



**CONSORTIUM POUR LA RECHERCHE  
ÉCONOMIQUE ET SOCIALE**



**Université Cheikh Anta Diop de Dakar  
FACULTÉ DES SCIENCES ÉCONOMIQUES ET DE GESTION  
Laboratoire d'Analyse des Politiques Publiques (LAPP)**

**SÉRIE DE DOCUMENTS DE RECHERCHE**

# **Déterminants des apprentissages dans l'éducation primaire au Sénégal**

Abdoulaye DIAGNE

**2007 / 01**

Consortium pour la Recherche Économique et Sociale  
Université Cheikh Anta Diop de Dakar  
Faculté des Sciences Économiques et de Gestion  
Laboratoire d'Analyse des Politiques Publiques (LAPP)

# Déterminants des apprentissages dans l'éducation primaire au Sénégal

Abdoulaye DIAGNE\*

Cette recherche a été réalisée grâce à l'appui de l'United States Agency  
for international Development (USAID) et le Centre de Recherches pour  
le Développement International (CRDI) dans le cadre  
du programme de recherche SAGA

---

\* [cres\\_ucad@yahoo.fr](mailto:cres_ucad@yahoo.fr)

Consortium pour la Recherche Économique et Sociale  
Rue de Kaolack x Rue F, Tour de l'Œuf – Point E  
En face de la Piscine Olympique - Dakar  
Tél. (221) 33 864 77 57 - Fax. (221) 33 864 77 58  
C.P : 12023 - BP 7988, Dakar-Médina  
E-mail : [cres@ucad.sn](mailto:cres@ucad.sn) / [cres\\_ucad@yahoo.fr](mailto:cres_ucad@yahoo.fr)  
Site web : [www.cres-ucad.org](http://www.cres-ucad.org)

Nous remercions Ismaël KAFANDO et Moussa H. OUNTENI  
pour leur précieuse assistance

## SOMMAIRE

|   |    |
|---|----|
| <i>Introduction</i> .....   | 5  |
| <b>1. Un cadre pour l'analyse des déterminants des apprentissages</b> .....             | 7  |
| <b>2. La qualité dans l'enseignement primaire : données et tendances majeures</b> ..... | 9  |
| 2.1 <i>Données</i> .....  | 9  |
| 2.2 <i>Principales tendances des performances scolaires des élèves</i> .....            | 12 |
| <b>3. Modèles et variables</b> .....  | 16 |
| 3.1 <i>Modèles</i> .....  | 16 |
| 3.2 <i>Variables</i> .....  | 21 |
| 3.2.1 <i>Caractéristiques de l'élève et du ménage auquel il appartient</i> .....        | 21 |
| 3.2.2 <i>Caractéristiques de l'offre éducative</i> .....                                | 23 |
| 3.2.3 <i>Caractéristiques de la communauté</i> .....                                    | 26 |
| <b>4. Résultats</b> .....   | 26 |
| 4.1 <i>Score moyen des deux disciplines</i> .....                                       | 29 |
| 4.2 <i>Score en français et en mathématiques</i> .....                                  | 37 |
| 4.2.1 <i>Les déterminants des apprentissages en français</i> .....                      | 37 |
| 4.2.2 <i>Les déterminants des apprentissages en maths</i> .....                         | 38 |
| <i>Conclusion</i> .....   | 38 |
| <i>Références bibliographiques</i> .....  | 40 |
| <i>Annexes</i> .....  | 42 |



## Introduction

Du début des années 80 au milieu des années 90, le système éducatif du Sénégal fut marqué par une mauvaise allocation des dépenses publiques entre les niveaux d'enseignement et une faible efficacité interne. Les taux bruts de scolarisation ont baissé dans l'enseignement primaire pendant cette période. Les différentes évaluations des acquisitions des élèves laissent apparaître un faible niveau de l'éducation primaire. Réagissant à ces contre-performances, le Gouvernement a élaboré un Programme décennal de l'éducation et de la formation (PDEF) qui a fait de la qualité des apprentissages son deuxième objectif majeur, après l'élargissement de l'accès visant notamment la scolarisation universelle à l'horizon 2010.

Des tests de connaissances acquises par les élèves du cycle primaire ont été organisés dans les années 90 et 2000 par la Conférence des ministres de l'éducation ayant le français en partage (CONFEMEN). Cet organisme qui regroupe des pays africains francophones, a conduit une enquête sur une cohorte d'élèves de 1996 à 2000 dans cinq pays membres (Burkina Faso, Cameroun, Côte d'Ivoire, Madagascar, Sénégal). Par ailleurs, en 1999, le Sénégal a participé au *Monitoring Learning Achievement* (MLA) conduit par l'UNESCO dans 18 pays dont la plupart sont africains. Enfin, le ministère de l'Éducation a procédé à sa propre évaluation dans le cadre du Système national d'évaluation des rendements scolaires (SNERS) en 2002.

Cependant, ces tests mettent l'accent sur l'évaluation des connaissances et accordent peu d'attention aux facteurs susceptibles d'expliquer les performances à ces tests. Or, il est important de combiner les tests de connaissances avec une collecte d'informations sur une gamme de variables permettant de bien comprendre les scores des élèves. C'est en mesurant les niveaux actuels d'acquisition de connaissances, et en identifiant les obstacles à la progression des élèves qu'il sera possible de connaître les facteurs sur lesquels agir pour relever le niveau des acquisitions des élèves.

Ce travail exploite les résultats des scores aux tests d'une cohorte d'élèves suivis pendant cinq ans (1996-2000) dans l'enseignement primaire au Sénégal dans le cadre du PASEC<sup>1</sup>. Les données de PASEC ont déjà été exploitées par Michaelowa (2001). Cependant, son travail diffère de celui-ci sur de nombreux points. D'abord, elle utilise un modèle hiérarchique statique à trois niveaux (élèves, écoles et pays), et seuls les scores de la cinquième année d'études sont étudiés.

Cette étude porte exclusivement sur le Sénégal. Elle a été menée à l'aide d'un modèle de croissance à trois niveaux (les scores aux tests, l'élève et son ménage, l'école et la communauté dans laquelle elle est implantée). Ainsi, les scores obtenus par chaque élève au cours des tests subis seront pris en compte. Les moyennes construites sur cinq observations reflètent mieux le niveau de la qualité de l'éducation que l'une de ces observations. Des variables aussi importantes que le niveau de vie des ménages auxquels appartiennent les élèves, le niveau d'éducation des parents et les caractéristiques de la communauté (dotation en infrastructures socio-économiques, importance du capital social, etc.) sont absentes de la base de données PASEC. Leur non prise en compte peut biaiser fortement les résultats des estimations faites avec les seules variables disponibles de la base de données PASEC. Ces lacunes ont été comblées par des enquêtes complémentaires menées en 2003. La modélisation linéaire hiérarchique à trois niveaux est utilisée pour analyser le niveau et la croissance des acquisitions scolaires des élèves (en français et en mathématiques) dans l'éducation primaire au Sénégal.

L'étude propose à la section 1 une brève revue des travaux consacrés aux facteurs qui influent sur la qualité de l'éducation primaire. Elle présente à la section 2 les sources des données utilisées ainsi que les tendances majeures des apprentissages dans l'éducation primaire sénégalaise à travers les résultats de l'enquête PASEC.

---

<sup>1</sup> Les données de PASEC ont déjà été exploitées par K. Michaelowa (2001). Cependant, son travail diffère de celui-ci sur de nombreux points. D'abord, elle utilise un modèle hiérarchique statique à trois niveaux (élèves, écoles et pays), ensuite, les scores de la cinquième année d'études sont étudiés.

Elle expose ensuite à la section 3 le modèle et décrit les variables retenues comme déterminants du niveau des acquisitions scolaires. Elle expose enfin les résultats des estimations à la section 4 avant de présenter quelques conclusions du point de vue de la politique éducative et économique.

## **1. Un cadre pour l'analyse des déterminants des apprentissages**

Les travaux de recherche menés dans le domaine des déterminants des apprentissages peuvent être divisés en trois groupes. Le premier vise à évaluer l'impact des intrants scolaires sur les apprentissages en recourant à des fonctions de production. Ces dernières années sont marquées par le débat relatif à l'effet de la taille de la classe sur les performances scolaires. Même si certaines études (Case et Deaton, 1999 ; Angrist et Lavy, 1999) semblent argumenter en faveur de l'effet positif et significatif des classes à petite taille, le consensus n'est pas encore établi sur la question. Cependant, Hanushek (1994) avait déjà montré qu'on ne peut pas conclure, de façon certaine, que la réduction de la taille des classes améliore les réussites scolaires Hoxby (2000), en utilisant des données sur les Etats-Unis, ne réussit pas à prouver l'effet de la taille de la classe sur les performances des élèves.

L'effet du ratio élèves/enseignant est également objet de controverses. Certains auteurs soutiennent que, lorsqu'il est faible, il peut avoir un effet positif sur la performance des élèves. Ainsi, sur la base des résultats en lecture et en mathématiques des enfants âgés de 16 ans dans 148 circonscriptions scolaires en Angleterre, Willms et Kerckhoff (1995) ont montré qu'une diminution de ce ratio de 25 à 16 augmente le rendement scolaire. En revanche, en utilisant des données collectées au niveau des écoles en Angleterre, entre 1992 et 1996, Bradley et Taylor (1998) ont trouvé que le nombre d'élèves par enseignant n'a pas d'effet sur la réussite aux examens. Cependant, ils ont obtenu un impact significatif, mais faible, lorsqu'ils ont étudié le rapport entre la variation de ce nombre entre 1992 et 1998, et celle des performances aux examens durant la même période.



Le second groupe de travaux, moins nombreux, examine l'effet des caractéristiques des ménages sur les acquisitions cognitives. Mc Lanahan et Sandefur (1994) ont constaté que les enfants vivant dans une famille monoparentale ou séparés de leurs parents biologiques ont des performances scolaires plus faibles que celles de leurs camarades vivant avec les leurs.

Les études sur l'impact de la taille de la famille, en général, et de la taille de la fratrie en particulier sur la scolarisation dans les pays en développement, ont abouti à des résultats très controversés qui n'autorisent pas des conclusions généralisables à d'autres pays. Au Kenya, par exemple, Buchman (2000) n'a pas trouvé de relation entre ces caractéristiques. Par ailleurs, Montgomery et Lloyd (1997) ont abouti au même résultat, lorsqu'ils ont mis en relation l'excès de fécondité et l'acquisition scolaire.

Les deux groupes de travaux présentés ci-dessus se fondent sur l'hypothèse selon laquelle les facteurs liés à l'école ont un effet sur les apprentissages scolaires, indépendamment de celui qu'exercent les caractéristiques propres à la famille. Todd et Wolpin (2003) ont remis en cause une telle indépendance, arguant du fait que les ménages prennent leurs décisions d'investissement dans l'éducation de leurs enfants en tenant en compte de la disponibilité et de la qualité des facteurs liés à l'école. Dans ce cas, l'utilisation de l'approche, en termes de fonction de production, conduirait à estimer un effet prenant en compte l'impact marginal des facteurs liés à l'école et les réponses des ménages à de tels intrants.

Das, Dercon, Habyarimana et Krishnan (2004) ont réexaminé la relation entre apprentissages scolaires et intrants propres à l'école en tenant compte de cette critique. Utilisant des données sur la Zambie, ils soutiennent qu'il existe une forte substituabilité entre les dépenses d'éducation des ménages et les subventions monétaires reçues par les écoles primaires de ce pays.

Le troisième groupe de travaux, plus récent, a mis en évidence les effets des facteurs plus étendus tels le voisinage et le quartier sur les rendements scolaires. Certains travaux ont établi un lien entre les résultats scolaires et la pauvreté de la famille et du quartier. Les enfants qui proviennent de milieux socio-économiques pauvres ont de moins bonnes performances scolaires que les enfants non pauvres (Conger *et al*, 1997 ; Entwisle et Alexander, 1990 ; Haveman et Wolfe, 1995).

Wilson (1987) a montré que le fait de vivre dans un quartier pauvre où les familles disposent de peu de ressources économiques, et qui affiche de forts taux de chômage et d'enfants monoparentaux, est associé à l'« isolement social ». Or, ce phénomène peut avoir des effets psychologiques et comportementaux négatifs sur les familles et leurs enfants. Les familles qui se trouvent dans cette situation ne valorisent pas toujours l'école.

Cette étude propose un meilleur éclairage des déterminants des apprentissages scolaires au Sénégal. En plus de l'influence des caractéristiques de l'élève, du ménage et de l'école, celle de la communauté sera prise en compte. S'il est vrai que les ménages peuvent pallier partiellement l'insuffisance des ressources de l'école, on peut supposer que la situation des écoles subit peu de changement au cours de l'année, si bien que les variations des intrants acquis par les ménages dues à des variations non anticipées des ressources des écoles sont d'ampleur très limitée. On ne tiendra donc pas compte d'une substituabilité possible entre intrants liés à l'école et ceux fournis par les ménages, suite à une variation non anticipée des premiers.

## **2. La qualité dans l'enseignement primaire : données et tendances majeures**

Le meilleur indicateur de la qualité interne d'un enseignement est ce que les élèves retiennent du curriculum suivi. C'est pourquoi les tests de connaissances constituent un instrument de premier choix pour l'évaluation de la qualité de l'éducation. Les tests menés au Sénégal sur l'apprentissage des élèves du cycle primaire sont utilisés pour apprécier le niveau de la qualité de l'enseignement primaire.

### **2.1 Données**

Deux bases de données sont combinées pour analyser des déterminants de la scolarisation primaire. La première a été produite par le Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC). Elle contient les scores aux tests de français et de mathématiques menés entre 1996 et 2000 sur une cohorte d'élèves de la deuxième année (CP) à la sixième année d'études primaires (CM2)<sup>2</sup>.

---

<sup>2</sup> Les données proviennent d'un échantillon aléatoire et stratifié d'écoles. Les tests standardisés ont été administrés également dans quatre pays africains francophones. Cette base contient aussi des informations sur les enseignants et les écoles.

Les tests standardisés ont été administrés aussi dans quatre pays africains francophones. Une limite importante de l'enquête PASEC est l'absence d'informations sur la situation socio-économique des élèves, le nombre limité de données recueillies sur les élèves, ainsi que les enseignants, enfin l'absence d'informations sur la communauté dans laquelle vivent les élèves.

C'est pourquoi une enquête rétrospective complémentaire appelée Enquête sur le bien-être des ménages au Sénégal (EBMS) a été menée en 2003 par le Centre de recherches économiques appliquées (CREA) de l'Université Cheikh Anta Diop de Dakar. L'EBMS couvre les ménages, les écoles et les communautés d'enfants âgés de 13 à 20 ans appartenant ou non à la cohorte d'élèves du PASEC. Les données recueillies portent sur les caractéristiques sociodémographiques et économiques des ménages, la trajectoire scolaire de l'enfant et des membres de sa famille, les chocs (naturels, de santé, de revenu, etc.) ayant affecté le ménage.

Sont collectées aussi des informations sur les facteurs propres à la collectivité et susceptibles d'influencer les performances des élèves (types d'écoles existantes et leur localisation, infrastructures et équipements, conditions de prise de décision concernant particulièrement l'éducation, ainsi que les conditions agro-climatiques et géographiques), la qualification des enseignants, la disponibilité de manuels et de matériels scolaires, les méthodes de gestion et les habitudes pédagogiques de l'école. Au total, deux sources de données ont été constituées et appariées. La première base renferme les informations sur les résultats aux tests du PASEC de 1995-96 à 1999-2000 ainsi que les informations sur les parcours scolaires. La deuxième source contient les caractéristiques individuelles et familiales des élèves et informe sur leurs écoles et les localités dans lesquelles elles sont implantées.

La stratégie d'échantillonnage du PASEC, qui conditionne celle de l'enquête EBMS, mérite d'être précisée. Le PASEC a utilisé la méthode d'échantillonnage en grappe. Les critères suivants ont été pris en compte lors de l'élaboration de l'échantillon : la représentativité des élèves de l'échantillon ; la couverture de tout le pays ; la prise en compte de la diversité des situations, y compris celles qui sont marginales ; la prise en compte de la taille des écoles dans la procédure de tirage et une taille considérable de l'échantillon afin de s'assurer de la précision des résultats.

La procédure de tirage des élèves a comporté trois étapes : d'abord un échantillon de 99 écoles a été tiré avec une probabilité égale au nombre total d'élèves en classe de CP, ensuite, une classe de CP dans chaque école (dans le cas où il y a plus d'une classe), et enfin, 20 élèves dans chaque classe de CP.

Une difficulté majeure de la collecte de données complémentaires est la recherche des enfants appartenant à la cohorte d'élèves du PASEC deux à trois ans après qu'ils ont quitté l'école primaire. En effet, dans le suivi de cohorte, le PASEC avait interviewé environ 1975 enfants répartis entre 99 écoles (soit 20 élèves environ par école). En définitive, 902 de ces enfants ont été retrouvés et parmi eux, seuls 701 ont été enquêtés. Cet échantillon réduit peut engendrer un biais de sélection si sa structure a subi de grandes modifications par rapport à l'échantillon global du PASEC. Le tableau 1 donne quelques indications sur l'importance de ce biais.

Tableau 1 : Comparaison entre l'échantillon prévu par le PASEC et l'échantillon observé par l'enquête EBMS

| CARACTERISTIQUES                               | Pasec          | EBMS           |
|--|----------------|----------------|
| <b>Taille</b>                                  | <b>1975</b>    | <b>700</b>     |
| Proportion des filles                          | 45,34%         | 39,23%         |
| Proportion des élèves en milieu rural          | <b>34,23%</b>  | <b>56,03%</b>  |
| Nombre d'écoles                                | <b>99</b>      | <b>59</b>      |
| Proportion d'écoles rurales                    | <b>34,34%</b>  | <b>54,24%</b>  |
| Pourcentage d'écoles pratiquant le double flux | 29,60%         | 27,60%         |
| Score général moyen 1995-1996                  | 44,39 (22,554) | 40,62 (21,717) |
| Score général moyen 1996-1997                  | 47,97 (22,611) | 46,52 (22,622) |
| Score général moyen 1997-1998                  | 40,93 (15,598) | 40,79 (15,886) |
| Score général moyen 1998-1999                  | 55,22 (18,074) | 54,17 (18,488) |
| Score général moyen 1999-2000                  | 42,63 (19,709) | 40,60 (19,017) |

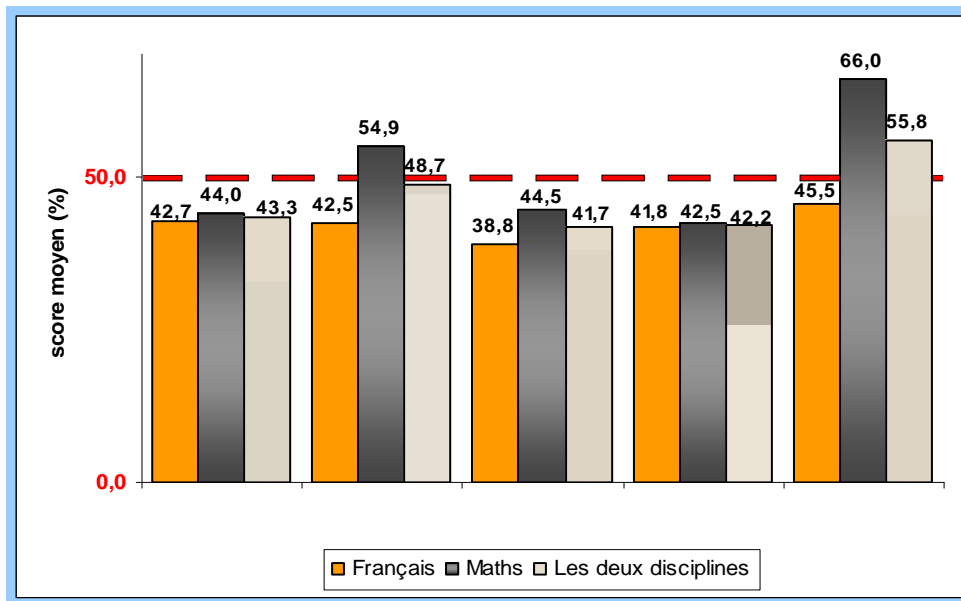
Note : Les valeurs mises entre parenthèses sont des écarts-types.

L'échantillon de départ a subi des modifications structurelles, notamment en ce qui concerne le milieu de résidence des élèves, la localisation des écoles et le nombre d'écoles. Cependant, les autres variables, précisément les distributions des résultats scolaires, n'ont pas véritablement changé. En somme, le sous échantillon d'élèves du PASEC enquêtés par l'EBMS a une structure relativement proche de celle de l'échantillon du PASEC et autorise, en conséquence, les mêmes conclusions.

## 2.2 Principales tendances des performances scolaires des élèves

Le graphique suivant présente les résultats aux tests d'évaluation en français et en mathématiques (maths) des élèves de la deuxième à la sixième année d'études primaires. Ces résultats sont présentés sous forme de pourcentages moyens de réponses correctes aux tests du PASEC.

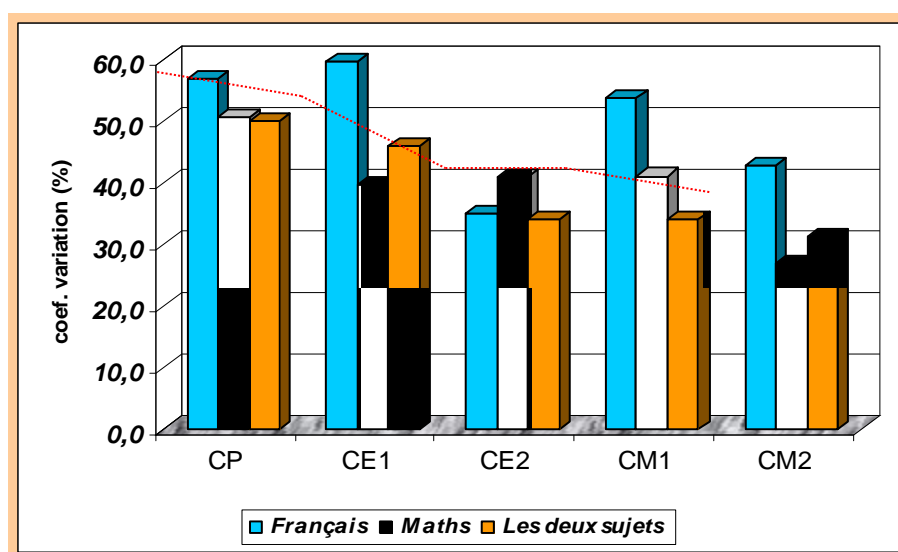
Graphique 1 : Score moyen par classe et par discipline (en %)



Globalement, le niveau moyen général des élèves est faible. C'est uniquement au CE1 (troisième année d'études) et au CM2 (sixième année d'études) que les élèves ont répondu correctement à plus de la moitié des questions en moyenne. Certes, les tests sont différents et ne sont donc pas directement comparables d'une classe à l'autre. Il se peut fort bien que le test du CP soit relativement plus difficile que celui du CE1 ou inversement. Le graphique indique aussi que le niveau moyen en maths est toujours supérieur au niveau moyen en français, quelle que soit la classe. Ce constat est très visible au CE1, au CE2 et au CM2 où il y a un décrochage assez important, de l'ordre de 5,7 à 20,5 points de pourcentage, entre les scores moyens en maths et en français. Cela laisse penser que les élèves s'en sortent plus en maths qu'en français.

Les distributions des scores laissent apparaître une forte hétérogénéité. Toutefois, ces disparités se réduisent au fur et à mesure que l'on avance dans le cycle primaire quelle que soit la discipline considérée. On peut mesurer ces disparités en calculant les coefficients de variation des scores. Ils décroissent d'une année d'études à l'autre, le coefficient de variation du score général passant de 50% au CP à 31% au CM2 (Graphique 2).

Graphique 2 : Coefficient de variation des scores par année d'étude et par discipline (en %)



Les niveaux d'acquisitions des connaissances des élèves tendent à converger au fil des classes. A l'entrée dans le cycle primaire (CP), les élèves ont des pré requis très disparates (niveaux très faibles, faibles, moyens, élevés et très élevés). Cependant, au fur et à mesure que l'on progresse en niveau, on assiste à un effet de sélection dû au fait que les élèves sont soumis à des phénomènes de redoublement et d'abandon. L'hétérogénéité caractérise aussi les performances de groupes d'élèves constitués selon différents critères. Le niveau d'instruction des parents influence très positivement les résultats de leurs enfants. En effectuant un test de comparaison des moyennes sur les scores en début et en fin de cycle, on constate que les résultats des élèves dont les parents ont un niveau supérieur ou égal à celui du collège sont significativement supérieurs au seuil de 1% (Tableau 2) à ceux de leurs camarades dont les parents ne sont pas instruits ou n'ont pas atteint le collège.

Tableau 2 : Test de comparaison des scores moyens des élèves selon leur environnement familial et le milieu de résidence

| Niveau d'instruction des parents |          |                        |                             |                |                          |
|----------------------------------|----------|------------------------|-----------------------------|----------------|--------------------------|
| <i>Score général</i>             |          | <i>Collège et plus</i> | <i>Primaire/Sans niveau</i> | <i>P-value</i> | <i>Conclusion</i>        |
| <b>CP</b>                        | Variance | 2,056                  | 0,930                       | 0,5008         | Variances égales         |
|                                  | Moyenne  | 50,80                  | 41,71                       | 0,0000**       | <b>Moyennes inégales</b> |
| <b>CM1</b>                       | Variance | 1,345                  | 0,800                       | 0,0924*        | Variances inégales       |
|                                  | Moyenne  | 46,11                  | 41,21                       | 0,0021**       | <b>Moyennes inégales</b> |
| Niveau de vie du ménage          |          |                        |                             |                |                          |
| <i>Score général</i>             |          | <i>Riche</i>           | <i>Pauvre</i>               | <i>P-value</i> | <i>Conclusion</i>        |
| <b>CP</b>                        | Variance | 1,201                  | 1,200                       | 0,2029         | Variances égales         |
|                                  | Moyenne  | 46,58                  | 40,62                       | 0,0005**       | <b>Moyennes inégales</b> |
| <b>CM1</b>                       | Variance | 0,941                  | 0,997                       | 0,1686         | Variances égales         |
|                                  | Moyenne  | 44,97                  | 39,71                       | 0,0002**       | <b>Moyennes inégales</b> |
| Possession de manuels scolaires  |          |                        |                             |                |                          |
| <i>Score général</i>             |          | <i>1 livre et plus</i> | <i>Aucun livre</i>          | <i>P-value</i> | <i>Conclusion</i>        |
| <b>CP</b>                        | Variance | 0,942                  | 1,676                       | 0,1076         | Variances égales         |
|                                  | Moyenne  | 46,43                  | 31,29                       | 0,0000**       | <b>Moyennes inégales</b> |
| <b>CM1</b>                       | Variance | 0,742                  | 1,740                       | 0,1559         | Variances égales         |
|                                  | Moyenne  | 43,26                  | 34,63                       | 0,0001**       | <b>Moyennes inégales</b> |
| Milieu de résidence              |          |                        |                             |                |                          |
| <i>Score général</i>             |          | <i>Urbain</i>          | <i>Rural</i>                | <i>P-value</i> | <i>Conclusion</i>        |
| <b>CP</b>                        | Variance | 1,248                  | 1,165                       | 0,2378         | Variances égales         |
|                                  | Moyenne  | 46,25                  | 41,20                       | 0,0034**       | <b>Moyennes inégales</b> |
| <b>CM1</b>                       | Variance | 1,017                  | 0,940                       | 0,8211         | Variances égales         |
|                                  | Moyenne  | 44,72                  | 40,19                       | 0,0012**       | <b>Moyennes inégales</b> |

Note : Nous avons considéré les résultats du CP et du CM1 afin de couvrir le début et la fin du cycle, sans toutefois tomber dans les spécificités fortes de la 1<sup>ère</sup> (C1) et de la dernière année (CM2).

\* = significativité au seuil de 10% ; \*\* = significativité au seuil de 1%.

Le niveau de vie des ménages influe fortement sur les performances scolaires de leurs enfants. Les élèves issus de familles riches ont en moyenne des résultats significativement supérieurs au seuil de 1% à ceux de leurs camarades appartenant à des ménages pauvres. Une comparaison des scores moyens en début et en fin de cycle montre que les élèves qui ont eu au moins un livre entre le CP et le CM1, ont un niveau moyen nettement supérieur à celui des élèves qui n'en ont pas possédé.

Le tableau 2 révèle que les élèves du milieu urbain ont, en moyenne, des scores plus élevés que ceux du milieu rural. Par ailleurs, le milieu de résidence a un effet modérateur sur l'influence de certains facteurs sur les apprentissages. Par exemple, le niveau de vie des ménages n'a plus la même influence dans les zones rurales que dans les zones urbaines. En effet, en effectuant un test de comparaison des scores moyens, on note une différence significative entre les résultats moyens des enfants de familles pauvres et ceux de familles riches en milieu rural ; ce qui n'est pas le cas en milieu urbain comme l'indique le tableau 3.

Tableau 3 : Test de comparaison des scores selon le niveau de vie et la zone de résidence

| Score général        | Riche | Pauvre | P-value | Conclusion               |
|----------------------|-------|--------|---------|--------------------------|
| <i>Milieu rural</i>  |       |        |         |                          |
| <b>CP</b>            | 45,66 | 40,43  | 0,0559* | <b>Moyennes inégales</b> |
| <b>CM1</b>           | 44,33 | 39,46  | 0,0318* | <b>Moyennes inégales</b> |
| <i>Milieu urbain</i> |       |        |         |                          |
| <b>CP</b>            | 46,78 | 42,40  | 0,2554  | <b>Moyennes égales</b>   |
| <b>CM1</b>           | 45,10 | 41,92  | 0,3116  | <b>Moyennes égales</b>   |

Note : \* = Significativité au seuil de 10%.

Les écoles diffèrent de par leurs caractéristiques physiques (taille, électricité, état des classes, ...), pédagogiques (double flux, classes multigrades, inspection des enseignants, ...), géographiques (région, département, ...) et par leur management (comité de gestion, association des parents d'élèves, partenariat, ...). Ces différences ont un impact sur les opportunités d'apprentissage des élèves et donc sur leurs performances. La significativité de l'hétérogénéité des résultats des élèves entre écoles est vérifiée par une analyse de la variance.

Tableau 4 : Test de significativité de la variance inter-écoles des résultats aux tests du PASEC selon l'année d'étude

| Année d'étude | Variance    |             | F-empirique | Prob > F |
|---------------|-------------|-------------|-------------|----------|
|               | Inter-école | Intra-école |             |          |
| <b>CP</b>     | 111378,19   | 186282,70   | 5,95        | 0,0000   |
| <b>CE1</b>    | 110895,88   | 157257,29   | 6,19        | 0,0000   |
| <b>CE2</b>    | 41711,00    | 51880,04    | 5,73        | 0,0000   |
| <b>CM1</b>    | 29150,12    | 58436,31    | 3,21        | 0,0000   |
| <b>CM2</b>    | 30737,41    | 21756,58    | 3,76        | 0,0000   |



Du CP au CM2, une part considérable des disparités observées dans les performances des élèves est liée aux différences entre écoles. Le tableau 4 fait apparaître une variance inter-écoles très significative au seuil de 1% (Prob>0,0000). En outre, l'importance de la variance intra-école attribuable aux élèves et à leurs ménages justifie l'utilisation des informations fournies par la base de l'EBMS. Les différences dans les performances des élèves s'expliquent par leurs caractéristiques individuelles et par celles de leur ménage. Ces différences sont dues aussi au fait qu'ils sont placés dans des environnements d'apprentissage très différents (différences entre écoles, différences entre quartiers/communautés).

### 3. Modèles et variables

#### 3.1 Modèles

Le cadre conceptuel généralement utilisé pour analyser les apprentissages scolaires présente les performances de l'élève à une date ou à une période donnée comme une fonction cumulative des facteurs relevant de l'élève lui-même, de la famille, de l'école et de la communauté. En interagissant entre eux et avec les capacités innées de l'élève, on peut déterminer le niveau et la progression de ses acquisitions scolaires. Ainsi, la fonction de production éducative peut être écrite comme suit :

$$Y_{ijk} = \alpha + V_{ijk} \cdot \lambda + M_{jk} \cdot \gamma + Z_k \cdot \delta + \xi_{ijk}$$

Avec  $Y_{ijk}$  le score à l'année  $i$  de l'élève  $j$  dans l'école  $k$  ;

$V_{ijk}$  les caractéristiques annuelles de l'élève  $j$  dans l'école  $k$  à l'année  $i$  ;

$M_{jk}$  les caractéristiques individuelles et familiales de l'élève  $j$  dans l'école  $k$  ;

$Z_k$  les facteurs relatifs à l'école et à la communauté  $k$  ;

$\xi_{ijk}$  le terme d'erreur.

Les données requises pour estimer ce type de modèle (modèle multi niveaux) présentent une structure hiérarchique : les scores sont regroupés par élève, les élèves par école et les écoles par communauté. Une observation sur l'école est attribuée à chacun des élèves de celle-ci ; de même une observation sur la communauté est attribuée à l'ensemble des écoles appartenant à ladite communauté. Dans ces conditions, l'utilisation des régressions multiples classiques risque de fournir des résultats biaisés.

Trois problèmes principaux limitent en effet l'utilisation des moindres carrés ordinaires (MCO) dans le traitement de ce type de données : l'hétérogénéité des régressions lorsque les relations entre les variables explicatives et la variable expliquée varient selon les répondants (Bryk et Raudenbush, 1992) ; l'existence d'une corrélation intra-école qui indique le degré de ressemblance entre individus appartenant au même groupe (Barcikowski, 1981) ; et la sous-estimation des écarts-types et donc une surestimation de la significativité statistique des résultats qui résulte de la non-vérification de l'hypothèse d'indépendance des observations (Hofman & Griffin, 1997).

Nous utiliserons donc dans ce travail les modèles linéaires hiérarchiques de croissance qui ont pour vertu de modéliser explicitement le problème des « niveaux d'analyse » et de donner les meilleures estimations linéaires non biaisées des paramètres. Ils sont spécifiquement conçus pour pallier les limites des MCO présentées ci-dessus.

Les variables explicatives retenues étant de trois types (les facteurs intra-individus, les facteurs inter-individus et les facteurs environnementaux), des modèles linéaires hiérarchiques à trois niveaux seront donc construits : le niveau intra-individu (niveau 1), le niveau inter-individus (niveau 2) et le niveau inter-groupes (niveau 3). Dans le premier modèle (niveau 1), c'est le temps utilisé pour estimer la différence entre les scores obtenus par chaque élève pendant les années 1996, 1997, 1998, 1999 et 2000.

Le second modèle (niveau inter-individus) permet de comparer les différences entre les élèves dans les résultats scolaires *en moyenne*, et le troisième, de saisir les effets des caractéristiques des écoles et de la communauté sur les scores moyens des élèves.

En raison de la complexité des modèles à trois niveaux, nous commencerons par un modèle simple correspondant à la première étape de l'analyse. Cette étape donne une décomposition de la variance excluant l'influence des variables explicatives. Les étapes suivantes consisteront à introduire dans le modèle initial d'abord les variables relatives à l'individu et au ménage, ensuite celles relatives à l'école et à la communauté.

### **Le modèle initial : analyse de la variance**

Les variables dépendantes ( $Y$ ) sont les scores aux tests administrés aux élèves pendant cinq années. La variable explicative relative au temps est désignée par  $X$ , les variables explicatives relatives à l'élève et à son ménage par  $W$ , celles relatives au niveau école et communauté par  $Z$ . Une précaution importante à prendre avant d'élaborer un modèle linéaire hiérarchique complexe pour expliquer la variance des scores entre les élèves consiste à vérifier si cette variance existe de façon significative. La première étape sera donc de mener une analyse de la variance avec effets aléatoires. Aucune variable explicative ne sera introduite dans ce modèle qui peut être représenté par les trois équations suivantes (Bryk et Raudenbush, 1992) :

$$\text{Niveau 1 (intra-individu) : } Y_{ijk} = \alpha_{0jk} + \varepsilon_{ijk} \quad (1)$$

avec  $i$  l'année du test,  $j$  l'élève et  $k$  l'école,

$Y_{ijk}$  : score au test de l'année  $i$  de l'élève  $j$  de l'école  $k$ ,

$\alpha_{0jk}$  : moyenne de l'élève  $j$  de l'école  $k$  aux différents tests sur les cinq années,

$\varepsilon_{ijk}$  : terme d'erreur mesurant l'écart entre le score au test de l'année  $i$  de l'élève  $j$  de l'école  $k$  et son score moyen ( $\alpha_{0jk}$ ) ; cette variable aléatoire est supposée suivre une loi normale de moyenne nulle et de variance constante,  $N(0, \sigma^2)$ .

$$\text{Niveau 2 (individuel et familial)} : \alpha_{0jk} = \beta_{00k} + \tau_{0jk} \quad (2)$$

$\alpha_{0jk}$  : moyenne de l'élève  $j$  de l'école  $k$  est utilisée comme variable expliquée dans le modèle du niveau 2,

$\beta_{00k}$  : moyenne de l'école  $k$  aux différents tests (niveau 2),

$\tau_{0jk}$  : terme d'erreur mesurant l'écart entre la moyenne de l'école  $k$  et de celle de son élève  $j$ , il est supposé distribué aléatoirement avec une moyenne nulle et une variance constante  $N(0, \tau_{\pi}^2)$ .

$$\text{Niveau 3 (école et communauté)} : \beta_{00k} = \gamma_{000} + \mu_{00k} \quad (3)$$

$\beta_{00k}$  : moyenne de l'école  $k$  aux différents tests du niveau 2

$\gamma_{000}$  : moyenne générale des élèves aux différents tests

$\mu_{00k}$  : terme d'erreur mesurant l'écart entre la moyenne générale des élèves et la moyenne de l'école  $k$  aux différents tests, il est supposé suivre une loi normale de moyenne nulle et de variance constante  $N(0, \tau_{\beta}^2)$ .

L'estimation de la variance à partir de paramètres d'interception aléatoires fournira des informations sur la partition de cette variance entre les trois niveaux, ainsi que sur le coefficient intra classe, et testera l'hypothèse d'égalité de la moyenne entre les niveaux.

### ***Le modèle final***

Dans la deuxième étape, nous nous intéressons uniquement aux relations entre les variables  $Y$  qui sont constituées des scores aux tests de 1996 à 2000 et des variables  $X$  de l'élève  $j$  dans l'école  $k$ . Le sous-modèle se présente comme suit :

$$\text{Niveau 1 : } Y_{ijk} = \alpha_{0jk} + \alpha_{1jk} X_{ijk} + \varepsilon_{ijk} \quad (4)$$

où  $\alpha_{0jk}$  est la valeur moyenne des performances de l'élève  $j$  et  $\alpha_{1jk}$  sont les effets des variables explicatives  $X$ . Ces effets varient d'un individu à l'autre. Ils dépendent des variables du niveau inter-individuel.

Les modèles aux niveaux 2 et 3 permettront de gérer le problème de la non indépendance des observations.

Dans la deuxième étape, nous allons expliquer le score moyen aux différents tests ( $\alpha_{0jk}$ ) et les effets ( $\alpha_{1jk}$ ) des variables explicatives du niveau 1 par les variables  $W$  du niveau 2. Cette opération permet de savoir si les variables  $W$  ont un impact non seulement sur la moyenne de l'élève mais aussi sur les paramètres estimés exprimant l'influence des variables indépendantes sur les scores de l'élève. La structure des sous-modèles est la suivante :

$$\text{Niveau 2 : } \alpha_{0jk} = \beta_{00k} + \beta_{01k} W_{jk} + \tau_{0jk} \quad (5)$$

$$\alpha_{1jk} = \beta_{10k} + \beta_{11k} W_{jk} + \tau_{1jk} \quad (6)$$

Où  $\beta_{00k}$  est la valeur moyenne des performances du groupe  $k$  et  $\beta_{01k}$ ,  $\beta_{10k}$  et  $\beta_{11k}$  sont les effets des variables du niveau 2 sur  $\alpha_{1jk}$ .

Ces effets dépendent des caractéristiques des écoles/communautés et varient d'une école à l'autre.

Dans la troisième étape, nous cherchons à expliquer  $\beta_{00k}$ ,  $\beta_{01k}$ ,  $\beta_{10k}$  et  $\beta_{11k}$  par les variables au niveau école et communauté. L'analyse se fera avec un effet modérateur des caractéristiques de celles-ci.

$$\text{Niveau 3 : } \beta_{00k} = \gamma_{000} + \gamma_{001} Z_k + \mu_{00k} \quad (7)$$

$$\beta_{01k} = \gamma_{010} + \gamma_{011} Z_k + \mu_{01k} \quad (8)$$

$$\beta_{10k} = \gamma_{100} + \gamma_{101}Z_k + \mu_{10k} \quad (9)$$

$$\beta_{11k} = \gamma_{110} + \gamma_{111}Z_k + \mu_{11k} \quad (10)$$

où  $\gamma_{000}$  la moyenne des performances commune à l'ensemble des élèves du PASEC.

### 3.2 Variables

Trois variables dépendantes sont retenues : les scores de l'élève en français, en mathématiques, et la moyenne de ces deux scores. Ces variables sont expliquées, d'une part, par les caractéristiques de l'élève et du ménage auquel il appartient, la qualité de l'offre d'éducation disponible et, d'autre part, par celles de la communauté dans laquelle l'école est implantée. Dans chacun de ces trois domaines, des variables ont été sélectionnées sur la base de leur disponibilité dans les bases de données appariées et de leur pertinence du point de vue de la littérature sur les déterminants de la qualité de l'éducation primaire.

#### 3.2.1 *Caractéristiques de l'élève et du ménage auquel il appartient*

Les caractéristiques de l'élève prises en compte sont : l'âge en deuxième année d'études primaires (cours préparatoire), le genre, le parcours scolaire (fréquentation préscolaire, nombre de redoublements), les aptitudes intellectuelles mesurées par le score au test organisé au début de la première année d'études primaires (cours d'initiation) par le PASEC, l'état de santé rétrospectif, le fait que l'enfant est confié ou non, et la disponibilité de livres de lecture et de calcul. On suppose que l'âge agit négativement sur les performances de l'élève. Plus il est âgé, plus il risque d'être exposé au travail des enfants, ce qui réduit considérablement le temps qu'il peut consacrer à ses études.

Le genre peut exercer aussi une influence significative surtout en interaction avec d'autres variables. Par exemple, une fille sera plus à l'aise à l'école si elle rencontre plus d'enseignantes durant son parcours scolaire.

Le parcours scolaire de l'enfant doit aussi être pris en compte. Beaucoup de travaux ont déjà montré que le redoublement, contrairement à l'effet attendu, loin d'améliorer les performances de l'élève, est une source de découragement pour lui et ses parents. En effet, le redoublement est supposé toucher uniquement les élèves en difficulté et, dans ce cas, ce sont les résultats scolaires qui l'expliquent. Or, certains parents n'hésitent pas à mettre en évidence une trop grande part de l'arbitraire et de la subjectivité des enseignants dans les décisions de redoublement.

D'autres auteurs (Crahay, Marcel, 1996 ; Aletta, 1984) montrent que l'évaluation des enseignants est fonction du groupe d'élèves auquel ils sont confrontés (loi de Posthumus). Ainsi, une grande variabilité des pratiques de redoublement existerait selon le lieu de scolarisation et la norme théorique du système se traduirait, en réalité, par des différences selon les écoles ou les classes porteuses d'inégalités de traitement pour les élèves. Dans ces conditions, le redoublement va avoir un impact certain sur la motivation des élèves.

Quant à la fréquentation préscolaire, on s'attend à ce qu'elle influe positivement sur les acquisitions de l'élève à l'école primaire, ce type d'éducation développant des capacités qui favorisent l'apprentissage scolaire. Une autre variable relative à l'élève est son statut d'enfant confié ou non. On présume qu'un enfant qui ne vit pas au moins avec l'un de ses parents bénéficie de conditions de vie moins favorables à l'apprentissage. Une dernière variable spécifique à l'élève est son état de santé rétrospectif. Une santé chancelante réduit le temps de fréquentation scolaire et la capacité à apprendre.

En ce qui concerne les caractéristiques du ménage, sont retenus : le niveau de bien-être, le niveau d'instruction du père, de la mère, l'usage du français à domicile, les chocs économiques et sociaux ayant affecté le ménage. Nous ne disposons pas dans l'enquête EBMS des données relatives aux indicateurs monétaires de niveau de vie tels que le revenu ou la dépense. C'est pourquoi, nous avons utilisé comme alternative les informations non monétaires inhérentes aux actifs et patrimoine des ménages pour construire un indice composite qui donne une bonne approximation de la richesse de ces derniers (Sahn et Stifel, 2003). Cet indicateur est construit par la méthode d'analyse des correspondances multiples (ACM).

Au départ, on disposait de 40 variables représentatives des actifs. Une première ACM a permis d'écartier celles qui ont de faibles contributions sur le premier facteur. Finalement, nous avons conservé 23 variables qui ont des contributions supérieures à 2. Une nouvelle ACM faite sur ces variables indique que le premier axe factoriel capte 20,10 % des disparités entre les modalités. Ce pourcentage est suffisant pour capter la différenciation entre les ménages en matière de conditions de vie.

L'indicateur synthétique du niveau de vie des ménages (ISNV) se présente donc comme une combinaison linéaire des modalités prises par les ménages sur les variables retenues, pondérées respectivement par leurs coordonnées sur le premier axe factoriel. Ainsi, pour chaque ménage, on a calculé une valeur numérique reflétant son niveau de vie.

Le niveau d'instruction des parents est une variable importante pour expliquer les performances réalisées par l'élève. Un environnement lettré exerce une influence positive sur la capacité de l'enfant à maîtriser la lecture, l'écriture et le calcul. De même, le fait que l'on parle français à la maison familiarise l'enfant avec la langue qu'il va apprendre à l'école et assure une continuité entre celle-ci et le milieu familial. Les autres variables relatives au ménage qui sont prises en compte sont les chocs économiques (chômage...) et sociaux (décès du père ou de la mère) subis par l'élève. Ces chocs ont un impact négatif sur la situation matérielle de la famille et peuvent déteindre négativement sur le parcours scolaire de ce dernier.

### **3.2.2 Caractéristiques de l'offre éducative**

De nombreuses caractéristiques de l'école sont prises en compte. Premièrement, par la méthode ACM, est construit un *indice de disponibilité et de qualité des infrastructures de l'école* (IDQIE). Cet indice est obtenu à partir de 11 variables (disponibilité de l'électricité, matériaux de construction des murs, du toit, du plancher ; état général des bâtiments de l'école, existence d'un point d'eau potable, d'une cantine, d'une bibliothèque, etc.). On prévoit que les écoles les mieux dotées en infrastructures ont les meilleurs résultats.



Deuxièmement, le caractère complet ou non du cycle primaire est pris en compte. Les faibles performances de l'élève peuvent être imputées au fait que l'école qu'il fréquente n'offre pas tous les niveaux de cours du cycle primaire. Sachant que les coûts qu'engendrera la poursuite des études de leur enfant dans un village éloigné sont hors de leur portée, les parents, même s'ils le scolarisent, investiront peu dans ses études puisqu'il devra prématurément les arrêter. Il est donc important de porter une attention particulière à la discontinuité de l'offre éducative.

Troisièmement, l'organisation pédagogique agit sur les acquisitions scolaires. Son effet est capté par l'introduction de deux variables indiquant si l'élève fréquente une classe à double flux ou une classe multigrade. Le double flux a un effet négatif sur les apprentissages des élèves en raison de son faible volume d'heures d'enseignement officiel (500), alors que le multigrade qui favorise l'auto-apprentissage aurait plutôt un effet positif.

Quatrièmement, le contrôle exercé sur l'école par l'administration publique ou les communautés a un impact important sur les performances scolaires. Celui-ci est appréhendé à travers les variables suivantes : le nombre de visites de l'inspecteur, l'existence d'une association de parents d'élèves, d'un comité de gestion ou d'une cellule école-milieu. On suppose que plus cette surveillance est forte, plus important est le temps effectif d'apprentissage, et plus grandes sont les acquisitions scolaires.

Cinquièmement, les caractéristiques de l'enseignant exercent une influence significative sur les résultats scolaires. Les variables retenues sont le genre, le niveau académique et la formation continue reçue par l'enseignant. On anticipe qu'une enseignante exerce une influence positive sur les filles qui s'identifient plus facilement à elle qu'à un enseignant. Pour le niveau d'instruction, une distinction est opérée entre avoir ou non le baccalauréat. On prévoit que le niveau d'éducation académique de l'enseignant(e) n'a pas un effet significatif sur les acquisitions scolaires de l'élève. Enfin, le fait de bénéficier d'une formation continue devrait améliorer les méthodes pédagogiques de l'enseignant(e) et agir positivement sur les acquisitions scolaires de ses élèves.

### 3.2.3 Caractéristiques de la communauté

Deux indicateurs ont servi à apprécier le niveau de développement des communautés dans lesquelles sont implantées les écoles. Le premier, appelé *Indicateur de Développement Économique* (IDE), synthétise par la méthode ACM les informations disponibles sur les infrastructures économiques (existence et qualité de l'électricité, accès à l'eau potable et au téléphone, système d'évacuation des eaux usées et des ordures, etc.). Une valeur élevée de cet indicateur reflète une dotation importante en infrastructures qui influent positivement sur les conditions de vie des ménages et favorisent leur investissement dans l'éducation de leurs enfants.

Le deuxième indicateur, appelé *Indice de Développement Social* (IDS), porte sur la disponibilité des infrastructures d'éducation et de santé dans la communauté. Afin de bien refléter l'importance relative des différentes catégories d'infrastructures, le système de pondération suivant a été adopté : école primaire = 1 ; collège = 2 ; école secondaire = 3 ; hôpital national = 20 ; hôpital régional = 10 ; clinique = 8 ; centre de santé = 7,5 ; poste de santé/dispensaire = 5 ; case de santé = 3 ; maternité = 1 ; cabinet médical = 1. L'existence dans une communauté de nombreuses écoles primaires crée un environnement lettré plus favorable aux études, et réduit considérablement la distance entre l'école et le domicile de l'élève. La présence de collèges et de lycées donne une assurance aux parents que leurs enfants pourront poursuivre leurs études à la fin du cycle primaire sans être obligés d'aller résider en ville ou dans une localité rurale éloignée. En outre, plus les écoles primaires, collèges et lycées sont nombreux dans la communauté, plus les élèves fréquentant le cycle primaire ont des chances de bénéficier d'un soutien à domicile fourni par des frères, sœurs, ou voisins qui ont atteint des niveaux d'études plus élevés.

Un *indicateur de mobilisation sociale* (IMS) est construit pour apprécier la capacité de la communauté à s'organiser pour produire un bien collectif. On peut supposer que c'est dans les communautés où l'action de l'État se fait le moins sentir que les initiatives communautaires se développent. Par conséquent, on doit s'attendre à ce que les conditions d'apprentissage y soient moins propices que dans les communautés où les populations éprouvent moins le besoin de s'auto organiser pour faire face à des besoins collectifs déjà relativement bien satisfaits.

Les variables prises en compte pour l'IMS sont l'existence ou non de programmes d'alphabétisation, de santé, de planning familial ou d'assistance mutuelle sponsorisés par la communauté, ainsi que l'existence ou non d'efforts pour rendre propre l'environnement, construire ou réparer des infrastructures. Ainsi, six variables sont considérées, chacune prenant la valeur 1 si la réponse est positive, et la valeur 0, si elle est négative. La valeur de l'IMS est calculée en faisant la somme des points obtenus par une communauté.

Un indicateur appelé *indice composite des chocs* (ICC) est construit pour chaque communauté. Il permet de prendre en compte les chocs qu'elle a subis au cours de la période 1996-2000. Ils peuvent être positifs (ouverture d'écoles et/ou de centres de soins, projets de développement, installation électrique...) ou négatifs (fermeture d'écoles et/ou de centres de santé, incendie, inondation, épidémie...). On anticipe qu'ils influencent les apprentissages des élèves, positivement dans le premier cas et négativement dans le second cas. Étant donné que ces chocs se présentent sous forme de variables dichotomiques, le chiffre -1 est attribué à un choc négatif et le chiffre +1 à un choc positif. La valeur de l'indice composite correspond à la somme des modalités obtenues par chaque communauté.

#### **4. Résultats**

L'estimation des modèles présentés dans la section précédente permet d'obtenir les effets individuels et les interactions entre les facteurs qui influencent les apprentissages des enfants. Pour faciliter l'interprétation de ces effets, les scores sont centrés et réduits. L'utilisation de ces scores standardisés offre un avantage : les coefficients des variables explicatives s'interprètent comme des pourcentages et sont directement comparables. Ces modèles sont estimés à partir du logiciel HLM 6.02 de Radenbush, Bryk, Cheong, Congdon et du Toit (2004) qui est conçu spécifiquement pour ces types de modèles.

Tableau 5 : Statistiques sommaires relatives aux variables des modèles

| VARIABLES  | NOMBRE | MOYENNE | ECART-TYPE | MINIMUM | MAXIMUM |
|--|--------|---------|------------|---------|---------|
| <b>Niveau intra-élève (1)</b>                                    |        |         |            |         |         |
| ANNEE  | 2654   | -0.13   | 1.43       | -2.00   | 2.00    |
| SCORE FRANÇAIS STANDARDISES                                      | 2654   | 0.00    | 1.00       | -1.82   | 2.73    |
| SCORE MATH STANDARDISES  | 2654   | 0.00    | 1.00       | -2.20   | 2.27    |
| SCORE GENERAL STANDARDISES                                       | 2654   | 0.00    | 1.00       | -2.16   | 2.56    |
| <b>Niveau inter-élèves</b>                                       |        |         |            |         |         |
| AGE AU CP (1)  | 700    | 8.16    | 0.73       | 6.00    | 11.00   |
| FILLE (1)  | 700    | 0.39    | 0.49       | 0.00    | 1.00    |
| SCORE FR_INITIAL (1)   | 700    | 39.39   | 23.22      | 0.00    | 100.00  |
| SCORE MATH_INITIAL (1)   | 700    | 42.17   | 21.61      | 0.00    | 92.31   |
| SCORE GENER_INITIAL (1)  | 700    | 22.57   | 15.03      | 0.00    | 84.77   |
| CLASSE SOCIO-ECONOMIQUE (2)                                      | 700    | 2.37    | 1.12       | 1.00    | 4.00    |
| USAGE FRANÇAIS A DOMICILE (2)                                    | 700    | 0.15    | 0.36       | 0.00    | 1.00    |
| EDUCATION DU CHEF MENAGE (2)                                     | 700    | 1.54    | 1.94       | 0.00    | 6.00    |
| CLASSE (1) RICHE (2)   | 700    | 0.18    | 0.38       | 0.00    | 1.00    |
| CLASSE (4) PAUVRE (2)  | 700    | 0.14    | 0.35       | 0.00    | 1.00    |
| MALADIE GRAVE (2)  | 700    | 0.03    | 0.18       | 0.00    | 1.00    |
| ENFANT CONFIE (2)  | 700    | 0.18    | 0.39       | 0.00    | 1.00    |
| ORPHELIN DE PERE (2)   | 700    | 0.10    | 0.30       | 0.00    | 1.00    |
| ORPHELIN DE MERE (2)   | 700    | 0.04    | 0.21       | 0.00    | 1.00    |
| CHOMAGE (2)  | 700    | 0.05    | 0.23       | 0.00    | 1.00    |
| ENFANT ORPHELIN (2)  | 700    | 0.14    | 0.34       | 0.00    | 1.00    |
| NBRE LIVRE DE CALCUL (1)   | 700    | 0.79    | 0.84       | 0.00    | 4.00    |
| NBRE LIVRE DE LECTURE (1)  | 700    | 1.45    | 1.04       | 0.00    | 4.00    |
| NBRE DE LIVRE TOTAL (1)  | 700    | 2.24    | 1.68       | 0.00    | 8.00    |
| FREQUENTATION PRESCOLAIRE (2)                                    | 700    | 0.07    | 0.25       | 0.00    | 1.00    |
| NOMBRE DE REDOUBLEMENT (1)                                       | 700    | 0.51    | 0.66       | 0.00    | 4.00    |
| <b>Niveau inter-écoles/communautés</b>                           |        |         |            |         |         |
| MILIEU RURAL (2)   | 59     | 0.54    | 0.50       | 0.00    | 1.00    |
| INDICE DEVELOP SOCIAL (2)  | 59     | 28.28   | 20.24      | 1.00    | 87.50   |
| IMMIGRANT (2)  | 59     | 0.66    | 0.48       | 0.00    | 1.00    |
| MOBILISATION SOCIALE (2)   | 59     | 2.56    | 1.60       | 0.00    | 6.00    |
| INDICE DES CHOCS (2)   | 59     | 1.12    | 1.33       | -2.00   | 5.00    |
| INFRASTRUCTURES ECONOMIQUES (2)                                  | 59     | -0.07   | 6.36       | -8.98   | 10.04   |
| INFRASTRUCTURES PHYSIQ_ECOL (2)                                  | 59     | -0.67   | 3.43       | -10.42  | 5.08    |
| ECOLE A CYCLE INCOMPLET (2)                                      | 59     | 0.14    | 0.35       | 0.00    | 1.00    |
| ASSOCIATION PARENTS D'ELEVES (2)                                 | 59     | 0.42    | 0.50       | 0.00    | 1.00    |
| CELLULE ECOLE-MILIEU (2)   | 59     | 0.19    | 0.39       | 0.00    | 1.00    |
| COMITE DE GESTION (2)  | 59     | 0.59    | 0.50       | 0.00    | 1.00    |
| ACTIVITE AGRICOLE (2)  | 59     | 0.59    | 0.50       | 0.00    | 1.00    |
| ACTIVITE DE COMMERCE (2)   | 59     | 0.37    | 0.49       | 0.00    | 1.00    |
| RATIO NBRE ELEVE/CLASSE (2)                                      | 59     | 62.03   | 11.60      | 34.33   | 90.67   |
| ECOLE A DOUBLE FLUX (1)  | 59     | 0.29    | 0.46       | 0.00    | 1.00    |
| ECOLE A MULTI-GRADE (1)  | 59     | 0.07    | 0.25       | 0.00    | 1.00    |
| NBRE ENSEIGNANTS AYANT BAC (1)                                   | 59     | 1.83    | 1.09       | 0.00    | 4.00    |
| NBRE ENSEIGNANTS AYANT SUIVI AU MOINS 2 FORMATIONS CONTINUES (1) | 59     | 2.12    | 1.10       | 0.00    | 4.00    |
| NOMBRE D'ENSEIGNANTES (1)  | 59     | 0.80    | 1.10       | 0.00    | 4.00    |

Notes : (1) = Variables provenant de la base du PASEC.

(2) = Variables provenant de l'enquête EBMS.

Avant l'estimation des modèles qui vont expliquer la variance des acquisitions scolaires des élèves, nous devons vérifier que cette variance existe de façon significative dans les trois niveaux que nous avons distingués. La source de variance des performances des élèves provient de trois niveaux : intra-individu (scores annuels), inter-individus, et inter écoles/communautés. La décomposition de la variance des scores en français et en maths aux tests du PASEC donne un aperçu de la variabilité expliquée par ces niveaux.

Tableau 6 : Variance résiduelle des scores en français, en mathématiques et du score moyen des deux disciplines

| Composante de la variance résiduelle | Score moyen |         |          | Score en Français |         |          | Score en Math |         |          |
|--------------------------------------|-------------|---------|----------|-------------------|---------|----------|---------------|---------|----------|
|                                      | Coefficient | P-value | Part (%) | Coefficient       | P-value | Part (%) | Coefficient   | P-value | Part (%) |
| <b>Niveau 1</b>                      |             |         |          |                   |         |          |               |         |          |
| $V(\varepsilon_{ijk}) = E$           | 0,3729      | -----   | 37,7     | 0,4797            | -----   | 47,5     | 0,4445        | -----   | 43,5     |
| <b>Niveau 2</b>                      |             |         |          |                   |         |          |               |         |          |
| $V(r_{0jk}) = R_0$                   | 0,4087      | 0,000   | 41,3     | 0,3202            | 0,000   | 31,7     | 0,4086        | 0,000   | 40,0     |
| <b>Niveau 3</b>                      |             |         |          |                   |         |          |               |         |          |
| $V(u_{00k}) = U_{00}$                | 0,207       | 0,000   | 21,0     | 0,209             | 0,000   | 20,7     | 0,1684        | 0,000   | 16,5     |
| <b>Total</b>                         | 1,0         | -----   | 100,0    | 1,0               | -----   | 100,0    | 1,0           | -----   | 100,0    |

Quelle que soit la discipline, la grande part de la variabilité des scores des élèves est expliquée par les caractéristiques intra et inter élèves. Par exemple, 47,5 % de la variance des résultats en français sont imputés au niveau 1 et 31,7 % au niveau 2. Le même constat est observé en maths où 43,5 % de la variance est expliquée par le niveau 1 et 40 % par le niveau 2. Toutefois, la part de la variance expliquée par le niveau 3 (inter écoles/communautés) n'est pas négligeable (20,7 % en français ; 16,5% en maths et 21,0 % pour le score moyen). Elle est d'ailleurs très significative au seuil de 1 % quelle que soit la discipline considérée. Ces constats permettent donc de conclure qu'il est nécessaire de tenir compte de la nature hiérarchique des facteurs susceptibles d'expliquer l'hétérogénéité des performances des élèves.

Pour analyser les acquisitions scolaires des élèves, nous allons d'abord modéliser leurs scores moyens (maths et français confondus) puis, nous compléterons les résultats obtenus en faisant une modélisation par discipline.

#### 4.1 Score moyen des deux disciplines

Les tableaux 7 et 8 montrent comment les caractéristiques individuelles des élèves, leur environnement familial et scolaire, ainsi que la communauté affectent le niveau moyen de leurs acquisitions scolaires.

Tableau 7 : Résultats d'estimation des effets aléatoires

| Composante de la variance résiduelle | Modèle 1:                |         | Modèle 2:                |         | Modèle final:            |         |
|--------------------------------------|--------------------------|---------|--------------------------|---------|--------------------------|---------|
|                                      | Coefficient              | P-value | Coefficient              | P-value | Coefficient              | P-value |
| <b>Niveau 1</b>                      |                          |         |                          |         |                          |         |
| $V(\varepsilon_{ijk}) = E$           | <b>0,373</b><br>(0,6107) | -----   | <b>0,372</b><br>(0,6099) | -----   | <b>0,371</b><br>(0,6090) | -----   |
| <b>Niveau 2</b>                      |                          |         |                          |         |                          |         |
| $V(r_{0jk}) = R0$                    | <b>0,409</b><br>(0,6393) | 0,000   | <b>0,149</b><br>(0,3859) | 0,000   | <b>0,144</b><br>(0,3799) | 0,000   |
| $V(r_{1jk}) = R1$                    | <b>0,001</b><br>(0,0276) | >0,500  | <b>0,000</b><br>(0,0157) | 0,206   | <b>0,000</b><br>(0,0163) | 0,217   |
| <b>Niveau 3</b>                      |                          |         |                          |         |                          |         |
| $V(u_{00k}) = U_{00}$                | <b>0,207</b><br>(0,4553) | 0,000   | <b>0,183</b><br>(0,4274) | 0,000   | <b>0,084</b><br>(0,2905) | 0,002   |
| $V(u_{06k}) = U_{06}$                | -----                    | -----   | <b>0,054</b><br>(0,2332) | 0,234   | <b>0,039</b><br>(0,1970) | 0,127   |
| $V(u_{010k}) = U_{010}$              | -----                    | -----   | <b>0,005</b><br>(0,0710) | 0,141   | <b>0,004</b><br>(0,0621) | 0,041   |
| $V(u_{10k}) = U_{10}$                | <b>0,007</b><br>(0,0831) | 0,000   | <b>0,008</b><br>(0,0901) | 0,194   | <b>0,009</b><br>(0,0941) | 0,126   |
| $V(u_{18k}) = U_{18}$                | -----                    | -----   | <b>0,001</b><br>(0,0386) | 0,147   | <b>0,002</b><br>(0,0451) | 0,128   |
| Variance expliquée (R2)              | <b>0,013</b>             |         | <b>0,272</b>             |         | <b>0,413</b>             |         |
| Variance totale                      | 1,0                      |         | 1,0                      |         | 1,0                      |         |
| Déviante du modèle                   | 6156,64                  |         | 5702,98                  |         | 5670,73                  |         |
| Nombre de paramètres estimés         | 9                        |         | 47                       |         | 59                       |         |

Note : Les écarts-types sont entre parenthèses.

$V(\varepsilon_{ijk})$  = Composante de la variance résiduelle du niveau 1 ;

$V(r_{0jk})$  et  $V(r_{1jk})$  sont les variances des résidus du niveau 2 ;

$V(u_{00k})$ ,  $V(u_{06k})$ ,  $V(u_{010k})$ ,  $V(u_{10k})$  et  $V(u_{18k})$  sont les variances des résidus du niveau 3.

Tableau 8 : Résultats de l'estimation des effets fixes

| VARIABLES   | Modèle 1:                  |         | Modèle 2:                   |         | Modèle final:               |         |
|---|----------------------------|---------|-----------------------------|---------|-----------------------------|---------|
|   | Coefficient                | P-value | Coefficient                 | P-value | Coefficient                 | P-value |
| <b>Niveau annuel</b>  |                            |         |                             |         |                             |         |
| Constante   | <b>-0,064</b><br>(0,0661)  | 0,339   | <b>-0,236</b><br>(0,3407)   | 0,491   | <b>-0,450</b><br>(0,3109)   | 0,154   |
| Rural (RURAL)   | -----                      | -----   | -----                       | -----   | <b>-0,152</b><br>(0,1154)   | 0,194   |
| Comité de gestion actif à l'école (GEST_ACT)                      | -----                      | -----   | -----                       | -----   | <b>0,231*</b><br>(0,1042)   | 0,031   |
| Double flux (DBLFLUX)   | -----                      | -----   | -----                       | -----   | <b>-0,061</b><br>(0,1213)   | 0,616   |
| Multigrade (MGRADE)   | -----                      | -----   | -----                       | -----   | <b>-0,616**</b><br>(0,1646) | 0,001   |
| Nbre d'enseignants bacheliers rencontrés entre CP - CM1 (NBR_BAC) | -----                      | -----   | -----                       | -----   | <b>0,141**</b><br>(0,0414)  | 0,002   |
| Age au CP (AGE_CP)  | -----                      | -----   | <b>-0,072~</b><br>(0,0392)  | 0,066   | <b>-0,078*</b><br>(0,0346)  | 0,025   |
| Fille (FILL)  | -----                      | -----   | <b>-0,031</b><br>(0,0497)   | 0,535   | <b>-0,021</b><br>(0,0494)   | 0,670   |
| Score initial (SG_INIT)   | -----                      | -----   | <b>0,027**</b><br>(0,0016)  | 0,000   | <b>0,027**</b><br>(0,0017)  | 0,000   |
| Enfant confié (CONFIE)  | -----                      | -----   | <b>-0,001</b><br>(0,0549)   | 0,986   | <b>-0,008</b><br>(0,0552)   | 0,884   |
| Orphélin de père/mère (ORPHEL)                                    | -----                      | -----   | <b>0,043</b><br>(0,0607)    | 0,481   | <b>0,056</b><br>(0,0579)    | 0,336   |
| Fréquentation préscolaire (F_PRESCO)                              | -----                      | -----   | <b>0,101</b><br>(0,0742)    | 0,173   | <b>0,058</b><br>(0,0687)    | 0,395   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)              | -----                      | -----   | <b>-0,270**</b><br>(0,0446) | 0,000   | <b>-0,270**</b><br>(0,0440) | 0,000   |
| Disponibilité manuels scolaires entre CP et CM2 (LIVR_TOT)        | -----                      | -----   | <b>0,107**</b><br>(0,0186)  | 0,000   | <b>0,099**</b><br>(0,0202)  | 0,000   |
| Disponibilité de manuels x double flux (LIVR_TOT x DBLFLUX)       | -----                      | -----   | -----                       | -----   | <b>-0,008</b><br>(0,0368)   | 0,833   |
| Disponibilité de manuels x multigrade (LIVR_TOT x MGRADE)         | -----                      | -----   | -----                       | -----   | <b>0,128**</b><br>(0,0395)  | 0,002   |
| Usage du français à domicile (F_DOMICI)                           | -----                      | -----   | <b>0,088</b><br>(0,0983)    | 0,369   | <b>0,100</b><br>(0,0937)    | 0,285   |
| Niveau d'éducation du CM (EDUC_CM)                                | -----                      | -----   | <b>0,012</b><br>(0,0111)    | 0,270   | <b>0,041**</b><br>(0,0132)  | 0,003   |
| Niveau d'éducation du CM x rural (EDUC_CM x RURAL)                | -----                      | -----   | -----                       | -----   | <b>-0,062**</b><br>(0,0205) | 0,003   |
| Catégorie très pauvre (PAUVR)                                     | -----                      | -----   | <b>0,069</b><br>(0,0726)    | 0,349   | <b>-0,244~</b><br>(-0,2444) | 0,070   |
| Catégorie très pauvre x milieu rural (PAUVR x RURAL)              | -----                      | -----   | -----                       | -----   | <b>0,524</b><br>(0,3272)    | 0,115   |
| Catégorie très pauvre x DIEC # (PAUVR x IDE)                      | -----                      | -----   | -----                       | -----   | <b>0,037~</b><br>(0,0193)   | 0,063   |
| Chômage CM/père (CHOMAGE)   | -----                      | -----   | <b>-0,146*</b><br>(0,0695)  | 0,036   | <b>-0,149*</b><br>(0,0677)  | 0,028   |
| <b>Progression annuelle (ANNE)</b>                                |                            |         |                             |         |                             |         |
| Constante   | <b>0,039**</b><br>(0,0140) | 0,008   | <b>0,209*</b><br>(0,0931)   | 0,028   | <b>0,217*</b><br>(0,0942)   | 0,025   |
| Infrastructure physique de l'école (IDQIE)                        | -----                      | -----   | -----                       | -----   | <b>0,001</b><br>(0,0044)    | 0,760   |
| Age au CP (AGE_CP)  | -----                      | -----   | <b>-0,017</b><br>(0,0113)   | 0,133   | <b>-0,018</b><br>(0,0112)   | 0,111   |
| Fille x enseignante (FILL x NB_FEM)                               | -----                      | -----   | -----                       | -----   | <b>0,021*</b><br>(0,0089)   | 0,020   |
| Enfant confié (CONFIE)  | -----                      | -----   | <b>-0,031~</b><br>(0,0171)  | 0,069   | <b>-0,036*</b><br>(0,0180)  | 0,045   |
| Chômage CM/père (CHOMAGE)   | -----                      | -----   | <b>-0,041</b><br>(0,0338)   | 0,223   | <b>-0,044</b><br>(0,0328)   | 0,181   |
| Disponibilité manuels scolaires entre CP et CM2 (LIVR_TOT)        | -----                      | -----   | <b>-0,020**</b><br>(0,0056) | 0,001   | <b>-0,021**</b><br>(0,0051) | 0,000   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)              | -----                      | -----   | <b>-0,003</b><br>(0,0104)   | 0,757   | <b>-0,051**</b><br>(0,0187) | 0,009   |
| Nbre de redoublements x enseignant bachelier (NB_RED x NBR_BAC)   | -----                      | -----   | -----                       | -----   | <b>0,026**</b><br>(0,0090)  | 0,007   |
| Niveau d'éducation du CM (EDUC_CM)                                | -----                      | -----   | <b>0,008~</b><br>(0,0046)   | 0,082   | <b>0,007</b><br>(0,0045)    | 0,126   |

Notes :

~ =Significativité au seuil de 10% ; \* = Significativité au seuil de 5% ;

\*\* =Significativité au seuil de 1%.

# DIEC = Disponibilité des Infrastructures Economiques dans la Communauté.

Le tableau 7 donne les résultats des composantes de la variance résiduelle obtenus à partir des différentes étapes d'estimation. Le concept de « proportion de la variance expliquée » est bien connu dans les analyses de régressions multiples, sa mesure usuelle est le  $R^2$ . Cependant, pour les modèles hiérarchiques linéaires, ce concept est beaucoup plus complexe (Snijders et Bosker, 1999). La proportion de la variance expliquée étant définie comme la réduction relative de la variance résiduelle entre le « modèle non contraint » et le « modèle contraint », elle est obtenue, pour chaque niveau, par la formule suivante :

$$R_i^2 = \frac{Var_i(\text{modèle non contraint}) - Var_i(\text{modèle contraint})}{Var_i(\text{modèle non contraint})} \text{ où } i \text{ représente le niveau}$$

Dans le modèle 1 où l'« année » est l'unique variable explicative des apprentissages, seul 1,3% de la variance des résultats scolaires a été expliquée. Dans le modèle 2, sont intégrées les caractéristiques des élèves et de leur environnement familial dans l'explication des résultats scolaires des élèves. On note alors une baisse des variances potentiellement explicables dans tous les niveaux (de 0,373 à 0,372 pour le niveau intra-élève ; de 0,409 à 0,149 pour le niveau inter-élève et de 0,207 à 0,183 pour le niveau inter-écoles), ce qui traduit une hausse importante de la variance expliquée. Celle-ci est passée de 1,3 % dans le modèle 1 à 27,2 % dans le modèle 2. Enfin, le modèle final, qui intègre toutes les variables explicatives des trois niveaux, offre le pouvoir explicatif le plus élevé en permettant d'expliquer 41,3 % de la variance totale des résultats aux tests du PASEC.

Abordons maintenant l'interprétation des effets des facteurs déterminants des performances scolaires. Le tableau 8 indique l'influence des caractéristiques des élèves et de leur environnement familial sur leurs résultats scolaires. Les scores des élèves dépendent positivement de la disponibilité des manuels scolaires (livres de français et de calcul), du niveau d'instruction de leurs parents et de la situation socio-économique de leur famille.

En effet, les résultats de l'estimation révèlent que les enfants qui ont disposé de plus de manuels scolaires durant les quatre années d'études (du CP au CM1), ont obtenu en moyenne des résultats plus élevés que leurs camarades qui n'en disposaient pas (soit un écart de 9,9 %).



De même, plus le niveau d'instruction des parents est élevé, meilleurs sont les résultats de leurs enfants, en raison sans doute de la qualité de l'encadrement dont ces derniers bénéficient à la maison. On constate aussi que le score obtenu par les élèves au pré-test effectué en début de CP a une influence positive et significative sur leur niveau moyen. En revanche, les résultats de l'estimation indiquent qu'il n'y a pas de différences notables entre les scores des filles et ceux des garçons.

A l'opposé, l'âge a un impact négatif sur les résultats scolaires : les enfants les plus âgés ont, en moyenne, un score moyen moins élevé que celui de leurs camarades plus jeunes. Les enfants sont aussi pénalisés par les chocs économiques qui affectent leur famille. En effet, quand le père (ou le tuteur) perd son emploi et reste au chômage pendant au moins un mois, les résultats de ses enfants diminuent en moyenne de 14,9% (-0,149). En outre, toutes choses étant égales par ailleurs, l'enfant qui vit dans une famille pauvre a un score moins élevé que celui de son camarade qui vit dans une famille aisée, et donc bénéficie de meilleures conditions d'apprentissage (disponibilité de livres à la maison, radio, télévision, cours de mise à niveau à domicile, une meilleure santé...). D'autre part, le redoublement a un impact négatif sur les scores des élèves (-0,270). Les élèves qui redoublent plusieurs fois au cours de leur cursus primaire ont, en moyenne, des résultats plus faibles.

Le modèle permet aussi d'estimer la progression moyenne des apprentissages et les effets des facteurs dont elle dépend. Le taux annuel de progression est de 21,7% (0,217). Cependant, il est plus faible chez les enfants confiés à un autre ménage (3,6 points de moins) et chez ceux qui ont redoublé au moins une fois (5,1 points de moins), ou chez ceux qui disposent de beaucoup plus de manuels scolaires (2,1 points de moins). Ce dernier résultat s'explique par le fait que les élèves qui ont suffisamment de manuels scolaires ont, en moyenne, des résultats élevés et donc la différence entre les performances d'une année à l'autre est plus faible.

De nombreux facteurs relatifs à l'école et à la communauté ont été retenus. Le premier est le niveau d'instruction des enseignants. Il a un impact positif sur les apprentissages des enfants. D'après le modèle final, les élèves qui ont eu, dans leur parcours scolaire (du CP au CM1), plus d'enseignants ayant au moins le bac comme diplôme académique, ont en moyenne les meilleurs résultats.

Le deuxième facteur est l'organisation des écoles. En effet, l'existence d'un comité de gestion actif à l'école favorise les apprentissages des élèves de 23,1 % (0,231). Il est vrai qu'un comité de gestion actif contribue par divers canaux au bon fonctionnement de l'école.

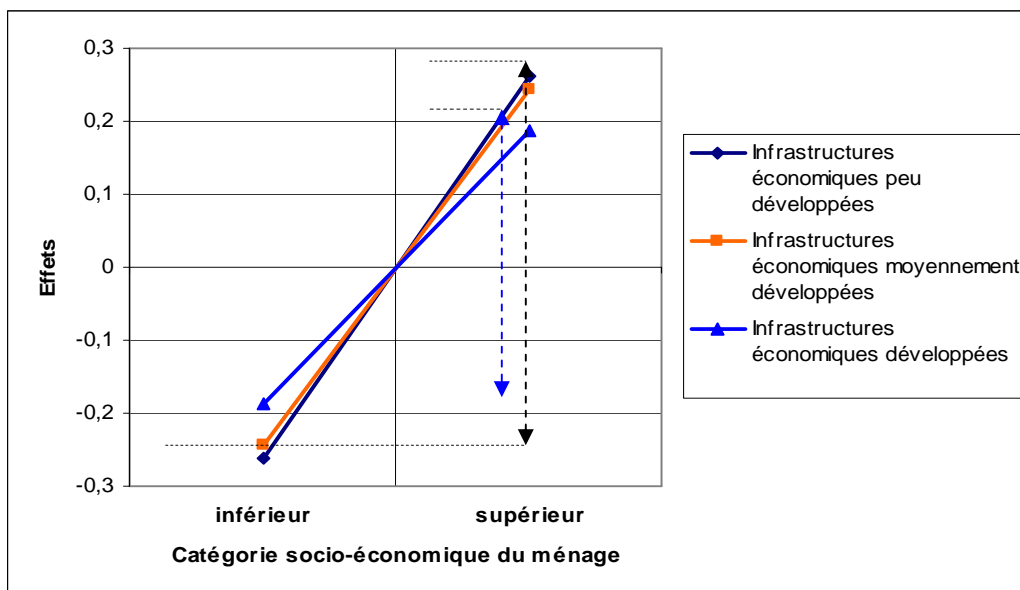
En ce qui concerne l'organisation pédagogique, l'existence de classes multigrades exerce une influence négative et très significative sur les performances scolaires. Le coefficient estimé montre que les scores des élèves qui fréquentent les classes multigrades sont de 61,6 % inférieurs aux scores de ceux qui sont dans une classe de flux unique. La pratique de la double vacation et l'appartenance au milieu rural ont un effet négatif mais non significatif sur les acquisitions scolaires des élèves. D'autres, comme le développement des infrastructures physiques de l'école et la disponibilité des infrastructures économiques dans la communauté, ont une influence positive mais non significative. Cependant, l'interaction de ces facteurs avec d'autres variables permet d'expliquer davantage certaines disparités de performances entre élèves.

En plus des effets directs, le tableau 8 présente aussi des effets d'interaction entre les variables explicatives des niveaux 2 et 3 sur les scores des élèves du PASEC. Trois interactions sont considérées.

***Influence conjointe du niveau de vie des ménages et des infrastructures économiques de la communauté sur les performances scolaires***

On a vu précédemment que les enfants qui vivent dans des familles aisées ont de meilleurs résultats, en moyenne, que leurs camarades qui vivent dans des ménages pauvres. Toutefois, l'influence du statut socio-économique de la famille sur les performances scolaires varie selon le niveau de développement économique de la communauté où elle réside (Graphique 3).

Graphique 3 : Impact combiné du niveau de vie et des infrastructures économiques de la communauté sur le score moyen des élèves



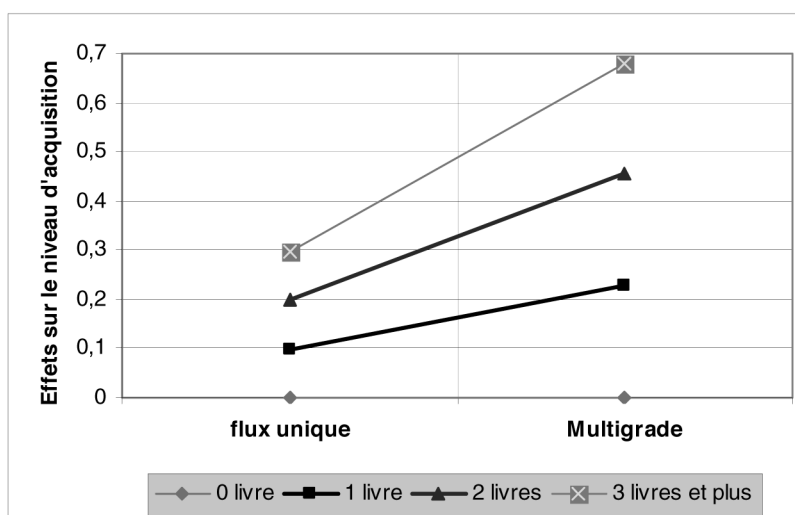
La construction du graphique suit la procédure employée par Lietz (1995) et par Mohandas (1999). Le calcul des coordonnées des effets d'interaction figure en annexe 2. Les disparités de performances induites par le niveau de vie des ménages (utilisation de l'électricité, existence de lignes téléphoniques, de systèmes de canalisation, d'évacuation des ordures et des eaux usées...) sont plus élevées dans les communautés à infrastructures économiques peu développées que dans celles à infrastructures économiques développées. En effet, dans les communautés disposant de peu d'infrastructures économiques, le score des élèves vivant dans des familles riches est supérieur de 52,1 % à celui des élèves habitant dans des ménages pauvres, contre seulement 37,5 % dans les communautés ayant beaucoup d'infrastructures. Ainsi, les élèves issus de familles pauvres bénéficient des externalités positives qu'offre la disponibilité des infrastructures économiques.

#### ***Effet de la disponibilité de manuels scolaires sur les performances scolaires selon la pratique du multigrade***

Comme l'indique le graphique 4, le nombre de manuels scolaires possédés par l'élève au cours de son cursus primaire influence très positivement ses apprentissages. Plus l'élève dispose de manuels scolaires, plus fort est l'impact sur ses résultats.

En outre, le graphique fait ressortir que la disponibilité de manuels scolaires est particulièrement bénéfique pour les élèves qui fréquentent les classes multigrades. Quel que soit le nombre de livres possédés, son impact sur les apprentissages est plus fort dans les classes multigrades que dans les classes à flux unique.

Graphique 4 : Influence de la possession de manuels scolaires sur les scores des élèves selon la pratique du multigrade



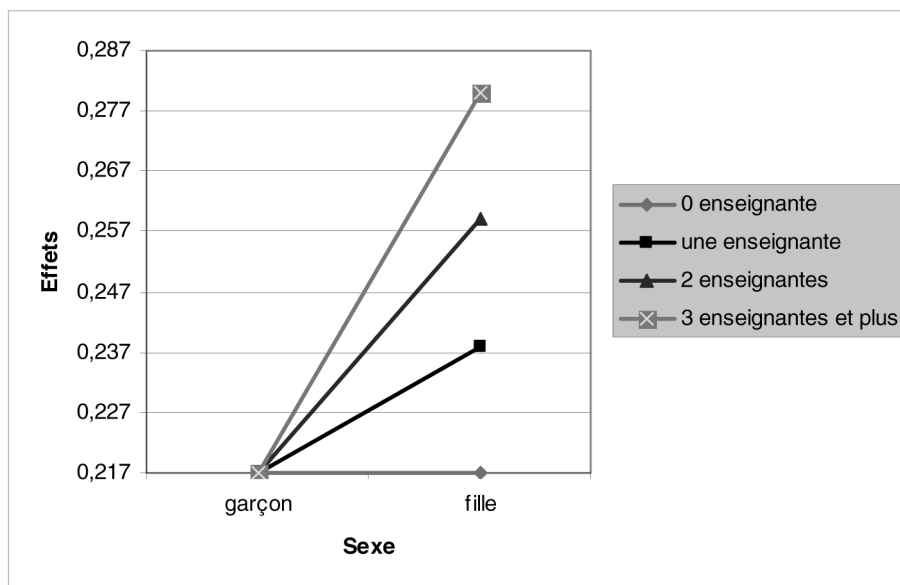
Etant donné que la pratique du multigrade a un effet négatif sur les performances scolaires (voir plus haut), une des solutions pour atténuer cet impact serait d'assurer la gratuité des livres et fournitures scolaires dans les écoles qui recourent à ce mode d'organisation pédagogique.

### ***Les différences d'apprentissage par sexe***

Le tableau 8 montre que l'introduction de la variable dichotomique « fille » ne révèle pas de différences significatives dans les acquisitions scolaires entre filles et garçons, même si le signe négatif du coefficient semble indiquer un avantage pour les garçons.

Pour mieux explorer l'impact du sexe sur les performances, nous nous sommes demandés si les caractéristiques de certaines écoles n'agissent pas davantage sur les performances des filles que sur celles des garçons. Ainsi, nous nous sommes intéressés à l'influence du sexe des enseignants sur les performances des filles. L'estimation donne alors un coefficient positif et significatif (Tableau 8). Comme l'indique le graphique 5, les filles progressent mieux, quand elles sont enseignées par des femmes que par des hommes.

Graphique 5 : Influence du sexe des enseignants sur les résultats scolaires des filles



En moyenne, c'est chez les filles qui ont eu, dans leur parcours scolaire, plus d'enseignantes que d'enseignants, que l'on enregistre les taux de progression annuelle des acquisitions scolaires les plus forts (jusqu'à 28 %). Ce résultat conforte l'idée selon laquelle les filles sont plus motivées quand elles sont enseignées par des femmes que par des hommes. L'explication pourrait être le rôle des enseignants et enseignantes comme modèles de leurs élèves et aussi la « discrimination » dont sont souvent victimes les filles dans les classes tenues par des hommes. Mingat et Suchaut (1998) ont abouti à des résultats similaires en montrant un effet positif des enseignantes sur la scolarisation des filles et sur la réduction des échecs scolaires.

## 4.2 Score en français et en mathématiques

La modélisation par discipline présente globalement des résultats similaires à ceux du modèle précédent. Nous allons donc ici nous intéresser uniquement aux effets propres à chaque discipline que le modèle général n'a pas détecté.

### 4.2.1 Les déterminants des apprentissages en français

En plus des effets déjà mis en évidence par le modèle général, l'estimation des scores en français fait apparaître d'autres influences importantes à prendre en compte (voir Annexe 1, Tableau A1). Par exemple, l'usage du français à la maison, qui n'était pas significatif dans le modèle général, le devient dans les résultats du modèle relatif aux performances en français. Les enfants qui vivent dans des familles où l'on parle français améliorent en moyenne leurs résultats en français de 13,3 % (0,133) par rapport à leurs autres camarades vivant dans des ménages qui ne font pas usage de cette langue.

En ce qui concerne l'environnement scolaire, l'existence d'une association de parents d'élèves active dans les écoles joue très positivement sur les résultats des élèves en français. En revanche, la pratique de la double vacation a un impact très négatif sur les performances des enfants en français. En moyenne, le modèle estime un écart de 11,4 % entre le niveau en français des élèves fréquentant des classes à double flux et celui de leurs camarades des classes à flux unique.

De même, le niveau moyen en français en milieu rural est inférieur de 87,2 % à celui du milieu urbain. En milieu rural, les enfants n'ont pas les mêmes opportunités d'apprentissage de cette langue que dans les villes (niveau d'instruction des parents plus élevé, usage plus fréquent du français à la maison, plus grand accès à la télévision et aux médias, enseignants plus qualifiés, existence de bibliothèques à l'école ou dans le quartier, niveau d'instruction de l'entourage plus élevé...).

#### **4.2.2 Les déterminants des apprentissages en mathématiques**

L'estimation du modèle relatif aux performances en mathématiques fait apparaître aussi des résultats qui méritent d'être soulignés. D'abord, on note que c'est en mathématiques que la valeur ajoutée annuelle globale est la plus importante. On enregistre une progression moyenne de 39,9 % (0,394) par an des acquisitions entre le CP et le CM1, contre 20,3 % (0,203) en français. Cela signifie que globalement, les élèves s'en sortent mieux en mathématiques qu'en français. Ensuite, il existe une influence positive et significative (mais faible) du score initial en français sur les résultats obtenus par les élèves en mathématiques. Si un élève a des difficultés en français, il ne lui sera pas aisé de répondre correctement aux questions qui lui sont posées en mathématiques. Par ailleurs, on observe dans ce modèle un impact positif (0,6 % de plus) des infrastructures physiques des écoles sur la progression des élèves en mathématiques.

#### **Conclusion**

Depuis 2000, l'éducation primaire au Sénégal a connu un bond spectaculaire. Le taux brut de scolarisation qui était de 68,3 % a atteint 82,5 % en 2005. De nombreux signaux indiquent que la qualité a légèrement progressé si l'on compare les résultats du PASEC en 2000 et ceux du SNER2 de 2004 qui affichent un taux de 50 % d'élèves qui sont parvenus à répondre correctement à la moitié des questions. Néanmoins, la qualité demeure faible et suscite une grande inquiétude. La définition d'une politique de relèvement de la qualité est devenue une priorité pour le Gouvernement.

La présente étude a cherché à identifier les facteurs qui agissent de façon significative sur le niveau d'acquisition des connaissances des élèves dans l'éducation primaire. L'appariement de trois bases de données a permis de disposer d'une large gamme de facteurs pouvant agir sur les acquisitions scolaires des élèves. Tenant compte de la nature hiérarchique des informations collectées et du fait que chaque élève a été évalué de la deuxième à la sixième année d'étude et dispose ainsi de cinq notes, un modèle de croissance à trois niveaux a été estimé.

Les résultats font apparaître une forte variabilité des performances scolaires et des différences marquées entre établissements, puisqu'elles expliquent une part relativement importante (20 %) de la variance du score moyen dans les deux disciplines sur lesquelles ont porté l'évaluation, à savoir le français et les maths. Les estimations ont aussi montré le poids relativement important des différences entre les élèves. Il ressort des résultats de l'étude que la politique éducative dispose de nombreux leviers pour relever sensiblement le niveau des apprentissages. Qu'il s'agisse de la disponibilité des manuels, de l'âge d'entrée à l'école, du niveau d'instruction des enseignants, de la proportion d'enseignantes, de l'organisation pédagogique, la politique éducative peut peser sur chacun de ces facteurs.

Par ailleurs, l'accent est mis sur l'effet positif que pourrait avoir sur la qualité de l'éducation, les politiques de lutte contre la pauvreté dans laquelle le Gouvernement s'est engagé. En effet, l'amélioration du niveau de vie des ménages, la réduction du chômage, et l'existence d'infrastructures économiques relativement développées dans les communautés contribuent significativement au relèvement du niveau et de la progression des acquisitions scolaires.



## Références bibliographiques

- Angrist, J. D. and Pischke, V. 1999. "Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quart. J. Econ.* 114 -2 : 533-75.
- Barcikowski, R.S. 1981. Statistical Power with Group Mean as the Unit of Analysis, *Journal of Educational Statistics*, 6(3): 267-285.
- Bradley, S. and Taylor, J. 1998. 'The effect of school size on exam performance in Secondary schools', *Oxford Bulletin of Economics and Statistics*, vol. 60, pp.291-324.
- Bryk, A.S. et. Raudenbush, S.W. 1992. "Hierarchical Linear Models : Application and data analysis methods".
- Buchman, C. 2000. "Family Structure, Parental Perceptions, and Child Labor in Kenya : What Factors Determine Who is Enrolled in School?" *Social Forces*.
- Case, A. and Deaton, A. 1999. "School Inputs and Educational Outcomes in South Africa," *Quart. J. Econ.* 114 -3 : 1047-84.
- Crahay, M. 1996. Tête bien faite, tête bien pleine, un recadrage constructiviste d'un vieux dilemme, *Perspectives*, 26 (1) : 55-84.
- Gamse, B., Conger, D., Elson, D., and McCarthy, M. 1997. Follow-up Study of Families in the Even Start In-Depth Study : Final Report. Cambridge, M.A : Abt Associates, Inc.
- Das, J., Dercon, S., Habyarimana, J., & Krishnan, P. 2004. "When can school inputs improve test scores? Policy Research Working Paper Series 3217, The World Bank.
- Entwisle, D.R. & Alexander, K.L. 1990. Beginning School math competence : Minority and majority comparisons. *Child development*, 61 : 454 – 471.
- Glewwe, P. 2002. "Schools and Skills in Developing Countries : Education Policies and Socioeconomic Outcomes," *J. Econ. Lit.*, Vol. XL: 436-82. University of Minnesota and the World Bank.
- Grisay, A. 1984. Les mirages de l'évaluation scolaire (1) - Rendement en français, notes et échecs à l'école primaire ? *Revue de la Direction Générale de l'Organisation des Études*, XIX, 5, pp. 29-42.
- Hanushek, E.A. 1994. "Making Schools Work : Improving Performance and Controlling Costs", Washington, D.C. : Brookings Institution.

- Haveman, R.H. and Wolfe, B.L. 1995. "The Determinants of Children's Attainments: A Review of Methods and Findings", *Journal of Economic Literature*.
- Hofman, D.A., & Griffin, M.A. 1997. Hierarchical linear models in organizational research : Cross-level interactions. *Research Methods Forum*, 2, 1-3. Retrieved February 10, 2003, from [http://www.aom.pace.edu/rmd/1997\\_forum\\_hlm\\_models.html](http://www.aom.pace.edu/rmd/1997_forum_hlm_models.html).
- Hoxby, C. 2000. "Does competition Among Public Schools Benefit Students And Taxpayers" *American Economic Review*, American Association, vol. 90(5) : 1209-1238.
- Lloyd, C.B. and Montgomery, M.R. 1997. "The consequences of unintended fertility for investments in children: Conceptual and methodological issues," in International Union for the Scientific Study of Population, *International Population Conference, Beijing, 1997*, Vol. 2. Liège, Belgium : 699-718.
- Mc Lanahan, S. and Sandfur, G. 1994. *Growing up with a Single Parent: What Hurts, What Helps*, Cambridge: Harvard University Press.
- Michaelowa, K. 2001. "Primary Education Quality in Francophone Sub-Saharan Africa : Determinants of Learning Achievement and Efficiency Considerations".
- Mingat, A. & Suchaut, B. 1998. "Une analyse économique comparative des systèmes éducatifs africains", Ministère français des affaires étrangères.
- Raudenbush, S.W., Bryk, A.S., Cheong, Yuk Fai, Congdon, R., du Toit, M. 2004. *HLM6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: Scientific Software International.
- Sahn, D.E. and Stifel, D. 2003. "Exploring Alternative Measures of Welfare in the Absence of Expenditure Data", *Review of Income and Wealth*, vol. 49, n°4, pp. 463-489.
- Snijders, T. & Bosker, R. 1999. "Multilevel Analysis : An Introduction to Basic and Advanced Multilevel Modeling".
- Todd, P. and Wolpin, K. I. 2003 : "On the Specification and Estimation of the Production Function for Cognitive Achievement," in *The Economic Journal*, Vol. 113, No. 485, F3-33.
- Willms, J.D., et Kerckhoff. A..C. 1995. "The challenge of developing new educational indicators", *Educational Evaluation and Policy Analysis*, vol. 17, n°1: pp. 113-131.

## ANNEXES

Annexe 1 : Estimation des résultats scolaires des élèves par discipline

Tableau A1 : Estimation des résultats en français des élèves du PASEC

| VARIABLES  | Modèle 1:                  |         | Modèle 2:                    |         | Modèle final:               |         |
|--|----------------------------|---------|------------------------------|---------|-----------------------------|---------|
|  | Coefficient                | P-value | Coefficient                  | P-value | Coefficient                 | P-value |
| <b>Niveau annuel</b>   |                            |         |                              |         |                             |         |
| Constante  | <b>-0,055</b><br>(0,0645)  | 0,397   | <b>-0,802**</b><br>(0,2306)  | 0,001   | <b>-0,525*</b><br>(0,2483)  | 0,039   |
| Rural (RURAL)  | ----                       | ----    | ----                         | ----    | <b>-0,872*</b><br>(0,4043)  | 0,035   |
| Comité de gestion actif à l'école (GEST_ACT)                       | ----                       | ----    | ----                         | ----    | <b>0,091</b><br>(0,0741)    | 0,226   |
| Double flux (DBLFLUX)  | ----                       | ----    | ----                         | ----    | <b>-0,114*</b><br>(0,0558)  | 0,046   |
| Développement des Infrastructures Economiques (IDE)                | ----                       | ----    | ----                         | ----    | <b>0,002</b><br>(0,0099)    | 0,857   |
| Infrastructure physique de l'école (IDQIE)                         | ----                       | ----    | ----                         | ----    | <b>0,001</b><br>(0,0138)    | 0,919   |
| Ecole à Cycle incomplet (CYCL_INC)                                 | ----                       | ----    | ----                         | ----    | <b>0,005</b><br>(-0,1230)   | 0,965   |
| Association des parents d'élèves Active (ASS_ACTI)                 | ----                       | ----    | ----                         | ----    | <b>0,11~</b><br>(-0,0641)   | 0,093   |
| Age au CP (AGE_CP)   | ----                       | ----    | <b>-0,06*</b><br>(0,0267)    | 0,026   | <b>-0,093**</b><br>(0,0266) | 0,001   |
| Age au CP × milieu rural (AGE_CP×RURAL)                            | ----                       | ----    | ----                         | ----    | <b>0,095*</b><br>(-0,0423)  | 0,025   |
| Fille (FILL)   | ----                       | ----    | <b>-0,030</b><br>(0,0251)    | 0,237   | <b>-0,029</b><br>(0,0252)   | 0,247   |
| Score initial en français (SF_INIT)                                | ----                       | ----    | <b>0,0278**</b><br>(0,0008)  | 0,000   | <b>0,028**</b><br>(0,0008)  | 0,000   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)               | ----                       | ----    | <b>-0,1224**</b><br>(0,0363) | 0,002   | <b>-0,229**</b><br>(0,0646) | 0,001   |
| Nbre de redoublements × enseignant bachelier (NB_RED × NBR_BAC)    | ----                       | ----    | ----                         | ----    | <b>0,0529*</b><br>(-0,0247) | 0,037   |
| Disponibilité manuels de lecture entre CP et CM2 (LECTURE)         | ----                       | ----    | <b>0,088**</b><br>(0,0195)   | 0,000   | <b>0,074**</b><br>(0,0261)  | 0,005   |
| Disponibilité de manuels de lecture × milieu rural (LECTURE×RURAL) | ----                       | ----    | ----                         | ----    | <b>0,029</b><br>(0,0361)    | 0,428   |
| Usage du français à domicile (F_DOMICI)                            | ----                       | ----    | <b>0,127203</b><br>(0,0796)  | 0,115   | <b>0,133~</b><br>(0,0761)   | 0,086   |
| Niveau d'éducation du CM (EDUC_CM)                                 | ----                       | ----    | <b>0,022*</b><br>(0,0089)    | 0,015   | <b>0,030**</b><br>(0,0093)  | 0,002   |
| Niveau d'éducation du CM × rural (EDUC_CM × RURAL)                 | ----                       | ----    | ----                         | ----    | <b>-0,026~</b><br>(0,0137)  | 0,062   |
| <b>Progression annuelle (ANNE)</b>                                 |                            |         |                              |         |                             |         |
| Constante  | <b>-0,040*</b><br>(0,0168) | 0,022   | <b>0,182</b><br>(0,1086)     | 0,100   | <b>0,203~</b><br>(0,1047)   | 0,053   |
| Milieu rural (RURAL)   | ----                       | ----    | ----                         | ----    | <b>-0,062</b><br>(-0,0468)  | 0,191   |
| Développement des Infrastructures Economiques (IDE)                | ----                       | ----    | ----                         | ----    | <b>0,191</b><br>(0,0042)    | 0,572   |
| Association des parents d'élèves Active (ASS_ACTI)                 | ----                       | ----    | ----                         | ----    | <b>0,055~</b><br>(-0,0295)  | 0,069   |
| Age au CP (AGE_CP)   | ----                       | ----    | <b>-0,024~</b><br>(0,0129)   | 0,067   | <b>-0,025*</b><br>(0,0127)  | 0,045   |
| Usage du français à domicile (F_DOMICI)                            | ----                       | ----    | <b>0,050</b><br>(0,0360)     | 0,163   | <b>0,053</b><br>(-0,0357)   | 0,139   |
| Niveau d'éducation du CM (EDUC_CM)                                 | ----                       | ----    | <b>0,011~</b><br>(0,0056)    | 0,052   | <b>0,009~</b><br>(0,0054)   | 0,090   |
| Enfant confié (CONFIE)   | ----                       | ----    | <b>-0,037*</b><br>(0,0175)   | 0,032   | <b>-0,040*</b><br>(0,0181)  | 0,026   |
| Disponibilité manuels de lecture entre CP et CM2 (LECTURE)         | ----                       | ----    | <b>-0,034**</b><br>(0,0103)  | 0,001   | <b>-0,034**</b><br>(0,0097) | 0,001   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)               | ----                       | ----    | <b>0,001</b><br>(0,0118)     | 0,957   | <b>0,001</b><br>(0,0118)    | 0,924   |
| Catégorie des Riches (RICHE)                                       | ----                       | ----    | <b>-0,004</b><br>(0,0259)    | 0,892   | <b>-0,054</b><br>(0,0472)   | 0,253   |
| Catégorie des Riches × milieu rural (RICHE×RURAL)                  | ----                       | ----    | ----                         | ----    | <b>0,058</b><br>(-0,0514)   | 0,259   |

*Effets aléatoires*

| <i>Composante de la variance résiduelle</i> | <b>Modèle 1:</b>           |         | <b>Modèle 2:</b>           |         | <b>Modèle final:</b>       |         |
|---|----------------------------|---------|----------------------------|---------|----------------------------|---------|
|   | Coefficient                | P-value | Coefficient                | P-value | Coefficient                | P-value |
| <b>Niveau 1</b>                             |                            |         |                            |         |                            |         |
| $V(\varepsilon_{ijk}) = E$                  | <b>0,443</b><br>(0,6654)   | ----    | <b>0,407</b><br>(0,6379)   | ----    | <b>0,407</b><br>(0,6381)   | ----    |
| <b>Niveau 2</b>                             |                            |         |                            |         |                            |         |
| $V(r_{0jk}) = R0$                           | <b>0,323**</b><br>(0,5686) | 0,000   | <b>0,051**</b><br>(0,2248) | 0,000   | <b>0,049**</b><br>(0,2214) | 0,000   |
| $V(r_{1jk}) = R1$                           | <b>0,001</b><br>(0,0385)   | >0,500  | <b>0,009*</b><br>(0,0930)  | 0,021   | <b>0,008*</b><br>(0,0915)  | 0,019   |
| <b>Niveau 3</b>                             |                            |         |                            |         |                            |         |
| $V(u_{00k}) = U_{00}$                       | <b>0,201**</b><br>(0,4485) | 0,000   | <b>0,044</b><br>(0,2087)   | 0,262   | <b>0,035*</b><br>(0,1873)  | 0,029   |
| $V(u_{06k}) = U_{02}$                       | ----                       | ----    | <b>0,002</b><br>(0,0439)   | >0,500  | <b>0,003</b><br>(0,0579)   | >0,500  |
| $V(u_{010k}) = U_{04}$                      | ----                       | ----    | <b>0,048</b><br>(0,2181)   | 0,189   | <b>0,042</b><br>(0,2059)   | 0,137   |
| $V(u_{10k}) = U_{10}$                       | <b>0,011**</b><br>(0,1045) | 0,000   | <b>0,012**</b><br>(0,1077) | 0,000   | <b>0,011**</b><br>(0,1025) | 0,000   |
| $V(u_{18k}) = U_{17}$                       | ----                       | ----    | <b>0,002</b><br>(0,0399)   | 0,265   | <b>0,002</b><br>(0,0407)   | 0,26    |
| <b>Variance expliquée</b>                   | <b>0,037</b>               |         | <b>0,342</b>               |         | <b>0,518</b>               |         |
| <b>Variance totale</b>                      | 1,0                        |         | 1,0                        |         | 1,0                        |         |
| Déviante du modèle                          | 6427,85                    |         | 5625,21                    |         | 5606,92                    |         |
| Nombre de paramètres estimés                | 9                          |         | 48                         |         | 63                         |         |

Tableau A2 : Estimation des résultats en maths des élèves Pasesc

*Effets fixes*

| VARIABLES  | Modèle 1:                  |         | Modèle 2:                   |         | Modèle final:               |         |
|--|----------------------------|---------|-----------------------------|---------|-----------------------------|---------|
|  | Coefficient                | P-value | Coefficient                 | P-value | Coefficient                 | P-value |
| <b>Niveau annuel</b>                                     |                            |         |                             |         |                             |         |
| Constante  | <b>-0,055</b><br>(0,0607)  | 0,371   | <b>-1,311**</b><br>(0,0581) | 0,000   | <b>-0,261**</b><br>(0,0695) | 0,000   |
| Rural (RURAL)  | -----                      | -----   | -----                       | -----   | <b>-0,028</b><br>(0,0386)   | 0,476   |
| Infrastructure physique de l'école (IDQIE)               | -----                      | -----   | -----                       | -----   | <b>0,011</b><br>(0,0092)    | 0,257   |
| Multigrade (MGRADE)                                      | -----                      | -----   | -----                       | -----   | <b>-0,131~</b><br>(0,0777)  | 0,097   |
| Fille (FILL)   | -----                      | -----   | <b>-0,030</b><br>(0,0315)   | 0,349   | <b>-0,033</b><br>(0,0310)   | 0,286   |
| Score initial en Math (SM_INIT)                          | -----                      | -----   | <b>0,026**</b><br>(0,0013)  | 0,000   | <b>0,026**</b><br>(0,0013)  | 0,000   |
| Score initial en Français (SF_INIT)                      | -----                      | -----   | <b>0,003**</b><br>(0,0011)  | 0,005   | <b>0,003*</b><br>(0,0011)   | 0,012   |
| Enfant confié (CONFIE)                                   | -----                      | -----   | <b>-0,012</b><br>(0,0407)   | 0,776   | <b>-0,011</b><br>(0,0409)   | 0,779   |
| Fréquentation préscolaire (F_PRESCO)                     | -----                      | -----   | <b>0,080</b><br>(0,0662)    | 0,226   | <b>0,068</b><br>(0,0660)    | 0,301   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)     | -----                      | -----   | <b>-0,061~</b><br>(0,0334)  | 0,068   | <b>-0,067~</b><br>(0,0339)  | 0,053   |
| Disponibilité manuels de calcul entre CP et CM2 (CALCUL) | -----                      | -----   | <b>0,066**</b><br>(0,0228)  | 0,005   | <b>0,064**</b><br>(0,0230)  | 0,006   |
| Niveau d'éducation du CM (EDUC_CM)                       | -----                      | -----   | <b>0,025**</b><br>(0,007)   | 0,001   | <b>0,024**</b><br>(0,0071)  | 0,001   |
| <b>Progression annuelle (ANNE)</b>                       |                            |         |                             |         |                             |         |
| Constante  | <b>0,111**</b><br>(0,0134) | 0,000   | <b>0,375**</b><br>(0,0261)  | 0,000   | <b>0,394**</b><br>(0,0292)  | 0,000   |
| Infrastructure physique de l'école (IDQIE)               | -----                      | -----   | -----                       | -----   | <b>0,006~</b><br>(0,0034)   | 0,091   |
| Multigrade (MGRADE)                                      | -----                      | -----   | -----                       | -----   | <b>-0,057~</b><br>(0,0311)  | 0,073   |
| Fille (FILL)   | -----                      | -----   | <b>-0,016</b><br>(0,0170)   | 0,337   | <b>-0,017</b><br>(0,0160)   | 0,282   |
| Score initial en Français (SF_INIT)                      | -----                      | -----   | <b>0,001**</b><br>(0,0004)  | 0,006   | <b>0,001*</b><br>(0,0004)   | 0,021   |
| Score initial en Math (SM_INIT)                          | -----                      | -----   | <b>-0,007**</b><br>(0,0005) | 0,000   | <b>-0,007**</b><br>(0,0005) | 0,000   |
| Enfant confié (CONFIE)                                   | -----                      | -----   | <b>-0,022</b><br>(0,0181)   | 0,233   | <b>-0,022</b><br>(0,0181)   | 0,221   |
| Chômage CM/père (CHOMAGE)                                | -----                      | -----   | <b>-0,057*</b><br>(0,0245)  | 0,021   | <b>-0,058*</b><br>(0,0248)  | 0,019   |
| Disponibilité manuels de calcul entre CP et CM2 (CALCUL) | -----                      | -----   | <b>0,005</b><br>(0,0104)    | 0,636   | <b>0,004</b><br>(0,0102)    | 0,700   |
| Nbre de redoublements entre le CP et le CM1 (NB_RED)     | -----                      | -----   | <b>-0,050**</b><br>(0,0169) | 0,004   | <b>-0,053**</b><br>(0,0172) | 0,003   |
| Fréquentation préscolaire (F_PRESCO)                     | -----                      | -----   | <b>0,056~</b><br>(0,0291)   | 0,052   | <b>0,049~</b><br>(0,0288)   | 0,086   |
| Catégorie des Riches (RICHE)                             | -----                      | -----   | <b>0,026</b><br>(0,0205)    | 0,200   | <b>0,022</b><br>(0,0208)    | 0,286   |

*Effets aléatoires*

| <i>Composante de la variance résiduelle</i> | <b>Modèle 1:</b>           |         | <b>Modèle 2:</b>           |         | <b>Modèle final:</b>       |         |
|---|----------------------------|---------|----------------------------|---------|----------------------------|---------|
|   | Coefficient                | P-value | Coefficient                | P-value | Coefficient                | P-value |
| <b>Niveau 1</b>                             |                            |         |                            |         |                            |         |
| V( $\epsilon_{ijk}$ ) = E                   | <b>0,404</b><br>(0,6357)   | -----   | <b>0,338</b><br>(0,5811)   | -----   | <b>0,337</b><br>(0,5809)   | -----   |
| <b>Niveau 2</b>                             |                            |         |                            |         |                            |         |
| V( $r_{0jk}$ ) = R0                         | <b>0,399**</b><br>(0,6319) | 0,000   | <b>0,072**</b><br>(0,2683) | 0,000   | <b>0,072**</b><br>(0,2687) | 0,000   |
| V( $r_{1jk}$ ) = R1                         | <b>0,001</b><br>(0,0223)   | >0,500  | <b>0,011</b><br>(0,1025)   | >0,500  | <b>0,011</b><br>(0,1027)   | >0,500  |
| <b>Niveau 3</b>                             |                            |         |                            |         |                            |         |
| V( $u_{00k}$ ) = U00                        | <b>0,167**</b><br>(0,4088) | 0,000   | <b>0,036**</b><br>(0,1889) | 0,000   | <b>0,032**</b><br>(0,1795) | 0,000   |
| V( $u_{010k}$ ) = U10                       | <b>0,006**</b><br>(0,0744) | 0,000   | <b>0,004**</b><br>(0,0597) | 0,005   | <b>0,003*</b><br>(0,0546)  | 0,041   |
| <b>Variance expliquée</b>                   | <b>0,040</b>               |         | <b>0,443</b>               |         | <b>0,580</b>               |         |
| <b>Variance totale</b>                      | 1,0                        |         | 1,0                        |         | 1,0                        |         |
| <b>Déviance du modèle</b>                   | 6299,07                    |         | 5165,93                    |         | 5160,36                    |         |
| <b>Nombre de paramètres estimés</b>         | 9                          |         | 26                         |         | 31                         |         |

**Annexe 2 : Effets d'interaction entre le niveau de vie des ménages et les infrastructures économiques de la communauté sur les performances scolaires**

Nous nous intéressons ici au calcul de l'effet conjoint du niveau de vie (variable du niveau 2) et des infrastructures économiques (variable du niveau 3) sur les résultats des élèves. Les autres effets d'interaction s'obtiennent de la même façon.

A partir de l'équation du niveau 2, l'effet du niveau de vie du ménage est donné par la relation suivante :

$$\pi_{0jk} = \beta_{00k} + \beta_{06k} (CLASSE4)_{jk} + r_{0jk} \quad (1)$$

A partir de l'équation du niveau 3, l'effet du niveau de développement des infrastructures économiques est donné par la relation suivante :

$$\beta_{06k} = \gamma_{060} + \gamma_{062} (INFRAS - E)_k + \mu_{06k} \quad (2)$$

On obtient une équation avec interaction en combinant les équations (1) et (2) :

$$\pi_{0jk} = \beta_{00k} + \gamma_{060} (CLASSE4)_{jk} + \gamma_{062} (INFRAS - E)_k (CLASSE4)_{jk} + r_{0jk}$$

Avec  $\beta_{00k} = \gamma_{000}$  (le niveau général moyen de l'ensemble des écoles) = 0 car non significatif (voir tableau des estimation des effets fixes) étant donné que les scores sont standardisés ;

$$\gamma_{060} = -0,244414 \quad \text{et} \quad \gamma_{062} = 0,036533$$

Ainsi :

$$\pi_{0,jk} = -0,244*(CLASSE4)_{jk} + 0,0365*(INFRAS\_E)_k(CLASSE4)_{jk} + r_{0,jk}$$

**Les effets sont obtenus de la façon suivante :**

La variable catégorie socio-économique a deux modalités : inférieure et supérieure, tandis que la variable « infrastructures économiques » en a trois : peu développées, moyennement développées, développées. Les interactions entre ces deux variables donnent six effets possibles.

- (1) Catégorie socio-économique inférieure et infrastructures économiques peu développées  
(PAUVRE = 1 ; INFRAS\_E=-1)

$$\pi_{0,jk} = -0,244*(1) + 0,0365*(-1) * (1) = -0,2605$$

- (2) Catégorie socio-économique supérieure et infrastructures économiques peu développées  
(PAUVRE = -1 ; INFRAS\_E=-1)

$$\pi_{0,jk} = -0,244*(-1) + 0,0365*(-1) * (-1) = 0,2605$$

- (3) Catégorie socio-économique inférieure et infrastructures économiques moyennement développées (PAUVRE = 1 ; INFRAS\_E=0)

$$\pi_{0,jk} = -0,244*(1) + 0,0365*(0) * (1) = -0,244$$

- (4) Catégorie socio-économique supérieure et infrastructures économiques moyennement développées (PAUVRE = -1 ; INFRAS\_E=0)

$$\pi_{0,jk} = -0,244*(-1) + 0,0365*(0) * (0) = 0,244$$

- (5) Catégorie socio-économique inférieure et infrastructures économiques développées  
(PAUVRE = 1 ; INFRAS\_E=1)

$$\pi_{0,jk} = -0,244*(1) + 0,0365*(1) * (1) = -0,1875$$

- (6) Catégorie socio-économique supérieure et infrastructures économiques développées  
(PAUVRE = -1 ; INFRAS\_E=1)

$$\pi_{0,jk} = -0,244*(-1) + 0,0365*(1) * (-1) = 0,1875$$