signifie que les chances de terminer sa scolarité primaire dans un délai « normal » sont plus grandes en milieu urbain qu'en milieu rural. On remarque cependant que ces chances sont moins grandes pour les filles que pour les garçons.

Intéressons-nous maintenant à l'impact marginal des variables explicatives sur les probabilités des choix successifs effectués par les ménages. Il est égal à la dérivée partielle de $F(X_{i\alpha}, \beta_{\alpha})$ par rapport aux éléments de $X_{i\alpha}$. Concrètement il est évalué évalué ici en faisant la différence, pour la décision α , entre la valeur de $F(X_{i\alpha}, \beta_{\alpha})$ associée à l'individu type et la valeur de cette fonction lorsqu'on fait varier d'une unité par rapport à l'individu-type une variable explicative donnée, les autres restant constantes. Les résultats de ce calcul pour les variables explicatives quantitatives proposées sont donnés dans le tableau 12. L'examen de ce tableau permet de faire les remarques suivantes :

1. L'impact marginal de la taille du ménage ne diffère significativement de 0 que pour le choix du type d'école (école publique et école privée confessionnelle). L'augmentation d'une unité de cette variable se traduit par une augmentation de 1,6% de la probabilité que le ménage préfère inscrire son enfant à l'école publique, et par une diminution de 1,8% de la probabilité que l'enfant soit inscrit dans une école privée confessionnelle, toutes choses restant égales par ailleurs. Mais elle a un effet négatif infime sur la probabilité que l'enfant termine (ou ait terminé) sa scolarité.

Tableau 12 : Impact marginal de quelques variables explicatives sur les choix successifs des ménages dans le cadre de la scolarisation de leurs enfants

Variables explicatives	Variation de Prob($\delta_1=1$). (Scolarisation)	Variation de Prob($\delta_2=1$). (Ecole publique)	Variation de Prob($\delta_3=1$). (Terminer l'école publique)	Variation de Prob($\delta_4=1$). (Ecole privée confess.-)	Variation de Prob($\delta_5=1$). (Terminer l'école confess.)	Variation de Prob($\delta_6=1$). (Terminer l'école laïque)
Taille du ménage	0	0,016	-0,0001	-0,0177	-0,0039	-0,0088
Age du chef de ménage	0	-0,00005	-0,0015	0,0043	0,007	0,0094
Classe achevée par le chef de ménage	0	-0,0097	-0,0004	0,0148	0,0208	0,0073
Indicateur de niveau de vie	0	-0,072	0,0577	0,1258	0,0677	-0,0228
Age de l'enfant	0	0,0158	0,1812	-0,0002	0,2297	0,2073

NB : Le calcul n'est effectué que pour les variables explicatives des différents choix significatives au risque de 10 % au moins.

2. L'âge du chef de ménage n'influence marginalement de façon significative aucune des décisions du ménage.

3. La classe achevée par le chef de ménage n'influence marginalement de façon significative que la probabilité de préférer pour l'enfant l'école privée confessionnelle.

4. Quant à l'indicateur de niveau de vie, sa variation d'une unité, signifie concrètement le passage du quatrième au cinquième quintile. Elle se traduit, toutes choses étant égales par ailleurs, par une baisse de la probabilité de choisir l'école publique et de terminer l'école privée laïque, et par une augmentation de la probabilité de terminer l'école publique ou privée confessionnelle, et de préférer l'école privée confessionnelle.

5. Enfin, on note l'influence marginale fortement positive de l'âge de l'enfant sur la probabilité qu'il ait terminé sa scolarité.

Les résultats ci-dessus concernent l'individu moyen qui appartient au quatrième quintile. Il importe de s'interroger maintenant sur la pertinence de ces résultats pour les individus des quintiles inférieurs. Nous nous contenterons d'examiner la question en modifiant l'indicateur de niveau de vie de l'individu-type de façon à le faire correspondre à celui de l'individu moyen du premier quintile, et en recalculant les quatre

distributions de probabilité de Y données dans le tableau 12. Les nouvelles distributions de probabilité de Y obtenues sont consignées dans le tableau 13.

Tableau 13 : Distribution de probabilités de Y pour l'individu type modifié (indicateur de niveau de vie égal à la moyenne de cet indicateur pour le premier inter-quintile, soit -1,003) et les individus qui ne diffèrent de lui que par le sexe et la zone de résidence.

Probabilités de Y	Individu-type modifié (fille d'une zone urbaine)	Fille d'une zone rurale	Garçon d'une zone urbaine	Garçon d'une zone rurale
Pr(Y=1)	0,1000	0,0288	0,1811	0,0721
Pr(Y=2)	0,6060	0,7731	0,4579	0,6738
Pr(Y=3)	0,0217	0,0064	0,0112	0,0022
Pr(Y=4)	0,0646	0,1402	0,0892	0,1819
Pr(Y=5)	0,0770	0,0395	0,1173	0,0576
Pr(Y=6)	0,1306	0,0120	0,1433	0,0123
Pr(Y=7)	0,0001	0,0000	0,0000	0,0001

Le commentaire du tableau 13 est identique à celui du tableau 11. Pour un enfant appartenant au groupe des 20% les plus pauvres, la zone de résidence et le sexe de l'enfant constituent un facteur déterminant de l'issue de sa scolarité. La différence avec les enfants du quatrième quintile réside dans le caractère un peu moins accentué de cette double discrimination pour le groupe le plus pauvre par rapport au groupe plus aisé du quatrième quintile.

VI. Conclusion

Cette étude relative à la demande d'éducation a eu recours à deux approches différentes mais complémentaires : une analyse descriptive des taux de scolarisation observés selon les régions, le niveau de vie des ménages et le sexe des enfants, et la construction d'un modèle de décisions séquentielles décrivant les choix successifs des ménages dans le cadre de la scolarisation de leurs enfants.

La partie descriptive s'est employée à construire un indicateur synthétique du niveau de vie des ménages fondé sur l'ACP, et à déterminer les quintiles à l'aide de cet indicateur. Les taux de scolarisation dans l'enseignement primaire et moyen, ont ensuite été calculés pour les différentes régions et pour les groupes économiques extrêmes. Il apparaît nettement que le taux de scolarisation est lié au niveau de pauvreté : il est beaucoup plus élevé dans le groupe des 20% les plus riches que dans celui des 40% les plus pauvres, dans toutes les régions. On enregistre un écart positif de 49% entre ces deux groupes extrêmes pour l'enseignement primaire dans l'ensemble du Sénégal. Mais cet écart n'augmente pas pour le niveau inférieur de l'enseignement secondaire. Une politique visant à accélérer la scolarisation des enfants issus des ménages contribuerait à réduire fortement les inégalités dans l'accès à l'éducation secondaire et supérieure.

La deuxième partie propose de modéliser les choix successifs des ménages dans le cadre de la scolarisation de leurs enfants. Le modèle probit séquentiel estimé a permis de mettre en évidence la pertinence d'une majorité des 12 variables proposées pour expliquer l'arbre de décisions des ménages. Les caractéristiques de l'enfant semblent particulièrement déterminantes dans ces choix, de même que certaines caractéristiques du ménage, dont l'indicateur de niveau de vie. Pour ce qui est du premier choix des ménages, à savoir le choix de scolariser ou non un enfant, il dépend significativement de 11 variables

sur les 12, la zone de résidence étant la seule variable non pertinente de ce choix. En particulier, la taille du ménage et le sexe de l'enfant ont un impact négatif sur ce choix (autrement dit les chances de scolarisation d'une fille appartenant à un ménage comprenant un grand nombre d'individus sont faibles), tandis que l'indicateur de niveau de vie, le niveau d'éducation du chef de ménage, la proximité de l'école et le sexe du chef de ménage, notamment, influencent positivement ce choix.

Le calcul de la distribution de probabilité de la variable de décision pour l'individu-type et pour les individus qui ne diffèrent de lui que par le sexe ou la zone de résidence, nous a permis de confirmer la discrimination subie par les enfants ruraux, quel que soit leur sexe, par rapport aux enfants des zones urbaines, et celle subie par les filles par rapport aux garçons, quelle que soit leur zone de résidence.

La lourdeur des calculs nous a obligés à limiter les simulations, mais ces quelques résultats témoignent de l'intérêt du modèle estimé, qui complète l'analyse descriptive effectuée dans la première partie.

Les résultats obtenus mettent en évidence la nécessité de rendre l'offre d'éducation plus accessible en construisant des écoles supplémentaires dans les différents quartiers et villages, la nécessité de généraliser l'éducation des adultes de manière à les sensibiliser d'avantage à l'importance de l'éducation de leurs enfants, garçons et filles, et l'urgence de la mise en place des programmes de lutte contre la pauvreté.

Bibliographie

- Alderman, H., Behrman, J., Ross, D. and R. Sabot, (1996), Decomposing the Gender Gap in Cognitive Achievement in a Poor Rural Economy, *Journal of Human Resources*, 31, 229-254.
- Amemiya, T. (1975), Qualitative Response Models, *Annals of Economic and Social Measurement*, 4 (3), 363-72.
- Becker, G., (1975), *Human Capital*, second edit., NBER, New York, NY.
- Filmer, D. and L. Pritchett, (1998), *Educational Enrollment and Attainment in India : Household Wealth, Gender, Village, and State Effects*. The World Bank Working Paper.
- Gouriéroux, C., (1989), *Econométrie des variables qualitatives*, Economica.
- Gertler, P. and P. Glewwe, (1990), The Willingness to pay for Education in Developing Countries : Evidence from Rural Peru, *Journal of Public Economics*, 42, 251-275.
- Handa, S., (2002), Raising Primary School Enrollment in Developing Countries : The Relative Importance of Supply and Demand, *Journal of Development Economics*, 69, 103-128.
- Liao, T. F. (1994), *Interpreting Probability models : Logit, Probit, and Other Generalized Linear Models* (Sage University Paper series on Quantitative Applications in the Social Sciences).
- Rivers, D., Vuong, Q.(1988), Limited Information Estimators and Exogeneity tests for Simultaneous Probit Models, *Journal of Econometrics*, 39 (3), 347-366.
- Sawada, Y. and M. Loshkin, (2001), *Household Schooling Decisions in Rural Pakistan*, The World Bank Working Paper.
- Skyt Nielsen, H. ,(2001), How Sensitive is the Demand for Primary Education to Changes in Economic Factors?, *Journal of African Economies*, 10, 191-218.