



HAL
open science

**La fonction de production éducative revisitée dans le
cadre de l'Education Pour Tous en Afrique
subsaharienne: des limites théoriques et
méthodologiques aux apports à la politique éducative**

Jean-Marc Bernard

► **To cite this version:**

Jean-Marc Bernard. La fonction de production éducative revisitée dans le cadre de l'Education Pour Tous en Afrique subsaharienne: des limites théoriques et méthodologiques aux apports à la politique éducative. Economies et finances. Université de Bourgogne, 2007. Français. NNT: . tel-00223023

HAL Id: tel-00223023

<https://theses.hal.science/tel-00223023>

Submitted on 29 Jan 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE BOURGOGNE
Faculté de Sciences économiques et de Gestion
Ecole Doctorale LANGAGES, IDÉES, SOCIÉTÉS, INSTITUTIONS, TERRITOIRES
(LISIT N°491)
IREDU (UMR CNRS 5225)

THÈSE

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ DE BOURGOGNE

Discipline : Sciences économiques (CNU 05)

Présentée et soutenue publiquement par

Jean-Marc BERNARD

Le 18 décembre 2007

**La fonction de production éducative revisitée dans le cadre de
l'Éducation Pour Tous en Afrique subsaharienne :
Des limites théoriques et méthodologiques aux apports à la
politique éducative**

Jury :

M. Jean BOURDON, Directeur de recherche, IREDU, UMR 5225, CNRS, Université de Bourgogne (directeur de thèse).

M. Marc DEMEUSE, Professeur à Faculté de Psychologie et des Sciences de l'Éducation, Institut d'Administration scolaire, Université de Mons-Hainaut

Mme Sylvie LAMBERT, Directrice de recherche, LEA-INRA et Paris School of Economics (rapporteur)

Mme Katharina MICHAELOWA, Professeur d'économie politique du développement et de la transition, Université de Zürich, Institut de science politique et Center for International and Comparative Studies (ETH/Université de Zurich)

Mme Marie-Claude PICHERY, Professeur, Laboratoire d'économie et de gestion (LEG), UMR 5118, CNRS, Université de Bourgogne.

M. Patrick PLANE, Directeur de recherche, CERDI, UMR 6587, CNRS, Université de Clermont-I (rapporteur)

L'université de Bourgogne n'entend donner aucune approbation ou improbation aux opinions émises dans les thèses. Ces opinions doivent être considérées comme propres à leurs auteurs.

Sommaire

Introduction générale	9
Chapitre 1 : L'analyse des déterminants des acquisitions scolaires : des débats théoriques aux analyses empiriques	15
1.1 La quête désespérée de la fonction de production éducative	17
1.2 La modélisation statistique des acquis scolaires : débats méthodologiques	33
Chapitre 2 : La problématique enseignante à l'école primaire en Afrique francophone : entre enjeux de scolarisation et qualité de l'éducation	60
2.1 Extension de la scolarisation et évolution du corps enseignant	62
2.2 Caractéristiques des enseignants et acquisitions scolaires : des convergences dans les résultats des recherches	76
2.3 Les caractéristiques des enseignants au cœur des enjeux de la scolarisation primaire universelle	95
Chapitre 3 : statut des enseignants et acquisitions à l'école primaire au Mali : faux problème et vraies questions	97
3.1 Protocole d'enquête et données disponibles	98
3.2 Statut et acquisitions scolaires : quelles évidences ?	108
3.3 Où se situent les marges de manœuvre pour l'amélioration de la qualité des apprentissages dans l'enseignement fondamental malien ?	121
3.4 Statut des enseignants et qualité des apprentissages : une question essentiellement politique	125
Chapitre 4 : l'amélioration de la qualité des apprentissages à l'école primaire en Afrique : les enjeux de la gestion des systèmes éducatifs	127
4.1 Le poids des disparités entre écoles et entre classes dans le processus d'apprentissage	129
4.2 De l'effet-classe à l'effet-école	145
4.3 Disparités entre écoles et gestion de la qualité des apprentissages	151
4.4 La gestion locale comme moyen de réduire les disparités entre écoles	162
Conclusion générale	164
 Annexes	 169
Bibliographie	266
Table des matières	278
Liste des annexes	280
Liste des tableaux	282
Liste des illustrations	285
Développement des sigles et acronymes utilisés	286

Remerciements

Un proverbe africain dit en substance que pour savoir où l'on va, il faut se souvenir d'où l'on vient. Aussi, je ne saurais entamer ces remerciements sans avoir une pensée pour ma famille et mes racines bourguignonnes.

Il est certain que ce travail n'aurait jamais abouti sans le soutien, la confiance et l'amitié de Jean Bourdon qui a été bien plus qu'un directeur de thèse tout au long de ces années.

Ce travail est indissociable d'une carrière professionnelle de près de 10 années en Afrique. La liste des personnes à remercier est donc longue et il m'est impossible de toutes les citer ici. J'ai une pensée particulière pour ceux avec qui j'ai travaillé au Cameroun, en Mauritanie et au Sénégal, ils et elles se reconnaîtront ici. J'ai beaucoup appris avec eux et cette thèse leur est redevable.

J'ai eu la chance de rencontrer Jean Coadou à mon arrivée au Cameroun en 1996 : il accompagna mes premières années africaines et m'entraîna dans ma première recherche sur le continent africain à travers l'évaluation des chantiers d'innovation pédagogique de l'enseignement primaire camerounais. Les voies impénétrables de la DRH du MAE ont voulu que je retrouve Jean quelques années plus tard en Mauritanie où ma rencontre avec mon compère François Robert fût à l'origine d'un ouvrage commun sur l'école moyenne africaine. Jean et François ont beaucoup contribué à ma réflexion sur les problématiques éducatives en Afrique, qu'ils en soient remerciés ici.

Mon arrivée au Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC) en 2001 fut l'opportunité de réaliser des études au cœur de ma problématique de thèse. Paul Coustère a joué un rôle déterminant dans cette étape. Outre la confiance et le soutien qu'il m'a accordés, il avait donné au PASEC des bases méthodologiques solides qui ont permis de mener la plupart des travaux utilisés dans ma thèse. Les « épaulés de géant » furent confortables et j'en remercie Paul. Ces presque cinq années au PASEC furent une expérience très riche mais parfois aussi très rude. Il aurait été impossible d'accomplir le travail réalisé sans une équipe soudée et déterminée. Mathieu Laroche, Alain Patrick Nkengne Nkengne, Odile Simon, Katia Vianou, Beifith Kouak Tiyab, Kenneth Hounbedji et El Hadj Ngom, le GS vous remercie. Je veux également rendre hommage aux équipes nationales du PASEC qui malgré le « pain sec » (surnom du PASEC) ont réalisé un travail remarquable (avec une mention particulière à Harouna Touré et Gérard Condat respectivement responsable des équipes Mali et Niger lors des études « sensibles » sur les enseignants contractuels). Naturellement, je ne peux ici omettre de remercier les « sages » du comité scientifique du

PASEC dont le soutien a été précieux au cours de ces années. Enfin, il faut remercier la CONFEMEN pour sa politique de libre diffusion des données du PASEC qui facilite grandement les travaux de recherche.

Je n'étais pas vraiment en terre inconnue au PASEC puisque celui-ci a été développé sur la base d'une étude réalisée au Togo par Jean-Pierre Jarousse et Alain Mingat, alors tous deux chercheurs à l'Irédu. J'avais eu la chance de les connaître lors de ma formation universitaire et nos chemins se sont souvent croisés depuis. Mes échanges réguliers avec eux au cours de ces années ont beaucoup contribué directement ou indirectement à ce travail, qu'ils en soient remerciés ici. Au-delà, mon attachement et mes relations avec l'Irédu tout au long de ces années ont joué un rôle important. L'Irédu fut incontestablement l'un des points de repère qui me permirent d'amener cette thèse à bon port. Un grand merci donc à toute l'équipe du laboratoire et aux trois directeurs qui se sont succédés depuis 1996. Comme tous les doctorants de la maison, je suis particulièrement redevable à Sylvie Gervreau et Bertille Tessé dont l'efficacité n'a d'égal que la gentillesse.

Je tiens à exprimer toute ma gratitude à Sylvie Lambert et Patrick Plane pour avoir accepté d'être rapporteurs de cette thèse ainsi qu'à Marc Demeuse, Katharina Michaelowa et Marie-Claude Pichery qui ont bien voulu être membres de mon jury.

« Il n'est de richesses que d'hommes »

Jean Bodin (1529-1596)

Introduction générale

La relation entre l'éducation et la productivité de l'individu, et donc ses gains, a été au cœur des préoccupations des économistes à l'origine du développement de l'économie de l'éducation (Mincer, 1958, Schultz, 1961 ; Becker, 1964). L'avènement de la théorie du capital humain est finalement la première théorisation d'un questionnement qui remonte à la préhistoire de la science économique. De nouvelles théories sont venues concurrencer la théorie du capital humain (théorie du filtre ou du signal, théorie des salaires d'efficience), mais aucune n'a réellement remis en cause la relation entre l'éducation et les gains des individus¹. Si cette relation au niveau individuel bénéficie d'un large consensus suite aux nombreux travaux empiriques réalisés (Psacharopoulos, 1994), elle apparaît plus difficile à confirmer au niveau macro-économique. C'est le paradoxe micro/macro cher à Pritchett (2001) qui souligne la difficulté de mettre en évidence une relation statistique forte entre éducation et croissance économique. Or, les nouvelles théories de la croissance, dites théories de la croissance endogène, dans la lignée des travaux de Lucas (1988), accordent à l'éducation une place tout à fait centrale dans l'explication de la croissance. Toutefois, il est possible que l'instabilité et la faiblesse de la relation s'expliquent par le fait que la très grande majorité des analyses macroéconomiques se concentrent le plus souvent sur les indicateurs de quantité de scolarisation en négligeant les différences de qualité. Ainsi, Hanushek et Kimko (2000) mettent en évidence la relation importante entre la qualité de la force de travail, mesurée par des scores de sciences et de mathématiques, et la croissance économique.

Des travaux récents mettent donc sur le devant de la scène la question de la qualité de l'éducation qui, comme le soulignent Hanushek et Woessman (2007), compte plus pour expliquer les différences de croissance de la productivité entre pays que les écarts en termes de nombre moyen d'années de scolarité ou de taux de scolarisation. De fait, la qualité de l'éducation, qu'on entend ici par la qualité des apprentissages scolaires, apparaît comme un enjeu majeur en termes de développement pour les pays les plus pauvres de la planète. Ce

¹ C'est plutôt le rôle de l'éducation dans cette relation, simple signal ou source d'amélioration de la productivité de l'individu, qui a été l'objet de remises en cause.

constat rejoint parfaitement les préoccupations et les objectifs de la communauté internationale formalisés par les Objectifs du millénaire et du Forum de Dakar (2000). Si offrir à tous les enfants une éducation de base de qualité relève du droit à l'éducation et se passe de toute justification économique, cela n'ôte pas pour autant l'intérêt que représente sa relation avec la croissance économique et donc le développement².

La conjonction de ces différents éléments fait que l'on accorde une importance de plus en plus grande à la qualité de l'éducation. Le mouvement est d'ailleurs mondial avec la montée en puissance d'évaluations internationales comme TIMSS ou PISA dont les résultats agitent parfois les opinions publiques mais contribuent également à la réflexion sur les politiques éducatives. Sur le continent africain aussi ce type d'évaluations s'est développé depuis le début des années 90, avec des programmes régionaux comme le PASEC et le SACMEQ, le premier portant sur les pays francophones et le second sur les pays anglophones. Il est intéressant de noter que si la qualité de l'éducation peut apparaître comme une question relativement complexe, une mesure s'est assez rapidement imposée, notamment sous l'impulsion de l'Association internationale pour l'évaluation du rendement scolaire (IEA). Il s'agit de la mesure des acquis cognitifs des élèves à travers des tests standardisés dans différentes disciplines. Si cette mesure n'épuise pas la question de la qualité de l'éducation, elle recouvre une dimension essentielle, celle de la qualité des apprentissages, notamment pour les premières années de scolarité (Hanushek, 1979). Toutefois, l'intérêt des économistes pour cette question n'est pas récent et ne se borne pas à connaître le niveau moyen des acquisitions scolaires dans les pays. En effet, à la suite du rapport Coleman (1966), les travaux sur les déterminants des acquisitions scolaires se sont rapidement développés. L'analogie à la fonction de production, par son caractère heuristique, s'est alors vite imposée, donnant naissance à la terminologie de « fonction de production éducative ». Le principal objectif de ces travaux consiste à identifier les facteurs les plus coût-efficaces afin d'essayer de déterminer la combinaison optimale d'intrants pour la production scolaire. Il faut bien reconnaître que ces travaux sont très majoritairement empiriques et que la fonction de production constitue plus une démarche d'analyse qui ne manque pas de vertus heuristiques qu'un véritable cadre théorique de référence. Cela évite d'ailleurs de s'interroger sur les limites de l'analogie qui importe tout de même dans un secteur essentiellement non marchand

² On peut d'ailleurs penser que la conviction que cette relation est déterminante pour le développement a influencé certains acteurs du développement.

un concept qui relève de la théorie microéconomique la plus épurée du fonctionnement des marchés en situation de concurrence pure et parfaite.

Les travaux sur les déterminants des acquisitions scolaires ont d'emblée été confrontés aux problèmes d'estimation à l'origine de nombreux débats suite au rapport Coleman. Les débats méthodologiques ont donc toujours été très nombreux dans ce champ de recherche et particulièrement sur la question des biais d'estimation. L'un des apports de ces débats est le recours aux modèles à valeur ajoutée qui se basent sur plusieurs mesures dans le temps des acquis des élèves. Cela implique un protocole d'enquête assez lourd avec au moins deux passages dans les écoles. Ce type d'études est en général l'apanage des chercheurs, à une exception près : le Programme d'Analyse des Systèmes Educatifs de la Confemen (PASEC). Depuis le début des années 90, ce programme réalise dans les pays d'Afrique francophone, des enquêtes avec une évaluation des acquis des élèves en début d'année et en fin d'année scolaire. La méthodologie suivie est issue d'une recherche de Jarousse et Mingat (1993) réalisée au Togo et directement inspirée de l'approche de la valeur ajoutée. Les données du PASEC, exceptionnelles dans le contexte des pays en développement et particulièrement dans le contexte africain, sont largement mobilisées dans ce travail. Elles permettent de mettre en perspective les résultats des recherches sur les déterminants des acquisitions scolaires, très majoritairement réalisées dans les pays développés, avec les problématiques spécifiques des systèmes éducatifs africains.

Les attentes sont énormes en Afrique qui reste le continent connaissant le plus de difficultés dans la poursuite des objectifs de l'Education pour tous, et notamment celui de la scolarisation primaire universelle. On estime, sur la base des données 2004-2005, que seuls une quinzaine de pays, sur les 43 pays considérés, seront susceptibles d'obtenir un taux d'achèvement du primaire supérieur à 90 % en 2015 (UNESCO-BREDA, 2007). A cela, il faut ajouter des niveaux de qualité des apprentissages extrêmement préoccupants dans certains pays. La recherche de nouvelles voies pour l'expansion de la scolarisation et l'amélioration de la qualité de l'éducation est donc clairement une priorité pour les responsables des systèmes éducatifs et les partenaires au développement. L'une des clés de la scolarisation primaire universelle en 2015, objectif emblématique du forum de Dakar, est de disposer d'ici 2015 de près de quatre millions d'enseignants en poste pour le seul secteur public dans les systèmes éducatifs africains. Le défi est considérable, tant sur les plans du recrutement et de la formation que sur le plan budgétaire. En Afrique francophone, où le salaire relatif des

enseignants fonctionnaires est relativement élevé, les états ont, depuis le début des années 2000, fait appel à de nouveaux types d'enseignants appelés, selon les pays contractuels, volontaires ou encore vacataires. Ces enseignants, souvent moins bien formés, moins expérimentés et surtout moins bien payés n'en finissent pas de susciter la polémique dans les débats sur les systèmes éducatifs africains. Non seulement ils constituent une remise en cause du statut de l'enseignant (il faut comprendre ici le statut social de l'enseignant), mais ils sont aussi accusés de contribuer à la dégradation de la qualité de l'éducation. On touche là au second défi majeur du continent africain : améliorer la qualité de l'enseignement qui apparaît à travers les différentes données disponibles très en deçà de ce qu'on observe dans les autres régions du monde. Il apparaît clairement que le profil des futurs enseignants, qui seront recrutés dans les années à venir, est au cœur de la problématique de l'Education Pour Tous en Afrique. Dire que cette question est sensible sur le plan politique est un euphémisme mais les enjeux sous-jacents n'en appellent pas moins des réponses claires. L'apport de la recherche en la matière, s'il n'est pas suffisant, est nécessaire pour éclairer le débat public.

Le premier chapitre tente de mettre en perspective les questions théoriques que soulèvent la fonction de production éducative et les problèmes d'estimations mis en évidence par la littérature empirique très abondante. Les problèmes sérieux que pose la fonction de production comme cadre d'analyse du processus d'acquisitions scolaires rendent la pauvreté de la littérature sur cet aspect d'autant plus étonnante. L'argument de la simple analogie revient souvent mais on constate l'influence de la notion de fonction de production sur l'interprétation des résultats. Ce n'est que récemment que de nouvelles pistes ont été explorées. A l'inverse, les travaux empiriques sur les déterminants des acquisitions scolaires sont foisonnants mais avec une forte prédominance, il est vrai, dans les pays développés et notamment aux Etats-Unis. L'absence de convergence des résultats entre études constitue l'un des problèmes marquant de ce type de travaux. Elle est souvent imputée à des problèmes méthodologiques des estimations. Si ces problèmes sont bien réels, ils ne sont peut être pas rédhibitoires. C'est ce que différentes analyses croisées tentent de mettre en évidence dans ce chapitre.

Le chapitre 2 plonge de plain-pied dans la problématique de la scolarisation primaire universelle et de ses implications en termes de recrutements d'enseignants en Afrique. Les sévères contraintes budgétaires que connaissent les états africains ont amené des évolutions sensibles de la composition de la population enseignante. L'émergence d'une nouvelle

catégorie d'enseignants recrutés par l'Etat mais n'appartenant pas à la fonction publique est le fait marquant du début des années 2000 en Afrique francophone. Cette évolution a tout de suite posé le problème de son impact sur la qualité de l'éducation, la nouvelle catégorie d'enseignants ayant été accusée de dégrader la qualité de l'éducation. Toutefois, la relation entre les caractéristiques des enseignants et les acquisitions des élèves s'avère relativement modérée dans la littérature, aussi bien dans les pays développés que dans les pays en développement, et notamment en Afrique.

Le chapitre 3 s'attache à estimer l'effet des enseignants contractuels sur la qualité des apprentissages à l'école fondamentale malienne. Une évaluation du PASEC, avec un protocole d'enquête spécifique fournit des données tout à fait exceptionnelles. Malgré ce protocole, on ne peut exclure totalement la possibilité de biais, dans la mesure où les enseignants contractuels ne se trouvent pas dans des écoles ayant tout à fait les mêmes caractéristiques que celles des enseignants fonctionnaires. Toutefois, l'impossibilité de démontrer une endogénéité et la robustesse des résultats obtenus permettent d'être relativement confiants. Les contractuels apparaissent à leur avantage dans cette comparaison avec les titulaires, notamment en 2^{ème} année. Toutefois, un constat important tient à la faible part de variance que permettent d'expliquer les différentes variables employées dans le modèle, ce qui tend à montrer qu'il faut chercher d'autres pistes pour l'amélioration de la qualité des apprentissages. Sur ce plan, les très fortes disparités entre classes expliquent une part bien plus importante de la variance. Il reste à savoir s'il s'agit d'une spécificité malienne.

Le chapitre 4 se consacre à cette question des disparités entre classes et entre établissements. Les fortes disparités entre classes sont avérées dans la plupart des pays africains avec, il est vrai, pas mal de variété. On note aussi avec toute la prudence nécessaire, que ces disparités semblent nettement moins marquées dans les pays développés. La littérature qualifie généralement cet effet-classe d'effet-maître mais cette assimilation est discutable. Il est clair que cette mesure en creux de ce qui se passe dans la classe laisse de nombreuses zones d'ombre et fournit peu de pistes pour la politique éducative. Toutefois, l'existence de fortes disparités entre classes et entre établissements constitue une indication forte. La gestion des systèmes éducatifs pourrait en effet constituer une réponse appropriée, si elle était en mesure d'identifier les établissements à problème et de leur apporter un soutien. Cela supposerait une évolution sensible des pratiques actuelles. Enfin notons que ce travail est orienté vers une approche de la fonction de production paramétrée, aussi le domaine de l'identification de la

production d'école par des méthodes non paramétriques d'optimisation, avec essentiellement les questions d'enveloppe type DEA, ne seront pas abordées dans ce travail.

La fonction de production éducative, ou plus exactement l'approche qu'on désigne sous ce terme, bien que décrite par certains chercheurs, se révèle utile au décryptage des systèmes éducatifs africains. Elle permet notamment d'analyser la relation entre les caractéristiques des enseignants, dont le statut, et les acquisitions des élèves. Elle permet surtout de mettre le doigt sur les grands enjeux en matière de qualité des apprentissages.

Chapitre 1 : L'analyse des déterminants des acquisitions scolaires : des débats théoriques aux analyses empiriques

La première étude à grande échelle visant à faire une analyse des déterminants des acquisitions scolaires a été réalisée en 1966³ et son principal auteur, James Coleman, est un sociologue. Les économistes n'étaient donc pas les pionniers dans ce domaine de recherche, d'autant que l'économie de l'éducation commençait seulement à émerger comme discipline à part entière. Pourtant, très rapidement, ils se sont montrés très productifs et très influents sur cette question des déterminants des acquis scolaires. Ainsi, la fonction de production éducative, en référence à la fonction de production de la microéconomie traditionnelle, est apparue assez rapidement dans le vocabulaire des chercheurs mais aussi assez paradoxalement. En effet, personne n'a vraiment considéré la fonction de production comme un véritable cadre théorique pour l'analyse de la production des systèmes éducatifs. Cependant, son influence est présente jusqu'à nos jours, un bref regard sur les titres des articles scientifiques dans ce domaine suffit pour s'en convaincre.

La première section de ce chapitre (1.1) s'interroge sur cette analogie à la fonction de production qui a dominé ce champ d'analyse. De façon intéressante, l'une des critiques les plus fréquentes tient à la spécificité du « produit éducatif » qui apparaît très éloigné du produit homogène de la concurrence pure et parfaite que requiert la fonction de production. D'emblée, c'est la non homogénéité du produit de l'éducation qui est au cœur des analyses et ce sont les différences de qualité de ce produit qu'on cherche à expliquer. Ceci conduit d'ailleurs à s'interroger sur les problèmes de mesure que cela pose et à se pencher sur les tests d'acquisitions généralement utilisés dans les études empiriques. Malgré l'inadéquation de la fonction de production et l'insuffisance notable de cadre théorique relevées par la plupart des chercheurs, la littérature économique comporte relativement peu de travaux sur cette question et encore sont-ils relativement récents. C'est à partir de la fin des années 1990 que quelques travaux ont commencé à ouvrir de nouvelles pistes de réflexion (Levin, 1997 ; Filmer et Pritchett, 1999 ; Akerlof et Kranton, 2002). Il faut dire que la multiplication des études

³ Equality of Educational Opportunity Study.

empiriques montrant la faiblesse voire l'inexistence du lien entre les moyens mis à disposition de l'éducation et les résultats en termes d'acquis scolaire a été plutôt stimulante. La fonction de production se révélant plus que jamais un cadre incapable d'expliquer les résultats des observations empiriques. La plupart de ces nouveaux travaux ont un point commun, ils mettent un accent particulier sur la dimension comportementale en abandonnant complètement la fonction de production éducative.

Si la fonction de production n'a jamais été vraiment considérée comme un cadre théorique en tant que tel, force est de constater son influence sur l'interprétation des résultats des études empiriques. Sinon, pourquoi s'étonnerait-on de la non convergence des résultats relatifs aux différents inputs entre les multiples études empiriques à travers le monde ? Seule une utilisation optimale des différents inputs comme le suppose la fonction de production pourrait aboutir à cette fameuse convergence. Certes, pour certains auteurs, l'absence de convergence des résultats des études empiriques peut en partie être due à différents problèmes méthodologiques qui se posent dans les études empiriques réalisées. C'est à cette question des problèmes méthodologiques de l'analyse empirique que s'attache la seconde section de ce chapitre (1.2). Il faut bien dire que la littérature empirique est bien plus abondante que la littérature théorique. Depuis l'étude pionnière de Coleman des centaines de travaux ont été réalisés et, d'ailleurs, les méta-analyses réalisées sur ces travaux ont nourri quelques controverses. La formalisation classique du processus d'acquisition met, entre autres, l'accent sur l'importance de la prise en compte des capacités innées de l'élève et de la dimension cumulative du processus d'apprentissage. Ceci amène à considérer les modèles à valeur ajoutée, et notamment les modèles de gains annuels, comme les plus appropriés pour la modélisation statistique. Ils n'en présentent pas moins des problèmes d'estimation qui laissent craindre l'existence d'un certain nombre de biais. Pour apprécier la sensibilité de ces modèles aux problèmes de biais, plusieurs estimations ont été effectuées en utilisant les données du PASEC de trois pays : Madagascar, Mali et Niger. Un modèle alternatif est aussi proposé se fondant sur l'évolution du rang des élèves pour contourner le problème d'endogénéité du score initial dans le modèle à valeur ajoutée. Enfin, le recours à la modélisation multiniveau pour les données scolaires vient compléter le panorama des différents défis que représente l'analyse statistique des déterminants des acquisitions scolaires.

1.1 La quête désespérée de la fonction de production éducative

La littérature sur la fonction de production éducative regroupe un ensemble assez vaste de travaux qui s'intéressent aux déterminants des acquisitions scolaires. La référence à la fonction de production est souvent qualifiée d'analogie par les auteurs. Toutefois, le cadre théorique sous-jacent aux analyses empiriques pose la question de l'interprétation des résultats qui a nourri de nombreuses controverses. Il n'est donc pas inutile de faire un bref détour par le concept de fonction de production de la microéconomie traditionnelle (1.1.1).

La spécificité du produit de l'éducation est certainement l'aspect qui a le plus suscité de questions voire de réserves chez les chercheurs. Si le caractère multidimensionnel de l'éducation soulève un problème de mesure, il convient aussi de s'interroger sur cette spécificité du produit éducatif comparativement aux produits plus classiques de la théorie économique (1.1.2).

Si l'on excepte les discussions autour du produit de l'éducation, les controverses autour des résultats de la littérature sur la fonction de production n'ont que tardivement débouché sur des évolutions du cadre théorique de référence. Ce dernier, loin d'être stabilisé, fait l'objet depuis quelques années de nouvelles réflexions axées sur le comportement des acteurs (1.1.3).

1.1.1 Le concept de fonction de production dans la théorie microéconomique

La fonction de production occupe une place importante dans la théorie microéconomique et plus particulièrement dans la théorie de la firme. C'est par analogie que les économistes ont utilisé ce terme dans l'analyse du fonctionnement des systèmes éducatifs. On peut néanmoins s'interroger sur les fondements de cette analogie notamment en raison du fait que la fonction de production se situe dans le cadre d'un marché fonctionnant en situation de concurrence pure et parfaite.

Dans l'étude du comportement de la firme, la théorie microéconomique considère que les entreprises sont soumises à des contraintes techniques : « *seules certaines combinaisons d'inputs permettent de produire une quantité donnée d'output et l'entreprise doit se limiter à des plans de production techniquement réalisables* » (Varian, 1992, p. 320). L'ensemble de

production renvoie alors aux différentes combinaisons d'inputs et d'outputs qu'autorisent les contraintes techniques. Un point qui, nous le verrons, a son importance est l'hypothèse sous-jacente d'homogénéité de chaque input et de l'output. La fonction de production correspond à la frontière de l'ensemble de production, elle « *décrit la relation qui existe entre les quantités utilisées des différents facteurs (inputs) et la quantité maximale du bien (output) qui peut être produite* » (Picard, 1990, p.128). Ainsi, la fonction de production ne rend pas compte de n'importe quelles relations mais seulement des plus efficaces. Sur le plan théorique, cela s'explique par le fait que les entreprises cherchent à maximiser leurs profits. Les inputs représentent un coût et il est donc logique, pour un coût donné, que les entreprises produisent l'output maximum. Cela est rendu possible par le fait que la fonction de production est issue de relations techniques connues qui découlent d'un processus technologique exogène.

Maximisation des profits, marché concurrentiel, processus technologique exogène, voilà qui semble bien éloigné de la réalité des systèmes éducatifs. Pourtant, c'est sur un autre aspect qu'insistent beaucoup d'auteurs pour expliquer la spécificité de la fonction de production éducative. C'est en effet l'hypothèse du produit homogène qu'implique la fonction de production classique qui est souvent remise en cause (Lemelin, 1998, Cohn et Geske, 1990, Hanushek, 1979).

1.1.2 Spécificité et mesure du produit de l'éducation

De nombreux auteurs soulignent que le produit de l'éducation se différencie d'un produit homogène tel qu'étudié classiquement par la théorie de la firme en microéconomie. Cohn et Geske (1990) vont plus loin et considèrent qu'il y a différents produits⁴. Ils citent cinq catégories : « basic skills », « vocational skills », « creativity », « attitudes » et « other outputs ». Lemelin (1998) parle plutôt de multiplicité du produit ; on pourrait parler aussi de produit multidimensionnel. Le produit de l'éducation serait donc composé de multiples facettes plus ou moins délicates à mesurer individuellement et quasiment impossibles à prendre en compte simultanément. Force est de constater que le concept de produit de l'éducation ressemble fort à une auberge espagnole où chacun amène sa propre définition. Ainsi, la définition du produit se confond avec les attentes tout aussi multiples de chacun.

⁴ Pour une discussion sur la question de l'estimation de plusieurs produits de l'éducation voir Hanushek (1979).

Naturellement, tout semble plus simple quand il s'agit de voitures, d'ordinateurs ou de machines à laver. Pourtant, il est bon de faire remarquer que chacun de ces produits présente aussi de multiples dimensions. En effet, un modèle de voiture se singularise par un ensemble de caractéristiques : vitesse de pointe, consommation d'essence, fiabilité, sécurité, esthétique, etc. Il ne serait pas nécessairement plus simple d'estimer une fonction de production qui tenterait de prendre en compte toutes ces dimensions, comment pondérer telle ou telle dimension et trouver une unité de mesure ? L'exercice serait assurément périlleux. C'est pourquoi la théorie de la firme en microéconomie se focalise sur un produit homogène et raisonne sur les quantités produites. La question de possibles différences de qualité⁵ des produits se trouve ainsi écartée de la réflexion alors qu'elle apparaît d'emblée centrale pour l'éducation. Si ce problème a pu être contourné en microéconomie, pour un temps au moins, c'est en partie lié au fait de l'existence d'une mesure du produit : le prix du marché. La monétarisation du produit constitue en effet une mesure *stricto sensu* qui est censée prendre en compte les différentes dimensions du produit. Toutefois, les travaux sur la différenciation des produits initiés par Lancaster (1966) ont montré dans ce domaine les insuffisances de l'approche microéconomique traditionnelle.

On voit se dégager, à travers la discussion précédente, trois aspects par lesquels se différencie nettement l'éducation. Tout d'abord, l'hypothèse de produit homogène est mise en cause et l'enjeu de l'analyse se situe clairement dans l'appréhension des différences de qualité. Alors que cette question est sensiblement de même nature pour les biens manufacturés, elle est apparue incontournable pour l'éducation. Ensuite, les caractéristiques qui définissent la qualité d'un produit industriel sont généralement observables, et même supposées observées (hypothèse d'information parfaite). Or, cela ne peut être retenu pour le produit de l'éducation pour la simple raison que le produit de l'éducation est incorporé à la personne et donc immatériel⁶ (Caner, 2000). Cette spécificité rend la question de la mesure très sensible. Or, c'est le troisième aspect, la mesure que fournit le marché (le prix) n'existe pas dans le domaine éducatif. La mesure est donc essentielle mais laquelle utiliser ? Il n'est

⁵ Pour l'International Organization for Standardization (ISO) : « La qualité est l'ensemble des propriétés et caractéristiques d'un produit, d'un processus ou d'un service qui lui confère son aptitude à satisfaire des besoins implicites et explicites ».

⁶ Cela a débouché sur des approches qui assimilent l'éducation à un service, on se réfèrera notamment sur ce point aux travaux de Gadrey. Notons que les approches économiques de la production de services peuvent, dans une vision forte, nier toute référence à une fonction de production, puisque la relation entre l'émetteur et le récepteur du service est considérée comme unique et non reproductible.

malheureusement pas de réponse entièrement satisfaisante à cette question. Même si on part d'une définition assez générale où l'éducation est considérée comme une activité de transmission de valeurs et de savoirs, le problème de la mesure reste délicat. Il n'a pas été possible de définir une mesure globale et exhaustive du produit de l'éducation. Ici, c'est le pragmatisme des chercheurs qui a prévalu, il est vrai en cohérence avec une vision dominante de l'éducation comme activité de transmission de savoirs. En effet, comme le souligne Hanushek (1979) le recours aux tests standardisés des acquis scolaires est devenu la pratique de référence pour mesurer le produit de l'éducation, même si des études ont été menées sur d'autres dimensions (attitudes, assiduité, poursuite d'études, etc.). Bien que les chercheurs soient peu prolixes pour expliquer cette direction prise par les recherches menées, un ensemble d'éléments d'explication plus techniques peut aussi être avancé :

- (i) La mesure des acquis des élèves, malgré ses limites, est un champ de recherche développé qui fournit des mesures relativement fiables dont on connaît les insuffisances⁷. Les outils (tests) sont facilement accessibles et utilisables.
- (ii) La mesure des acquis des élèves correspond bien à l'idée d'incorporation à l'individu qu'on associe à l'éducation. On identifie pour chaque individu des qualités différentes, ici des niveaux d'acquisitions. On est donc très proche du concept de capital humain tel que défini par Becker (1964).
- (iii) Lemelin (1998) rappelle aussi que les tests sont utilisés en interne dans les systèmes éducatifs, il s'agit donc d'outils courants pour la communauté éducative et en lien direct avec les objectifs fixés par le système éducatif (les tests se fondent sur les programmes scolaires). Les études basées sur des tests d'acquis suscitent donc généralement l'intérêt des acteurs et ont leur légitimité.

Il y a donc différentes catégories de raisons qui peuvent contribuer à expliquer l'engouement pour les tests d'acquisitions standardisés dont la représentation collective de l'éducation, où les enfants vont acquérir des savoirs, n'est pas la moindre.

Pour revenir à la question de la mesure, il ne faut pas conclure trop rapidement que le fait de limiter le produit de l'éducation à la dimension des acquis scolaires résout tous les problèmes de mesure. En effet, les débats autour des théories des tests sont très vivaces. Il semble d'ailleurs utile de faire un bref détour par ces questions, car elles éclairent les limites de la

⁷ Les premiers travaux de l'IEA (Association Internationale pour l'évaluation du rendement scolaire) notamment relatifs à l'élaboration de tests comparatifs internationaux datent de la fin des années 50.

mesure généralement utilisée dans les études se référant à la « fonction de production éducative ».

L'un des critères déterminants d'appréciation d'un test est sa fiabilité, c'est-à-dire la précision de la mesure qu'il permet. Le principe est que, si le test est répété, il devrait fournir sensiblement les mêmes résultats. La fiabilité est un prérequis à la discussion sur la validité qui s'attache à interroger l'écart entre ce qu'on souhaite mesurer et ce qu'on mesure réellement avec le test. Ainsi, pour Laveault et Grégoire (1997) « sans fiabilité ou, si l'on préfère, sans mesure précise, toute discussion sur la validité devient futile ». Or, pour être fiable, un test doit en théorie mesurer une seule et même chose (une aptitude scolaire comme la lecture par exemple). Ainsi, l'un des indicateurs de fiabilité les plus utilisés mis au point par Cronbach (1951), l'alpha de Cronbach⁸, suppose que plus les corrélations entre les items qui composent le test, pris deux à deux, sont fortes et plus ces items sont homogènes et mesurent la même chose. Par conséquent, l'unidimensionnalité apparaît comme une conditionnalité de la fiabilité d'un test. Certains tenants de l'approche par les compétences, importée du monde de l'entreprise, soulèvent ainsi les limites de la démarche. Pour De Ketele et Gérard (2005) les épreuves classiques « *permettent bien d'évaluer les ressources jugées nécessaires (savoir-reproduire et savoir-faire), elles ne permettent pas (ou peu) d'évaluer la faculté de mobiliser celles qui sont pertinentes pour résoudre des problèmes ou effectuer des tâches complexes* ». Les tests utilisés ne sauraient donc pas rendre compte de la complexité d'une compétence ou en tout cas pas suffisamment. Il convient toutefois de relativiser. En effet, l'expérience montre que des dimensions différentes peuvent être intégrées dans un même test sans remettre en cause sa fiabilité. Par exemple, le test de mathématiques de fin de 5^{ème} année de l'école primaire utilisé par le Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC) intègre plusieurs dimensions, dont la géométrie et la résolution de problème, ce qui ne l'empêche pas d'obtenir des alpha de Cronbach souvent supérieurs à 0,80 indiquant une très bonne fiabilité du test puisqu'on considère généralement qu'au-delà de 0,70 la fiabilité du test est convenable.

⁸
$$\alpha = \frac{J}{J-1} \left[1 - \frac{\sum s_j^2}{s_x^2} \right]$$
 où J est le nombre d'items, $\sum s_j^2$, la somme des variances des j items et s_x^2 la variance des scores totaux au test.

Tableau 1.1 : Fiabilité des tests de mathématiques de 5^{ème} année du PASEC

	Burkina Faso	Madagascar	Mauritanie	Tchad
Alpha de Cronbach	0,82	0,84	0,87	0,82

Source : calculs à partir des données PASEC

Les statistiques nous rappellent simplement que, si la distinction sur le plan pédagogique est bien claire, les connaissances dans un domaine d'acquisition ne sont pas totalement déconnectées avec celles de domaines connexes. D'autres difficultés sont cependant soulevées pour la mesure d'une compétence par les tests classiques et le débat autour de l'évaluation des connaissances/compétences est loin d'être clos. Il nous rappelle les limites des instruments de mesure utilisés et donc de la portée des analyses en termes de « fonction de production éducative » qui peuvent en découler.

La critique sur la mauvaise prise en compte des compétences par les tests standardisés recoupe quelque peu celle qu'adresse Hanushek (1979) aux tests standardisés. Il reproche à ces derniers d'être généralement construits sur la base de critères internes comme le pouvoir discriminant des questions et la fiabilité. Ces critères, selon lui, ne prennent pas en compte directement les connaissances ou les compétences valorisées par la société. Le trait est excessif, d'une part, les tests sont construits sur la base des programmes scolaires et intègrent donc les objectifs fixés à l'éducation par la société, d'autre part, le débat présenté auparavant montre que désormais cet aspect est bien au cœur de la réflexion. D'ailleurs, comme le soulignent De Ketele et Delory (2000), l'approche par les compétences est initialement issue du monde de l'entreprise. Néanmoins, cela révèle une problématique intéressante par rapport aux attentes relatives à ce que mesurent les tests. Pour Hanushek (1979), l'intérêt d'observer la performance du système éducatif est généralement lié à l'importance qu'on accorde à la scolarisation pour l'insertion future des élèves dans la société. Pour l'économiste, les effets de l'éducation sur les gains futurs, l'insertion professionnelle et la croissance économique constituent des objets d'étude privilégiés. De ce point de vue, il est nécessaire que les objectifs des tests soient en relation avec les performances en dehors de l'école. On attend donc des tests qu'ils mesurent des aspects directement pertinents pour l'insertion dans la société. On rejoint bien ici l'approche par les compétences et notamment l'approche par les compétences de base (De Ketele, 1993) qui définit un nombre très limité de compétences essentielles à la vie courante. Il y a toutefois ici un glissement qu'il convient de bien

considérer. Un test standardisé mesure des acquis scolaires quelle que soit la nature de ces acquis (savoir faire, savoir être ou compétence de base). Si l'intégration des compétences dans les tests permet la prise en compte d'aptitudes des élèves qui jusque-là étaient absentes, c'est un progrès. Toutefois, la tentation paraît grande ici de donner aux tests la responsabilité de prédire le devenir de l'élève dans la société. En quelque sorte, ils seraient la boule de cristal du pédagogue et de l'économiste. Il faut s'inscrire en faux contre cette vision des choses pour au moins deux raisons. Tout d'abord, si les tests peuvent identifier un potentiel, en termes de compétences, ils ne nous diront jamais rien sur son utilisation. Autrement dit, ce n'est pas parce qu'un test indique que tel élève possède telle compétence que celui-ci la mettra automatiquement en œuvre dans sa vie ultérieure. On peut même aller plus loin et affirmer que les élèves apprennent à faire des exercices selon l'approche par les compétences. Les tests nous indiquent ceux qui parviennent à mobiliser les connaissances nécessaires pour réussir les exercices, les compétences ainsi identifiées sont des compétences scolaires. La compétence réelle de la vie courante ne peut être mesurée au sens strict par des tests scolaires, sa mesure repose sur sa mise en œuvre dans la réalité. Il y a donc une différence de nature entre une situation-problème dans un exercice scolaire et un problème de même type dans la vie quotidienne⁹.

Ainsi, les tests font surtout le bilan de ce que les élèves ont appris à l'école et constituent à ce titre un outil privilégié de l'analyse du fonctionnement des systèmes éducatifs. Naturellement, on attend que ces acquis contribuent utilement au devenir de chacun même si on sait que beaucoup d'autres facteurs interviennent. On se rapproche ici des notions de capital humain général et spécifique de Becker (1964), même s'il serait plus approprié de parler de capital humain scolaire et extra-scolaire. Les tests demeurent donc avant tout des instruments de mesure de la production scolaire.

On peut bien sûr contester le fait qu'un seul produit de l'éducation soit pris en compte dans cette perspective. Après tout, des aspects importants comme la socialisation, la citoyenneté, etc. disparaissent ainsi du champ d'analyse. Cela dit, comme le rappellent Duru-Bellat et al. (2004) « *Quoi qu'il en soit, s'il est certain que l'opérationnalisation des "produits" de*

⁹ Cette question est commune à d'autres préoccupations des économistes ; ainsi en économie du travail, des théories majeures, déclinées par Spence et Thurow en particulier, ont souligné que le système éducatif apporte des éléments pour signaler ou filtrer les individus. Ces signalements ne peuvent en aucun cas garantir la maîtrise de compétences mais pour le moins diminuer le risque de l'incompétence.

l'école est nécessairement réductrice, on ne peut contester l'importance des acquis scolaires, même s'ils sont loin de résumer la totalité des objectifs poursuivis par l'éducation ». D'ailleurs, comme le souligne Hanushek (1979) ce constat est d'autant plus vrai qu'on s'intéresse à l'école primaire, c'est-à-dire aux premières classes de la scolarité où les acquis cognitifs sont éminemment prioritaires¹⁰.

Si la mesure du produit de l'éducation est complexe et soulève des difficultés méthodologiques, certaines insuffisances du cadre conceptuel de la fonction de production ont émergé de la discussion. Tout d'abord, l'hypothèse d'homogénéité de l'output qui doit être, *de facto*, abandonnée pour l'analyse des déterminants des acquisitions. Ensuite, l'hypothèse d'information parfaite apparaît inappropriée tant les problèmes d'information, entre autres liés à l'incorporation des acquis à l'individu, semblent prégnants en éducation. D'autres limites sont à considérer.

1.1.3 La remise en cause de la fonction de production éducative

Le point de départ des débats autour de la fonction de production éducative est la difficulté à réconcilier les résultats des travaux empiriques avec le cadre théorique présumé. Un numéro spécial de «The Economic Journal» (113 [485], février 2003) est consacré aux controverses de la littérature relative à la fonction de production éducative, avec en point d'orgue celle opposant Hanushek (1997, 2003) et Krueger (2003) qui, à partir des mêmes données, n'aboutissent pas aux mêmes conclusions. Leclercq (2005) constate que la principale conclusion qu'on puisse tirer de la littérature sur la fonction de production éducative est précisément qu'elle n'est pas conclusive. La lecture sous l'angle de la fonction de production éducative des résultats des travaux empiriques pose incontestablement de grosses difficultés.

Dans le prolongement, une source importante de débats aux Etats-Unis porte sur le paradoxe observé au cours des dernières décennies qui ont vu les dépenses éducatives augmenter sensiblement sans évolution notable des acquisitions des élèves. Hanushek (2003) pose clairement le problème : « Les tailles de classes ont diminué, les qualifications des enseignants se sont accrues, et les dépenses ont augmenté. Malheureusement, peu d'éléments

¹⁰ Il considère par ailleurs que si le poids relatif accordé à l'un des outputs est très différent de ceux des autres outputs, le problème lié à l'estimation de plusieurs outputs sera empiriquement insignifiant.

existent pour suggérer qu'un quelconque changement des résultats des élèves ait accompagné cette croissance des moyens consacrés aux écoles »¹¹. Il observe notamment que les résultats des élèves de 17 ans au National Assessment of Educational Progress entre 1969 et 1999 en lecture, mathématiques et science n'ont connu que des évolutions très limitées sans commune mesure avec la croissance des ressources. Ce type de constat n'est pas propre aux Etats-Unis, Gundlach, Woessmann et Gmelin (2001) signalent que ce phénomène a été encore plus marqué pour les pays de l'OCDE¹², ce qui les amène à parler de déclin de la productivité scolaire.

Ces résultats interpellent fortement sur la question de la relation entre les ressources et les résultats scolaires et dépassent ainsi largement le cadre des seuls pays développés. Ils apparaissent aussi comme une remise en cause de l'utilisation du concept de fonction de production dans le domaine éducatif. En effet, comme le rappelait Hanushek (1979) : « La distinction clé entre une « fonction de production » et toute autre description alternative des relations entre input et output réside dans la notion qu'elle représente l'output maximum possible pour les inputs donnés »¹³. Il rappelle ainsi que l'analogie de la fonction de production renvoie à la maximisation des profits des firmes sur un marché concurrentiel, chère à la microéconomie traditionnelle. Certains auteurs, comme Hoxby (2000), contestent ouvertement la légitimité et la pertinence d'appliquer un modèle conçu pour expliquer le comportement de firmes cherchant à maximiser leurs profits sur un marché concurrentiel au secteur éducatif. Pourtant, ce n'est qu'assez récemment que le cadre théorique de la fonction de production a fait l'objet de réflexions nouvelles.

L'analogie entre la maximisation des profits et la maximisation des acquisitions est la base d'une analyse intéressante de Pritchett et Filmer (1999). Les auteurs constatent que les règles qui gouvernent l'allocation des ressources éducatives ne sont que peu concernées par les résultats scolaires. En effet, les observations contredisent l'hypothèse d'une affectation des ressources qui viserait à maximiser les résultats scolaires. Ils en conçoivent une hypothèse

¹¹ Traduction de l'auteur, texte original : "Class sizes have fallen, qualifications of teachers have risen, and expenditures have increased. Unfortunately, little evidence exists to suggest that any significant changes in student outcomes have accompanied this growth in resources devoted to schools".

¹² Pour le cas français, on peut également se reporter au rapport de la Cour des comptes (2003) consacré à la gestion du système éducatif.

¹³ Traduction de l'auteur, texte original : "The key distinction between a « production function » and any number of alternative descriptions of input and output relationships is the notion that it represents the maximum achievable output for given inputs".

alternative où l'allocation des ressources vise à optimiser à la fois les résultats scolaires et l'utilité des enseignants, ceux-ci étant très influents dans le processus de décision. Ils concluent qu'il y aura un plus haut niveau de dépenses pour les facteurs que privilégient les enseignants. Si le produit marginal est décroissant avec le niveau d'utilisation du facteur, cela mènera à un produit marginal plus faible pour ce facteur. Pritchett et Filmer envisagent trois modèles théoriques de dépenses d'éducation basés sur ce modèle général. Le premier est basé sur la théorie de l'agence : les parents, ou les gestionnaires censés les représenter, ne connaissent pas les paramètres de la fonction de production. Les enseignants profitent de cette situation pour favoriser les facteurs qui ont leur préférence (à commencer par leur niveau de rémunération). Le 2^{ème} modèle considère que, dans la compétition pour l'affectation des ressources, les enseignants sont plus influents que les parents sur le plan politique ce qui va aussi se traduire sur l'allocation des ressources. Enfin, dans le 3^{ème} modèle, le décideur politique utilise l'affectation des ressources dans une stratégie pour renforcer sa base politique. Dans les trois modèles, l'affectation et l'utilisation des ressources vont être fortement influencées par les enseignants, ce qui induit une sur-dotation des facteurs qui ont la préférence de ces derniers. L'interprétation des résultats des études empiriques doit donc prendre en compte cet aspect, l'apport d'un facteur va dépendre du niveau d'utilisation auquel il est évalué. La relecture des études empiriques par les auteurs corrobore leur hypothèse. Toutefois, d'autres clarifications, pas nécessairement incompatibles, peuvent également expliquer les résultats observés. L'apport majeur de Pritchett et Filmer, qu'ils revendiquent, porte sur la nécessité d'une théorie comportementale pour comprendre les résultats empiriques¹⁴. On voit en effet que l'allocation des ressources n'étant pas optimale, l'interprétation des résultats doit tenir compte du niveau d'utilisation des différents facteurs. Dès lors, contrairement à ce que suppose la fonction de production éducative, on peut s'attendre à des résultats différents pour les mêmes facteurs en fonction du contexte, ce qui permet de réconcilier quelque peu le cadre théorique avec les résultats empiriques.

Il convient de souligner que cette analyse met en évidence une allocation non optimale des ressources pour la production d'acquis scolaires, en revanche, elle ne remet pas en cause la notion même de fonction de production. Ainsi, dans ce modèle, l'école produit toujours l'output maximum en fonction des facteurs dont elle dispose. On ne sort donc pas du cadre

¹⁴ "the production function is derived from technical, not behavioral relationships, one needs a behavioral theory to understand the results of empirically estimating a production function (...)" p.224.

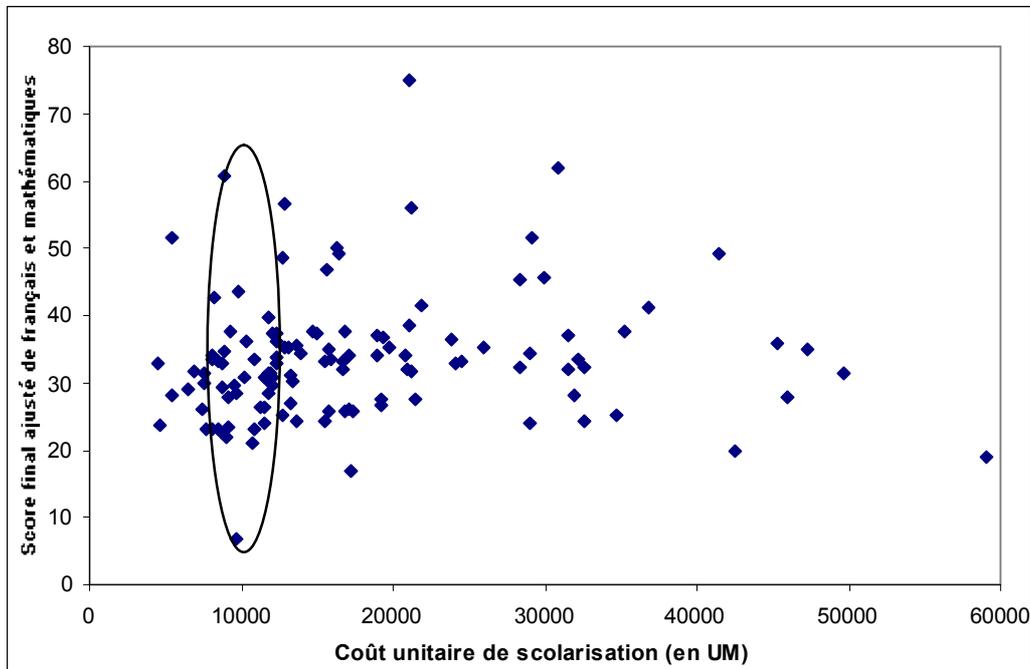
strict de la fonction de production. Le problème se situe au niveau de la dotation des différents facteurs qui sera fortement dépendante de l'influence des enseignants, ce qui l'écarte de l'optimalité revendiquée par la théorie. Si ce modèle est beaucoup plus en cohérence avec les résultats empiriques, il laisse toutefois une partie de ces résultats non expliquée. En effet, comme le rappellent Akerlof et Kranton (2002), on observe fréquemment de grandes différences de résultats pour des niveaux de ressources comparables. Les nombreuses analyses sectorielles menées récemment ont montré que le fait est particulièrement marqué dans les systèmes éducatifs africains¹⁵. Dans le graphique 1.1 (PASEC, 2006a), à partir d'une modélisation statistique des acquis des élèves en 5^{ème} année de l'enseignement fondamental (l'équivalent de l'enseignement primaire) en français et mathématiques une simulation a été faite du score qu'obtiendrait un élève moyen, dans des conditions moyennes, mais différencié selon la classe où il est scolarisé¹⁶. On met en suite en relation ce score ajusté avec le coût unitaire qui inclut les manuels scolaires, le coût annualisé de la salle de classe et le salaire de l'enseignant. Si la précision de la mesure peut toujours être discutée¹⁷, l'ampleur des écarts observés confirme qu'à ressources semblables, les établissements scolaires peuvent avoir des résultats très différents. Même sans considérer les cas extrêmes, les scores peuvent varier du simple au double d'une classe à l'autre, ce qui traduit des inégalités sévères en fonction du lieu de scolarisation. Les différenciations dans les ressources ne suffisent donc pas à expliquer la variété des résultats observés. Cet aspect ne saurait être oublié dans toute réflexion théorique sur le fonctionnement des systèmes éducatifs.

¹⁵ On se réfèrera dans ce domaine aux Rapports d'état des systèmes éducatifs nationaux (RESEN) initiés par la Banque mondiale. Plusieurs de ces rapports sont accessibles sur les sites de la Banque Mondiale : www.worldbank.org et du Pôle de Dakar : www.poledakar.org.

¹⁶ Cette information est produite grâce à l'introduction des variables muettes classe dans le modèle standard.

¹⁷ Plusieurs mesures du même type sont utilisées dans les analyses de secteur en éducation, le plus souvent ce sont les résultats aux examens d'une école qui sont mis en relation avec les coûts. Contrairement à la mesure proposée ici, cela ne permet pas de prendre en compte le niveau initial des élèves et rend la mesure plus sensible à des biais de sélection.

Graphique 1.1 : Score moyen ajusté de français et mathématiques et coût unitaire dans chaque classe de 5^{ème} année de l'enseignement fondamental de l'échantillon PASEC en Mauritanie



Akerlof et Kranton (2002), en cherchant à concilier les travaux sociologiques et économiques, proposent un modèle qui se focalise sur l'utilité de l'élève où l'identité de l'élève joue un rôle majeur. Dans cette perspective, un individu connaît des gains d'utilité quand ses actions ou celles des autres renforcent son image de soi. De plus, l'image de soi, ou l'identité, est associée avec un environnement social : les individus se pensent eux-mêmes et pensent aux autres en termes de différentes catégories sociales. Pour chaque catégorie, il existe un idéal correspondant à des stéréotypes physiques et comportementaux. Les individus gagnent ou perdent de l'utilité en fonction de la catégorie sociale à laquelle ils appartiennent, avec un haut ou bas statut social, et du comportement qui correspond à l'idéal de leur catégorie. Leur modèle prédit que l'utilisation des ressources risque d'être inefficace quand le background de l'élève est antithétique avec les valeurs académiques que les écoles devraient promouvoir, ce qui contraint les écoles à investir dans les questions d'identité de leurs élèves.

Si ce modèle a le grand mérite de rentrer plus avant dans la boîte noire de l'éducation, il présente plusieurs faiblesses dont certaines sont relevées par les auteurs eux-mêmes. La

prédominance de l'élève dans l'analyse laisse peu de place aux enseignants et à l'environnement scolaire en général. Akerlof et Kranton reconnaissent que la motivation des enseignants et des administrateurs sont aussi des éléments clés de la réussite des écoles. En éludant cet aspect, ils passent à côté de questions majeures pour la politique éducative notamment en termes de management des systèmes éducatifs. Par ailleurs, le modèle suppose une maximisation des compétences des élèves, ce qu'ils considèrent comme une faiblesse du fait que les objectifs de l'école ne se limitent pas aux compétences. On peut ajouter que les questions d'identité telles qu'elles sont abordées sont plus prégnantes chez des adolescents et de jeunes adultes mais moins prononcées chez les enfants plus jeunes. D'ailleurs, on peut se demander si le constat sur lequel les auteurs élaborent leur modèle, n'est pas avant tout un échec de l'école aux niveaux inférieurs. En outre, si les questions d'identité sont souvent au cœur des débats éducatifs, elles ne constituent certainement pas un cas général. Ainsi, le modèle d'Akerlof et Kranton n'apparaît pas comme un modèle général mais plutôt spécifique qui laisse de grandes interrogations notamment en matière de gestion scolaire.

La question de l'efficacité dans l'utilisation des ressources telle que décrite reste posée, il faut alors envisager que les écoles ne produisent pas le maximum d'output compte tenu des inputs disponibles. C'est la voie qu'a empruntée Leibenstein (1966) dans sa réflexion sur l'efficacité des firmes qui a abouti à la fameuse « x-eficiency ». Il souligne que les gains qu'on peut attendre d'une amélioration de l'efficacité dans l'allocation des ressources, qui est le propos de Pritchett et Filmer (1999), sont minimes par rapport à ceux qu'on peut attendre d'une amélioration de l'efficacité dans l'utilisation des ressources¹⁸. Leibenstein considère que la motivation des employés est un élément majeur, même si ce n'est pas le seul, de l'efficacité-X. Pour Levin (1997), l'approche peut s'étendre au secteur public en général¹⁹ et à l'éducation en particulier. En effet, comme le rappelle Simon (1991), la plupart des producteurs sont des employés des firmes pas des propriétaires. Ils n'ont donc aucune raison de maximiser le profit de la société sauf s'ils peuvent être contrôlés par les propriétaires. Ainsi, les firmes privées, les organisations à but non lucratif et les organisations bureaucratiques sont toutes confrontées

¹⁸ "The data suggest that in a great many instances the amount to be gained by increasing allocative efficiency is trivial while the amount to be gained by increasing X-efficiency is frequently significant." p. 413 Leibenstein (1966).

¹⁹ "Certainly, if dramatic changes in the organization of private enterprises can result in greater productivity, the concept is likely to be doubly applicable to schools where the usual problems of public sector production prevail. In the production of government services, outputs are often diffuse, incentives are demonstrably unrelated to productivity, production functions are uncertain, and standard operating procedures reign supreme." p.304.

au même problème, celui de faire travailler leurs employés dans le sens des objectifs de l'organisation. Il n'y a pas de raison, *a priori*, que ce soit plus facile (ou plus difficile) de motiver le personnel dans les organisations qui cherchent à faire des profits que dans les autres organisations.

Levin (1997), s'appuyant sur les travaux initiés suite à l'article de Leibenstein, retient cinq dimensions essentielles pour l'efficacité d'une firme. (i) La firme doit être claire sur ce qu'elle cherche à réaliser afin qu'il y ait une compréhension commune et un accord de tous les acteurs du processus productif sur les objectifs. En outre, des résultats mesurables doivent être disponibles. (ii) Il doit exister des incitations pour que les salariés poursuivent les mêmes objectifs que la firme. (iii) Les firmes ont besoin d'une information continue et systématique sur leur réussite pour voir si elles atteignent leurs objectifs. (iv) Pour faire face aux contextes évoluant rapidement les firmes doivent s'adapter et pour cela fournir de la formation continue et encourager l'initiative des employés pour résoudre les problèmes. (v) Les firmes doivent adopter les technologies les plus productives en fonction des contraintes de coût. Pour l'auteur, dans la plupart des écoles, on est très loin de remplir les conditions qu'impliquent chacune des cinq dimensions, ce qui suggère que des gains substantiels d'efficacité sont possibles. Levin présente un projet réalisé aux Etats-Unis, Accelerated Schools Project, qui s'est focalisé sur les cinq dimensions. Le projet a commencé avec 2 écoles pilotes en 1986-87 pour comprendre plus de 800 écoles élémentaires et collèges dans 39 états US en 1995-96. Les résultats enregistrés dans les écoles sont appréciables et laissent entrevoir des gains possibles d'efficacité tout à fait conséquents. Cela dit, comme le rappelle l'auteur, la principale difficulté est de savoir comment généraliser la démarche suivie dans ce projet.

La question de l'efficacité des écoles, au sens de Leibenstein, n'est pas vraiment nouvelle pour les économistes et elle souligne, à l'instar des autres travaux présentés, les limites du cadre théorique de la fonction de production. Il est toujours possible de dire que la fonction de production n'a jamais été un cadre théorique à proprement parler mais plutôt une référence commode, voire une métaphore, pour les travaux empiriques. Il est cependant permis d'en douter quand on regarde les controverses animées qu'a suscitées la littérature sur cette question. Il faut insister sur le fait que cette référence, implicite ou explicite, n'est pas neutre dans la compréhension des résultats des analyses empiriques. C'est, entre autres, pour cette raison que Pritchett et Filmer (1999) insistent sur la nécessité de disposer d'une théorie

« comportementale » dans l'analyse de la production éducative, suivis en cela par Akerlof et Kranton (2002) et Leclercq (2005).

1.1.4 Débats théoriques et enjeux empiriques

La littérature relative à la fonction de production éducative est très majoritairement une littérature empirique qui se réfère plus ou moins explicitement au concept de fonction de production de la théorie microéconomique. Les débats autour du cadre théorique qu'offre celle-ci sont finalement assez récents, si l'on excepte les discussions autour de la nature du produit de l'éducation. Si le produit de l'éducation présente des spécificités, notamment son incorporation à l'individu, sa nature ne semble pas fondamentalement différente de celle des autres produits auxquels se réfèrent traditionnellement les économistes. Pourtant, un certain nombre d'hypothèses inhérentes à la fonction de production dont l'homogénéité du produit et la perfection de l'information paraissent d'emblée contradictoires avec la démarche de l'analyse du produit éducatif. Ces problèmes ont débouché sur une attention accrue pour les questions de mesure signalant un certain pragmatisme des chercheurs qui se sont largement focalisés sur des travaux empiriques. La mesure du produit éducatif est incontestablement complexe, et la voie privilégiée par les recherches, à travers des tests standardisés des acquis scolaires, est nécessairement réductrice compte tenu des multiples facettes de l'éducation. Malgré tout, il est difficile de contester l'importance des acquis scolaires, notamment les acquis de base au cycle primaire, même si chacun s'accorde à reconnaître qu'ils ne sauraient résumer l'ensemble des objectifs assignés à l'éducation.

La littérature empirique, très abondante, s'est largement appuyée sur les tests standardisés des acquis des élèves. Il faut reconnaître qu'avec la multiplication des travaux, il est devenu de plus en plus difficile de concilier les résultats avec le cadre théorique de la fonction de production, ce qui a donné lieu à de nombreuses controverses. La relation incertaine entre les ressources affectées à l'éducation et les résultats obtenus a été au cœur des préoccupations des chercheurs. Assez récemment, des travaux ont proposé un élargissement du cadre théorique pouvant être appliqué à l'éducation. Ils ont notamment remis en cause l'allocation optimale des ressources (Pritchett et Filmer, 1999) mais aussi l'efficacité dans l'utilisation des ressources (Levin, 1997). Ces travaux tendent à concentrer leurs analyses sur le comportement des acteurs, on parle d'approche comportementale ou de « behavioral

economics » (Leclercq, 2005). Une première conséquence est que lorsque les résultats scolaires dépendent des comportements des acteurs, les recommandations de politique éducative sont beaucoup moins évidentes que lorsqu'il s'agit d'augmenter la dotation de tel ou tel facteur comme c'est le cas avec la fonction de production traditionnelle. Cela implique donc une nouvelle lecture des résultats empiriques.

En effet, les résultats des travaux empiriques doivent être considérés avec certaines précautions. Ainsi, l'influence d'un facteur va dépendre de sa dotation et de son utilisation effective. Pour reprendre l'exemple souvent utilisé des manuels scolaires, si tous les élèves disposent d'un manuel de lecture, il n'y a pas de progrès à attendre à fournir un second manuel de lecture. En revanche, si une partie des élèves ne dispose pas du manuel et qu'on n'observe pas d'effet significatif sur les acquisitions pour ceux qui en ont un, il faut alors s'interroger sur l'utilisation qui en est faite. Il se peut par exemple que l'enseignant n'utilise pas le manuel dans son enseignement car trop peu d'élèves en disposent. Les résultats de la modélisation doivent donc être questionnés pour aboutir à une interprétation correcte. Il résulte de ce qui précède qu'on ne saurait généraliser les résultats d'une étude, ces derniers ne valent qu'à contexte donné. Il faut donc se référer aux résultats des études avec précaution notamment quand on veut élargir les conclusions à d'autres contextes. Il est alors souhaitable de pouvoir s'appuyer sur la récurrence de résultats pour un même facteur pour envisager avec prudence une extension des conclusions.

Cependant, il faut signaler que les débats dans la littérature ne se limitent pas au cadre théorique ; bien au contraire, les problèmes méthodologiques rencontrés dans les estimations économétriques ont retenu depuis bien plus longtemps l'attention des chercheurs. Pour Todd et Wolpin (2003), l'une des raisons majeures pour expliquer les différences de résultats entre études empiriques réside dans les problèmes de spécification des modèles statistiques. Il s'agit donc d'un aspect essentiel à considérer quand on s'intéresse aux déterminants des acquisitions des élèves. La section suivante sera consacrée à cet aspect.

1.2 La modélisation statistique des acquis scolaires : débats méthodologiques

Les problèmes d'estimation ont toujours été au cœur des débats sur les fonctions de production éducatives bien plus que les aspects théoriques. Avec le fameux rapport Coleman, en 1966, les méthodes d'estimation ont fait l'objet d'âpres débats. Dans ce rapport, il ressortait que les facteurs scolaires avaient finalement une importance bien moindre que les facteurs extrascolaires dont l'environnement familial. Il est apparu que ce résultat pouvait être lié à la technique d'estimation utilisée. Les modèles à valeur ajoutée sont ensuite apparus peu à peu pour modéliser le processus d'acquis cognitifs (2.1). Dès l'origine donc, une attention accrue a été portée sur les méthodes d'estimation utilisées pour l'analyse des facteurs agissant sur les acquisitions scolaires. Ce qui a permis d'identifier progressivement les différents biais d'estimation potentiels qui reviennent régulièrement dans les débats méthodologiques (2.2).

Pour certains auteurs, comme Todd et Wolpin (2003), les différences de résultats constatées entre les études sont pour une large part imputables à des problèmes d'estimation. Beaucoup d'auteurs considèrent désormais qu'il n'y a plus beaucoup à attendre de l'estimation de fonctions de production éducative traditionnelles. Pourtant, si les biais potentiels des modèles explicatifs des acquis cognitifs sont désormais bien connus, en revanche peu d'études se sont penchées sur l'importance de ces biais dans les estimations empiriques. A partir des données du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC), nous essaierons d'apprécier la sensibilité des modèles classiques aux problèmes de variables omises et d'endogénéité (2.3).

1.2.1 Formalisation du processus d'acquisitions scolaires

Leclercq (2005) relève que l'économie n'a pas de théorie pour modéliser le comportement des élèves et des enseignants et commence à peine à s'intéresser aux aspects pédagogiques. La modélisation statistique du processus d'apprentissage est donc avant tout une démarche empirique. Le point majeur sur lequel s'accordent les économistes est que l'éducation est un processus cumulatif, ce qui implique de prendre en compte la dimension temporelle dans la modélisation statistique. En dehors de cet aspect, le principe de modélisation retenu est la mise en relation d'inputs avec un output qui est généralement les acquisitions scolaires de

l'élève. Nous reprenons ici la formalisation proposée par Hanushek (1979) qui est assez générale pour rendre compte de la plupart des estimations de fonctions de production éducative.

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i)$$

Où A_{it} est le niveau des acquis scolaires à la période t

$B_i^{(t)}$ est le vecteur des influences de l'environnement familial cumulées au temps t

$P_i^{(t)}$ est le vecteur des influences des pairs cumulées au temps t

$S_i^{(t)}$ est le vecteur de l'influence cumulative des inputs scolaires au temps t

I_i est le vecteur des capacités cognitives innées

Ce premier modèle suppose d'avoir des informations dont on ne dispose généralement pas sur l'historique des individus et des inputs scolaires. D'où l'intérêt potentiel du modèle à valeur ajoutée :

$$A_{it} = f^*(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, I_i, A_{it^*})$$

Où les inputs sont mesurés sur une période allant de t^* à t .

Deux points méritent d'être soulignés par rapport à ce modèle théorique. Le premier concerne la prise en compte dans le modèle des capacités cognitives innées de l'élève, qui en pratique se révèle très délicate faute de mesure pertinente²⁰ mais aussi contestable sur le plan scientifique. Le second point est relatif à la période sur laquelle est construit le modèle à valeur ajoutée. Dans la pratique, cela varie entre une année scolaire et plusieurs années, souvent un cycle d'enseignement (4 à 5 ans). Cependant, nous savons que plus la période de temps de la valeur ajoutée est importante plus il est difficile de recueillir toutes les données nécessaires. On retrouve alors le problème de disponibilité des données historiques. Il est vrai que les enquêtes sur plusieurs années tendent à devenir la référence dans les travaux américains (McCaffrey & al., 2003) mais elles sont coûteuses et extrêmement complexes à mettre en œuvre particulièrement dans des contextes de développement. Les modèles à valeur ajoutée portant sur une seule année scolaire, aussi appelés modèles de gains annuels, sont les

²⁰ La plupart des tests prétendant mesurer les capacités cognitives innées de l'enfant incorpore en fait des facteurs environnementaux (American Psychological Association, 1995).

plus simples à mettre en œuvre. Toutefois, ils risquent de sous-estimer des effets diffus dans le temps comme ceux des enseignants (McCaffrey & al., 2003). Nous nous référerons en priorité aux modèles de gains annuels dans ce qui suit.

1.2.2 Les problèmes d'estimation des modèles explicatifs des acquis scolaires : quid des modèles à valeur ajoutée ?

Plusieurs problèmes, bien connus, se posent dans l'estimation des modèles explicatifs des acquis scolaires. Tout d'abord, la formalisation présentée précédemment fournit une description du processus de production des acquisitions qui prend en compte sa dimension cumulative. L'analyse du processus ne dépend donc pas seulement des données contemporaines, mais nécessite des informations historiques rarement disponibles ou complètes (Hanushek, 1979, Leclercq, 2005, Glewwe & Kremer, 2006) comme les capacités cognitives innées de l'élève ou les informations sur l'ensemble du parcours scolaire de l'individu. D'emblée, l'estimation empirique est donc confrontée à l'insuffisance des données disponibles. Dans la littérature, l'accent est souvent mis sur l'absence de mesure des capacités innées qui aboutit à des biais dans l'estimation. Comme le rappellent Glewwe et Kremer (2006), une variable omise dans un modèle entraîne des biais dans les coefficients estimés quand elle est influente et reliée aux variables qui sont dans le modèle. Hanushek (1979) souligne que les capacités cognitives de l'élève et l'environnement familial sont corrélés positivement. Paxson et Schady (2007) montrent à partir de données équatoriennes que le développement cognitif des enfants est lié à la richesse du ménage et à l'éducation des parents²¹. Or, on dispose dans les études sur les acquis scolaires d'informations sur l'origine sociale de l'élève et au moins une variable est généralement introduite dans les modèles, alors qu'il n'y a pas de variable relative aux capacités cognitives antérieures de l'élève. En l'absence d'une mesure des capacités cognitives de l'élève, on doit s'attendre à un biais sur le coefficient de la variable d'origine sociale qui va surestimer l'effet de l'environnement familial puisqu'il tend à incorporer, en partie, l'effet des capacités cognitives antérieures de l'élève.

²¹ Leurs résultats suggèrent même un effet cumulatif de ces facteurs.

Un autre point fréquemment abordé dans la littérature porte sur l'absence ou, tout au moins, l'insuffisance de mesures historiques des variables exogènes et notamment des inputs scolaires. Les modélisateurs doivent se contenter le plus souvent de mesures contemporaines. Cela peut biaiser les coefficients de façon imprévisible quand les mesures contemporaines ne rendent pas correctement compte de l'historique. Un exemple parlant tient aux caractéristiques des enseignants. Rien ne garantit que les caractéristiques d'un enseignant lors de l'année de l'étude correspondent à celles de tous les enseignants que ses élèves ont eus pendant leur scolarité antérieure. L'estimation des coefficients peut alors aboutir à deux types de biais. Le premier, en l'absence de variable omise influente, va donner l'effet de la caractéristique de l'enseignant pour une seule année, celle observée, ce qui amène à sous-estimer l'effet par rapport à l'ensemble des années de scolarité antérieures. Il faut donc être vigilant dans l'interprétation que l'on fait du coefficient. Cependant, on peut aussi redouter l'influence d'une variable omise qui pourrait modifier le biais de façon imprévisible. Prenons l'exemple de l'ancienneté de l'enseignant. Dans les systèmes éducatifs, l'ancienneté est souvent un critère important pour la mobilité professionnelle des enseignants, et il est fréquent que les enseignants les plus expérimentés cherchent à rejoindre les meilleures écoles situées dans les zones perçues comme les plus favorisées. Si c'est effectivement le cas, le coefficient de l'ancienneté intégrerait le fait que, plus on est ancien, plus on a de bons élèves, mais cela serait en large partie imputable à la mobilité professionnelle et aux élèves plutôt qu'à la seule ancienneté de l'enseignant. Ainsi, sans contrôle du niveau initial des élèves, il est extrêmement délicat d'interpréter les coefficients associés aux caractéristiques des enseignants. On considère généralement que ce type de problèmes liés à la disponibilité d'informations historiques est plus prégnant pour les variables scolaires (Hanushek, 1979). Il est en effet probable que l'origine sociale de l'élève et l'environnement de l'école connaissent moins de fluctuations sur quelques années que les variables scolaires dont une partie varie chaque année. Il faut donc considérer que les coefficients des variables scolaires ont tendance à être plus biaisés que ceux des variables extrascolaires. Il convient toutefois de nuancer cette conclusion dans les pays pauvres où la fiabilité de l'information sur les caractéristiques socio-économiques des personnes est discutable (Glewwe & Kremer, 2006). Ainsi, dans une enquête scolaire, les données qu'on recueille sur les familles des élèves dépendent de l'élève lui-même et des informations disponibles sur ses parents au niveau de l'école. Dans les pays en développement, on ne peut ordinairement compter que sur les informations fournies par les

élèves qui sont d'autant plus imprécises que les enfants sont jeunes²². Il est particulièrement difficile d'obtenir une information précise sur les ressources des parents. Cela conduit à reconsidérer la qualité des données dont on peut disposer sur l'environnement familial de l'élève et donc l'imprécision de l'estimation qui peut en découler. Or, une erreur de mesure va mener à une sous-estimation de l'effet d'une variable si elle est aléatoire, ou à un biais imprévisible si elle n'est pas aléatoire. En outre, « *une variable mal mesurée « contamine » tous les coefficients estimés par les moindres carrés* » (Greene, 2005, p.83).

Ces problèmes d'estimation qui viennent d'être examinés sont parfois utilisés comme des arguments en faveur du modèle à valeur ajoutée. En effet, d'une part, l'introduction du score initial comme variable explicative dans un modèle peut permettre de prendre en compte, tout au moins partiellement, les capacités innées de l'élève et une partie des effets de son environnement familial. De ce fait, on peut en attendre une atténuation des biais engendrés par ces deux problèmes. D'autre part, le score initial, dans le cadre des modèles à valeur ajoutée, en limitant la période de temps considérée, permet de réduire les biais liés à l'insuffisance de données historiques omises notamment en ce qui concerne les inputs scolaires. Les modèles à valeur ajoutée sont donc généralement préférés par les analystes à des modèles plus simples qui ne prennent pas en compte le score initial de l'élève. Néanmoins, on voit bien que cette variable de score initial est une variable éponge censée capter beaucoup d'effets imputables à la période antérieure. Elle est par définition endogène et l'interprétation de son coefficient doit être très circonspecte. De plus, comme le rappellent Todd et Wolpin (2003), si l'endogénéité n'est pas prise en compte, le biais n'affectera pas seulement le coefficient du score initial mais peut-être aussi les autres coefficients. Or, le faible nombre de variables disponibles dans les enquêtes scolaires ne permet pas toujours de mettre en œuvre la méthode des variables instrumentales pour contourner les problèmes d'endogénéité. Si l'on ajoute l'éventualité de variables omises sur les inputs contemporains et les problèmes d'erreurs de mesure, on peut relativiser la fiabilité des estimations des modèles à valeur ajoutée. Il est clair que potentiellement ce type de modèle incorpore des biais, encore faut-il apprécier leur importance. Bien que l'on connaisse assez mal la sensibilité des modèles à valeur ajoutée à ces biais (McCaffrey & al., 2003), de plus en plus de chercheurs suggèrent leur abandon (Leclercq, 2005 ; Glewwe, 2002 ; Kremer, 1995). Il est toutefois souhaitable de

²² A titre d'exemple, le PASEC, pour évaluer les ressources du ménage, demande à l'élève des informations sur sa maison et les biens que possèdent ses parents.

ne pas tomber dans l'excès en rejetant purement et simplement ce type de modèle. Un argument, déjà évoqué, est qu'on ne peut pas facilement remplacer le modèle à valeur ajoutée, et surtout le modèle de gains annuels, par d'autres types de modélisation. Pour recourir à une modélisation plus performante, il faut disposer de données plus complètes sur le plan historique, ce qui implique des enquêtes sur plusieurs années beaucoup plus lourdes (Todd & Wolpin, 2003) et elles ne sont pas elles-mêmes sans poser de problèmes méthodologiques pour les analyses. En outre, il est très difficile et très coûteux de mettre en place de telles enquêtes particulièrement dans les pays pauvres. D'ailleurs, l'approche expérimentale (Duflo, 2006) n'échappe pas à cette critique même si le nombre d'études utilisant cette approche tend à s'accroître et compléter utilement les travaux existants. De plus, comme le rappellent Todd et Wolpin, si les résultats des études expérimentales sont utiles pour comprendre les effets d'interventions particulières, ils ne permettent pas de résoudre le problème de spécification de la modélisation de la production des acquis cognitifs. Il apparaît donc utile de tenter de mieux cerner les contours de ce que peuvent apporter les modèles à valeur ajoutée ainsi que leurs limites. Une mesure imparfaite est parfois préférable à pas de mesure du tout, à condition qu'on sache apprécier les marges d'erreur dans lesquelles on se situe.

1.2.3 La sensibilité des estimations des modèles de gains annuels : le cas des données du PASEC

Une démarche pragmatique invite à tenter d'identifier l'importance des problèmes de biais, et particulièrement ceux d'endogénéité, en s'appuyant sur des données appropriées. Pour cela, il est préférable de disposer de plusieurs enquêtes répondant aux besoins des modèles à valeur ajoutée et ayant une solide base comparative. La multiplication des résultats peut en effet rendre plus confiant dans les tendances observées. L'objectif ici est de procéder à une comparaison entre des modèles à valeur ajoutée et des modèles identiques à l'exception d'une variable, le score initial des élèves. Les données du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN se prêtent bien à ce type d'exercice. En effet, les données du PASEC ont l'avantage d'être collectées selon un même protocole dans les différents pays étudiés (Cf. annexe 1.2.1), ce qui permet de mettre en perspective les résultats obtenus dans plusieurs pays. Le choix a été arrêté à trois pays : Madagascar, Mali et Tchad. Le cas du Mali est spécifique dans la mesure où cette évaluation, qui visait à comparer les enseignants contractuels et les enseignants fonctionnaires, a donné lieu à un appariement entre les

classes²³ ce qui théoriquement limite les problèmes de biais liés à des variables de contexte omises. Les modèles auxquels nous allons nous référer concernent la 5^{ème} année de l'enseignement primaire. Il s'agit de modèles de gains annuels. Des données sont disponibles pour la 2^{ème} année, mais les problèmes de biais dus à l'insuffisance de données historiques sont supposés être plus marqués au terme de cinq années de scolarisation plutôt que de deux, ce qui apparaît plus intéressant dans l'examen de la sensibilité des modèles à ces problèmes. Par ailleurs, on utilisera pour la modélisation les moindres carrés ordinaires avec estimation robuste des intervalles de confiance (Cf. annexe 1.2.2).

Si les problèmes d'endogénéité sont importants, il faut s'attendre à observer une relative variabilité des coefficients lors de l'introduction de la variable de score initial dans les modèles qui peut être difficile à expliquer. Le caractère cumulatif du processus d'acquisition nous invite néanmoins à anticiper certaines variations des coefficients. On peut, d'une part, s'attendre à ce que les facteurs qui sont censés avoir une influence continue au cours de la scolarité, comme les caractéristiques individuelles de l'élève, aient un coefficient plus important dans le modèle sans score initial si l'influence est avérée. D'autre part, les facteurs qui ne jouent que lors d'une année scolaire, comme les caractéristiques de l'enseignant ou de la classe, devraient être plus influents dans le modèle à valeur ajoutée.

A la lumière des estimations réalisées (Cf. tableau 1.2.1), une première observation porte sur la grande variabilité du coefficient de détermination qui, selon les pays, est multiplié par un facteur deux à quatre quand on passe d'un modèle sans score initial (MSI) à un modèle à valeur ajoutée (MVA). En effet, la part de variance expliquée ne dépasse pas 10 % avec les MSI alors qu'elle peut dépasser les 30 % dans un MVA. Naturellement, de tels écarts suggèrent que la spécification avec le score initial devrait avoir la primeur pour peu que l'introduction de cette variable n'introduise pas de nouveaux biais dans l'estimation. Une question est donc de savoir comment déterminer si la variabilité des coefficients qui survient avec la prise en compte du score initial correspond à une diminution ou, au contraire, à une aggravation des biais ou à l'émergence de nouveaux biais dans le modèle. Dans la mesure où l'existence de variables omises reliées aux variables du modèle est une source de biais importante, une variable comme le score initial, qui permet d'augmenter sensiblement le

²³ Pour chaque enseignant contractuel, un enseignant titulaire était choisi dans une école voisine, aussi proche que possible, de l'école du contractuel et vice versa.

pouvoir explicatif, diminue donc nécessairement certains biais. D'un autre côté, le score initial est une variable endogène (Cf. annexe 1.2.6) et il est donc attendu qu'elle engendre des problèmes de colinéarités et/ou d'endogénéité dans le modèle. Si on peut détecter assez facilement les problèmes de colinéarité (Cf. annexe 1.2.3), il n'en va pas de même pour les questions d'endogénéité. Il faut donc s'interroger sur chaque variation de coefficient quand on passe d'un modèle sans score initial à un modèle à valeur ajoutée.

**Tableau 1.2.1 : Estimation de modèles avec et sans score initial
en 5^{ème} année d'école primaire dans trois pays du PASEC²⁴**

Variables	Madagascar		Mali		Tchad	
	M1	M2	M3	M4	M5	M6
L'élève						
est une fille	0,04 (0,04)	0,04 (0,04)	-0,19*** (0,05)	-0,09** (0,05)	-0,04 (0,06)	0,04 (0,06)
Age	-0,08*** (0,02)	-0,08*** (0,02)	0,02 (0,02)	-0,00 (0,02)	0,01 (0,02)	0,00 (0,02)
Appartient à une famille de niveau de vie élevé	0,11 (0,10)	0,17* (0,09)	-0,03 (0,09)	-0,01 (0,07)	-0,11 (0,13)	-0,18* (0,10)
a redoublé au moins une fois	-0,15** (0,06)	-0,08 (0,06)	-0,17* (0,09)	0,01 (0,07)	-0,10 (0,08)	-0,01 (0,06)
possède les manuels de français et mathématiques	0,12 (0,08)	0,10 (0,07)	0,17* (0,09)	0,04 (0,07)	0,21* (0,11)	0,13 (0,11)
Le maître						
Années d'ancienneté	0,02 (0,01)	0,02 (0,01)	-0,00 (0,01)	0,00 (0,01)	-0,01 (0,01)	-0,01 (0,01)
niveau second cycle secondaire	0,29* (0,15)	0,12 (0,13)	0,05 (0,17)	0,07 (0,14)	0,04 (0,19)	0,01 (0,17)
titulaire du BAC	0,41 (0,28)	0,30 (0,24)	/	/	0,23 (0,24)	0,13 (0,22)
formation professionnelle initiale ≥ 1 an	0,18 (0,17)	0,04 (0,14)	-0,08 (0,19)	-0,14 (0,13)	-0,22 (0,19)	-0,18 (0,16)
La classe						
double flux	-0,28 (0,22)	-0,31* (0,16)	-0,45*** (0,14)	-0,40*** (0,12)	-0,71** (0,33)	-0,63* (0,34)
multigrade	0,32 (0,45)	0,55 (0,42)	0,25 (0,29)	0,04 (0,25)	0,04 (0,18)	0,07 (0,16)
Nombre d'élèves	0,00 (0,00)	-0,00 (0,00)	-0,00 (0,00)	-0,00 (0,00)	-0,01*** (0,00)	-0,01*** (0,00)
L'école						
est en zone rurale	-0,14 (0,14)	0,02 (0,12)	-0,11 (0,16)	-0,11 (0,13)	-0,45** (0,18)	-0,24* (0,14)
Score de début d'année (français et mathématiques)	/	0,43*** (0,06)	/	0,52*** (0,05)	/	0,52*** (0,05)
Observations	2063	2063	1761	1761	1235	1235
R ²	0,10	0,26	0,08	0,34	0,10	0,34

Erreurs types robustes entre parenthèses

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

²⁴ Les scores de début et de fin d'année utilisés dans ces modèles sont standardisés avec une moyenne de 0 et un écart-type de 1.

On observe dans les tableaux 1.2.1 et 1.2.2 que la fréquence et l'ampleur de la variation des coefficients sont relativement modérées. Les modèles apparaissent beaucoup plus stables qu'on aurait pu le supposer. Il n'y a pas de bouleversements majeurs avec la prise en compte du score initial dans les modèles. Il faut toutefois rentrer dans des discussions au cas par cas pour détecter d'éventuels problèmes. Pour cela, il peut être utile de distinguer trois grandes catégories de variables : (i) les variables caractérisant l'élève, (ii) celles caractérisant l'enseignant et (iii) celles caractérisant la classe et l'école. Ce découpage rappelle que les données d'enquête sur les acquisitions scolaires ont une structure hiérarchique qu'il faut prendre en compte sous peine d'erreurs sur les intervalles de confiance des coefficients. Sur ce point, la prise en compte du seul coefficient donne parfois une précision illusoire au regard des intervalles de confiance. A titre d'exemple, pour le modèle 1 (M1), l'effet de l'âge de l'élève au seuil de 95 % est compris entre -0,05 et -0,11 écart-type, ce qui est relativement précis. En revanche, pour le diplôme de niveau secondaire (NIVCYCLB) alors que le coefficient est de 0,29 écart-type, l'intervalle de confiance est compris entre 0 et 0,58 écart-type, ce qui est considérable. Cette différence tient au fait qu'il y a beaucoup moins d'observations pour les variables caractérisant les enseignants et les classes (de 108 à 134) que pour les élèves (de 1235 à 2063). Il s'agit donc d'une dimension à considérer quand on s'intéresse à la variation des coefficients d'un modèle. C'est aussi pour cela qu'il est indispensable d'avoir recours à des modèles permettant une estimation robuste des intervalles de confiance en prenant en compte la structure hiérarchique des données.

**Tableau 1.2.2 : Différences entre les coefficients des deux modèles
(valeurs absolues de MSI-MVA)**

Variabes	Madagascar	Mali	Tchad
L'élève est une fille	0	0,10	0,08
Age de l'élève	0	0,02	0,01
Niveau de vie	0,06	0,02	0,07
L'élève a redoublé au moins une fois	0,07	0,18***	0,09
L'élève possède les manuels de français et mathématiques	0,02	0,13*	0,08
Ancienneté du maître (en années)	0	0	0
Le maître a fréquenté le second cycle secondaire	0,17	0,02	0,03
Le maître est titulaire du BAC	0,11	/	0,1
Le maître a suivi une formation professionnelle initiale longue (1 an et plus)	0,14	0,06	0,04
La classe fonctionne en double flux	0,03	0,05	0,08
La classe est multigrade	0,23	0,21	0,03
Nombre d'élèves dans la classe	0	0	0
L'école est en zone rurale	0,16	0	0,21

* La différence est significative au seuil de 10 %

*** La différence est significative au seuil de 1 % (Cf. annexe 1.2.3)

Les variations des différents coefficients ont été résumées dans le tableau 1.2.3 pour donner une vision d'ensemble. On constate qu'il y a seulement deux différences de coefficients qui sont statistiquement significatives, toutes deux dans le modèle du Mali. Cela montre avant tout la faible précision des estimations puisque des écarts supérieurs à 20 % d'écart-type n'apparaissent pas significatifs pour les variables de niveau classe. Il est malgré tout intéressant d'essayer d'analyser les différentes variations observées pour juger les coefficients fournis par les MCO. Trois possibilités ont été considérées. Tout d'abord, une absence de variation des coefficients qui regroupe les coefficients non significatifs dans les deux modèles et les coefficients significatifs qui ne varient pas ou peu (<0,05 écart-type) d'un modèle à l'autre. Ensuite, une variation qui va dans le sens attendu sur la base de la dimension cumulative du processus d'acquisition énoncée plus tôt. Enfin, une variation qui ne correspond pas à ce qui est attendu.

Tableau 1.2.3 : Tableau récapitulatif de la variabilité des coefficients

Variables	Absence de Variation	Variation attendue	Variation non attendue
L'élève est une fille	2	1	0
Age	3	0	0
Niveau de vie	1	0	2
L'élève a redoublé au moins une fois	1	2	0
L'élève possède les manuels de français et mathématiques	1	2	0
Ancienneté du maître (en années)	3	0	0
Le maître a fréquenté le second cycle secondaire	2	0	1
Le maître est titulaire du BAC	3	0	0
Le maître a suivi une formation professionnelle initiale longue (1 an et plus)	3	0	0
La classe fonctionne en double flux	0	2	1
La classe est multigrade	3	0	0
Nombre d'élèves dans la classe	3	0	0
L'école est en zone rurale	2	1	0
Total	27 (69,2 %)	8 (20,5 %)	4 (10,3 %)

On observe seulement quatre cas de variations non attendues dont deux concernent la même variable, à savoir la richesse de la famille. Il a déjà été signalé les difficultés relatives à la mesure de cette variable. Il faut ajouter ici que les mêmes critères ont été retenus dans les trois pays pour définir le niveau de richesse (Cf. annexe 1.2.4) ce qui peut être problématique compte tenu de l'hétérogénéité des contextes nationaux. Dans les modèles, le coefficient correspond aux élèves issus des familles les plus riches sur la base des déclarations des élèves (niveau de vie). Pour Madagascar, on remarque un coefficient positif non significatif dans le modèle 1 qui devient significatif au seuil de 10 % dans le modèle 2. Or, on s'attend bien sûr à avoir un effet cumulé plus marqué dans le modèle 1 que celui d'une seule année dans le modèle 2. Ce résultat pourrait être lié à un problème d'endogénéité.

Tableau 1.2.4 : Comparaisons des scores moyens²⁵ selon l'origine sociale des élèves à Madagascar (Cf. annexe 1.2.4)

	Elèves issus des familles les plus riches	Autres élèves	Différence
Score moyen de début d'année	0,017	0,035	-0,018
Score moyen de fin d'année	0,216	-0,05	0,266***

*** différence significative au seuil de 1 %

Toutefois, quand on regarde les scores moyens en début et en fin d'année pour chaque catégorie d'élèves (Cf. tableau 1.2.4), on constate que le coefficient du modèle 2 reflète bien leur évolution. La variation observée ne serait donc pas imputable à une endogénéité latente du modèle. Pour le Tchad, on remarque une variation similaire, mais avec cette fois un coefficient négatif (Cf. tableau 1.2.5). Naturellement, on s'attend à ce que le niveau de richesse soit corrélé positivement avec les résultats scolaires des élèves ; on peut donc s'interroger sur la pertinence de la variable de ressources des parents dans le contexte tchadien. Cependant, encore une fois, le coefficient reflète bien l'évolution du score entre le début et la fin de l'année éloignant le soupçon d'instabilité du coefficient en raison d'une endogénéité. Il semble donc que les variations suspectes que nous observons soient un problème de pertinence et de mesure de la variable plutôt que d'endogénéité. Cela nous renvoie au problème de l'information sur les caractéristiques socio-économiques des parents soulevé précédemment. Par ailleurs, ces différences de coefficients ne sont pas significatives.

Tableau 1.2.5 : Comparaisons des scores moyens selon l'origine sociale des élèves au Tchad (Cf. annexe 1.2.4)

	Elèves issus des familles les plus riches	Autres élèves	Différence
Score moyen de début d'année	0,294	-0,025	0,320***
Score moyen de fin d'année	0,079	-0,015	0,093

*** différence significative au seuil de 1 %

²⁵ Les scores sont standardisés avec une moyenne de 0 et un écart-type de 1.

Une autre variation inattendue est constatée pour Madagascar et concerne la variable relative au niveau de formation académique de l'enseignant, lorsque celui-ci a fréquenté le second cycle secondaire mais n'a pas obtenu son baccalauréat. On observe un coefficient de 0,29 dans le modèle sans score initial significatif seulement au seuil de 10 %, alors que ce coefficient est de 0,12 et non significatif dans le modèle à valeur ajoutée. Comme cela a été signalé, il faut être très prudent dans l'interprétation des coefficients relatifs aux caractéristiques des enseignants dans les modèles qui ne prennent pas en compte le score initial des élèves. Il est possible que le coefficient du MSI capte d'autres effets que celui propre à la formation académique du maître. On peut soupçonner que ces enseignants aient eu en moyenne des élèves de niveau plus élevé que leurs collègues, ce qui expliquerait le résultat observé. Il faut alors s'intéresser au niveau initial des élèves de cette catégorie d'enseignants en début d'année, c'est-à-dire avant qu'ils ne bénéficient de l'enseignement de ces maîtres.

Tableau 1.2.6 : Comparaisons des scores moyens selon la formation académique de l'enseignant à Madagascar (Cf. annexe 1.2.4)

	Elèves avec un enseignant de niveau second cycle secondaire	Autres enseignants	Différence
Score moyen de début d'année	0,101	-0,120	0,221***

*** différence significative au seuil de 1 %

La comparaison de moyenne (tableau 1.2.6) confirme cette hypothèse. Ainsi, c'est plus du côté des élèves que des enseignants qu'il faut chercher les meilleurs résultats des élèves dans des groupes pédagogiques tenus par des maîtres de niveau second cycle secondaire. On voit ici que les variables attachées aux enseignants peuvent induire des conclusions erronées dans les modèles où le score initial des élèves n'est pas pris en compte.

Un dernier résultat inattendu touche aux classes qui fonctionnent en double flux. Le coefficient qui leur est attaché varie peu (de -0,28 à -0,31) mais devient significatif, seulement au seuil de 10 %, dans le modèle à valeur ajoutée. On peut relativiser le caractère inattendu de ce résultat en considérant la mesure que représente la variable. En effet, l'effet cumulé sur plusieurs années d'un type d'organisation de classe devrait être plus marqué que

l'effet au cours d'une seule année scolaire. Cependant, la variable ne caractérise que l'année de l'étude ; aussi nous ne savons pas dans quel type de classe les élèves étaient les années précédentes, rien ne nous garantit qu'ils étaient dans des classes à double flux. Ainsi, comme pour les enseignants, le mode d'organisation de la classe, une année donnée, ne fournit pas une mesure historique satisfaisante. Les comparaisons de moyenne (Cf. tableau 1.2.7) confirment l'évolution décrite par le coefficient du modèle à valeur ajoutée. On peut donc considérer qu'en contrôlant avec le score initial des élèves, on dispose d'une estimation plus fiable de l'incidence du double flux sur les acquisitions scolaires.

Tableau 1.2.7 : Comparaisons des scores moyens selon l'organisation de la classe en double flux à Madagascar (Cf. annexe 1.2.4)

	Elèves en classe à double flux	Autres élèves	Différence
Score moyen de début d'année	0,019	0,033	0,014
Score moyen de fin d'année	-0,377	0,032	- 0,409***

*** différence significative au seuil de 1 %

Une attention particulière vient d'être accordée aux variables dont les résultats apparaissaient suspects comme c'est généralement le cas dans un exercice de modélisation. Cependant, on ne peut écarter l'hypothèse que les autres variables puissent aussi être biaisées et le coefficient proposé surestimé ou sous-estimé. En effet, il est possible qu'une variable omise affecte de façon sensiblement identique les deux types de modèle utilisés ce qui introduirait un biais qu'on ne saurait pas détecter. Il est alors très difficile de savoir dans quelle mesure le coefficient est biaisé en procédant aux mêmes vérifications que précédemment. Il y a donc clairement un problème de précision de l'estimation. Plus généralement, le problème de variables omises reste posé. Toutefois, les constats effectués avec la variable de score initial permettent de relativiser l'importance éventuelle de ces biais. Si l'absence d'une variable aussi influente et très liée aux autres variables du modèle que le score initial n'entraîne pas plus de changements que ce que nous avons pu observer, on peut être relativement optimiste sur ce point. Malgré tout, cela laisse planer un doute sur la précision des coefficients que nous fournit le modèle à valeur ajoutée. La présence d'une variable endogène comme le score initial au sein des variables indépendantes est en soi une source d'incertitude pour le modèle.

1.2.4 L'estimation alternative des modèles à valeur ajoutée : progrès ou régression ?

L'un des problèmes majeurs des modèles à valeur ajoutée intégrant la variable de score initiale dans les variables explicatives du modèle est le caractère endogène de celle-ci. Une voie pour contourner cette limite consiste à raisonner sur la progression des élèves, c'est-à-dire en faisant la différence entre les scores de fin d'année et les scores de début d'année des élèves. Ainsi, la variable de score initiale disparaît des variables explicatives ce qui évite les biais qui lui sont attachés sans pour autant perdre l'information essentielle qu'elle apporte. Toutefois, sur le plan méthodologique, si on s'en réfère à la théorie des tests (Laveault et Grégoire, 1997), cette pratique soulève de sérieux problèmes. En effet, deux tests différents correspondent à deux échelles de mesure qui ne sont pas directement comparables, on ne peut donc pas faire une différence des scores à chacun de ces tests. Si on effectue une différence entre les scores de deux tests distincts, on ne sait pas exactement ce qu'on mesure. Il faut donc ramener les deux tests à une échelle de mesure commune pour pouvoir effectuer une différence entre les scores. Ceci implique que les tests aient été élaborés spécialement pour cela et qu'une modélisation statistique spécifique soit mise en œuvre pour créer l'échelle commune (type modèles de Birnbaum ou de Rasch).

Il faut alors évoquer une seconde objection, plus radicale, par rapport à cette démarche. Supposons que nos deux scores soient ramenés à une même échelle, une autre difficulté se présente. En reprenant la terminologie utilisée dans la formalisation du modèle à valeur ajoutée :

$$A_{it} = f(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)}, A_{it^*})$$

$$\text{Soit } A_{it} = \alpha A_{it^*} + \beta_1 B_i^{(t-t^*)} + \beta_2 P_i^{(t-t^*)} + \beta_3 S_i^{(t-t^*)}$$

On voit que $A_{it} - A_{it^*} = f(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)})$ n'est pas une formalisation équivalente puisque cela suppose que $\alpha=1$, or les α mesurés dans les études oscillent généralement entre 0,4 et 0,6. On voit donc que les deux modèles sont différents ; ils renvoient à des conceptions différentes du processus d'apprentissage. On ne peut donc pas substituer l'un à l'autre.

Une autre approche de ce problème, mais qui n'est pas équivalente, serait de prendre en compte le rang des élèves dans l'échantillon à chacun des tests. On peut affecter à l'élève qui détient le meilleur score en début d'année le rang 1 puis regarder son rang au test de fin d'année. S'il est classé troisième, il aura perdu deux places. Il est ainsi possible d'effectuer une différence des rangs des élèves entre le test de début d'année et celui de fin d'année. On

compare en fait deux classements. La variable ainsi créée permet de mesurer la progression relative des élèves les uns par rapport aux autres et c'est cette variable que l'on va chercher à expliquer. On ne raisonne plus en termes d'acquis scolaires mais en termes de classements basés sur les scores aux tests, il s'agit d'une autre démarche de modélisation. L'ambition est moindre, on ne cherche plus à modéliser le processus d'acquisition des connaissances mais plutôt à modéliser l'évolution du classement des élèves au sein d'un même groupe entre deux tests. Certes, le classement est étroitement lié aux scores aux tests mais il ne mesure pas les acquis scolaires des élèves et n'est donc pas soumis aux mêmes exigences méthodologiques. On ne se réfère pas non plus directement à une théorie de la connaissance.

On peut retenir la formalisation suivante :

$$R_{it^*} - R_{it} = f(B_i^{(t-t^*)}, P_i^{(t-t^*)}, S_i^{(t-t^*)})$$

Où R_{it} est le rang de l'individu i au test de la période t

Le tableau 1.2.8 présente les principales caractéristiques de la variable créée selon cette démarche (Cf. annexe 1.2.5). Un premier constat est que la moyenne de la différence de rang entre le début et la fin de l'année n'est pas nulle. Cela vient du fait que deux élèves avec la même note ont le même rang, les changements de classement ne s'annulent donc pas comme on aurait pu l'attendre si chaque élève avait un classement unique pour chaque test.

Tableau 1.2.8 : Caractéristiques des variables de rang des élèves selon les pays

	Madagascar	Mali	Tchad
Moyenne de la différence de rang entre le début et la fin de l'année	54,8	-23,1	161,4
Plage de variation	-2056 / 1912	-1598 / 1531	-1050 / 1224
Corrélation entre le rang et le score en début d'année	-0,97	-0,96	-0,94
Corrélation entre le rang et le score en fin d'année	-0,97	-0,95	-0,96

On remarque aussi que les différences de rang entre les tests de début et de fin d'année sont comprises sur des plages de variation très importantes²⁶ et qui sont directement liées à la taille

²⁶ Il a d'ailleurs été jugé préférable d'exclure les valeurs extrêmes susceptibles de perturber les analyses (Cf. annexe 1.2.5).

de l'échantillon. Pour favoriser les comparaisons avec les modèles explicatifs des scores mais aussi entre pays pour les modèles explicatifs des classements, nous avons standardisé la variable avec une moyenne de zéro et un écart-type de 1. Enfin, la très forte corrélation entre les scores et les rangs découle de la construction même de la variable, elle nous confirme la proximité entre les variables qui nous intéressent ici.

Le tableau 1.2.9 présente les résultats de deux types de modèles. Le premier est le modèle à valeur ajoutée classique qui considère le score initial de l'élève comme une variable explicative. Le second, le modèle de rang, cherche à expliquer les évolutions du classement des élèves entre le début et la fin de l'année. Il faut toutefois attirer l'attention sur le fait que l'échelle de mesure n'est pas la même et que les coefficients ne peuvent pas s'interpréter de la même façon. On peut cependant voir, en comparant les coefficients, si les conclusions sont différentes selon le modèle. On remarque une différence notable pour les R^2 qui sont nettement plus faibles avec les modèles de rang, ce qui n'est pas une surprise compte tenu du rôle qu'occupe la variable de score initial dans les modèles à valeur ajoutée. Il faut quand même relever que dans le cas du Mali, avec un R^2 de 3 %, le pouvoir explicatif du modèle est particulièrement faible. Pour ce qui concerne les coefficients, il n'y a pas de différences majeures avec les résultats de ce nouveau modèle. A quelques exceptions près, on retrouve les mêmes variables significatives avec des ordres de grandeur assez comparables. Les changements de significativité observés concernent des variables à la limite de la significativité dans l'un ou l'autre des modèles.

Les modèles de rang permettent de contourner certaines questions méthodologiques liées à la théorie de la connaissance ainsi qu'à la théorie des tests. Cependant, il ne s'agit que d'un artifice dont l'un des intérêts est de pouvoir mesurer une différence (en places dans le classement entre deux tests) tout en évitant les problèmes d'endogénéité soulevés par la présence de la variable de score initial dans les modèles à valeur ajoutée traditionnels. Il n'est pas clairement établi que ces avantages compensent les inconvénients liés à l'abandon de la modélisation du processus d'apprentissage.

Tableau 1.2.9 :
Comparaison des résultats des modèles à valeur ajoutée et des modèles de rang

Variables	Madagascar		Mali		Tchad	
	M1	M2	M3	M4	M5	M6
Score de début d'année	0,47*** (0,06)	/	0,56*** (0,04)	/	0,54*** (0,05)	/
L'élève						
est une fille	0,04 (0,04)	0,04 (0,05)	-0,11** (0,05)	-0,05 (0,06)	0,05 (0,06)	0,13* (0,06)
Age	-0,08*** (0,02)	-0,08*** (0,02)	-0,01 (0,02)	-0,02 (0,02)	-0,00 (0,02)	-0,01 (0,02)
Appartient à une famille de niveau de vie élevé	0,16* (0,09)	0,20** (0,10)	0,01 (0,07)	0,05 (0,08)	-0,17* (0,10)	-0,28** (0,11)
a redoublé au moins une fois	-0,08 (0,06)	0,00 (0,06)	0,02 (0,07)	0,11 (0,07)	-0,03 (0,06)	0,02 (0,07)
possède les manuels de français et mathématiques	0,08 (0,07)	0,03 (0,08)	0,03 (0,07)	-0,03 (0,07)	0,11 (0,11)	0,08 (0,16)
Le maître						
Ancienneté	0,02 (0,01)	0,02 (0,01)	-0,00 (0,00)	-0,00 (0,01)	-0,01 (0,01)	-0,01 (0,01)
de niveau second cycle secondaire	0,15 (0,13)	-0,08 (0,13)	0,08 (0,13)	0,11 (0,14)	-0,00 (0,17)	-0,05 (0,23)
titulaire du BAC	0,30 (0,23)	0,06 (0,26)	/	/	0,08 (0,22)	0,05 (0,27)
formation professionnelle ≥ 1 an	0,07 (0,14)	-0,10 (0,15)	-0,14 (0,13)	-0,13 (0,13)	-0,12 (0,14)	-0,16 (0,16)
La classe						
double flux	-0,30* (0,16)	-0,28 (0,18)	-0,37*** (0,11)	-0,30** (0,13)	-0,63* (0,33)	-0,63 (0,47)
multigrade	0,57 (0,43)	0,79* (0,43)	0,07 (0,25)	-0,14 (0,19)	0,03 (0,15)	0,07 (0,17)
Nombre d'élèves	-0,00 (0,00)	-0,00 (0,01)	-0,00 (0,00)	-0,00 (0,00)	-0,01*** (0,00)	-0,01** (0,00)
L'école						
est en zone rurale	0,02 (0,12)	0,19 (0,13)	-0,14 (0,12)	-0,11 (0,12)	-0,22 (0,14)	-0,03 (0,14)
Observations ²⁷	2044	2044	1740	1740	1226	1226
R ²	0,29	0,08	0,37	0,03	0,36	0,08

Erreurs types robustes entre parenthèses

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

²⁷ Les valeurs extrêmes identifiées avec les évolutions de rang ont été exclues des analyses pour les deux types de modèles afin d'avoir une base identique dans les deux cas.

1.2.5 L'utilisation des modèles hiérarchiques ou multiniveaux

La littérature sur les modèles hiérarchiques (Cf. Bryk et Raudenbush, 1992 ; Goldstein, 1995 ; Bressoux, 2000) souligne les insuffisances des MCO pour traiter des données ayant une structure hiérarchique comme, typiquement, les données scolaires. Une de ces limites, particulièrement importante pour l'analyse des données scolaires, touche à l'une des hypothèses fondamentales sur lesquelles reposent les MCO : l'indépendance des résidus. Les MCO supposent que le résidu attaché à l'individu i n'est pas lié au résidu de tout autre individu j . Cette hypothèse est bien souvent problématique. C'est le cas pour les séries temporelles où chaque individu est observé plusieurs fois dans le temps, la valeur au temps $t+1$ dépend de la valeur au temps t , on parle d'autocorrélation : les individus dépendent d'une variable explicative particulière, le temps. L'hypothèse est également intenable dans le cas de données hiérarchiques comme les données scolaires. En effet, les élèves qui appartiennent à une même classe tendent généralement à être plus semblables entre eux que ceux de classes différentes (Goldstein, 1995). Une façon de mesurer cette ressemblance selon les classes est de calculer le coefficient de corrélation intra-classe (ρ).

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2} \quad \text{où } \sigma_{u_0}^2 \text{ est la variance au niveau classe et } \sigma_e^2 \text{ est la variance au niveau élève.}$$

Goldstein (1995) estime que l'existence d'un ρ non nul remet en cause l'utilisation des MCO et il devient préférable d'avoir recours au modèle multiniveau. Or, on observe dans les enquêtes internationales sur les acquis scolaires comme PASEC, PIRLS, SACMEQ et TIMSS des valeurs de ρ qui varient entre 0,1 et 0,7 (Cf. tableau 1.2.10).

Tableau 1.2.10 : Coefficient de corrélation intra-classe dans les données PASEC²⁸

Pays	Rho
Burkina Faso (1996)	0,4
Côte d'Ivoire (1996)	0,4
Sénégal (1996)	0,2
Mauritanie (2004)	0,5
Tchad (2004)	0,5
Cameroun (2005)	0,4
Madagascar (2005)	0,3

Source : PASEC

Il est donc bien difficile de maintenir l'hypothèse d'indépendance des résidus, d'autant que la violation de cette hypothèse entraîne des biais dans les estimations. Traditionnellement, les économistes ont l'habitude de ce type de problème avec les données temporelles et utilisent les moindres carrés généralisés (MCG) pour les surmonter (Greene, 2005). C'est aussi en partie pour faire face à ce problème que les modèles hiérarchiques ou multiniveaux se sont développés dans les sciences sociales (Bressoux, 2000).

La présentation du modèle multiniveau classique à deux niveaux qui suit reprend celle des principaux manuels dont celui de Hox (2002).

En partant des MCO :
$$Y_{ij} = \beta_0 + \beta_1 x_{ij} + e_{ij}$$

Y_{ij} est la variable expliquée associée à l'individu i du groupe j

x_{ij} est la variable explicative associée à l'individu i du groupe j

β_0 est la constante et β_1 est la pente de la droite de régression de y sur x .

e_{ij} représente une erreur aléatoire pour chaque individu i du groupe j de moyenne nulle et de variance σ_e^2

Le modèle multiniveau s'écrit :

$$Y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij}$$

On constate que les pentes et les constantes ont un indice j indiquant qu'elles peuvent varier d'un groupe à l'autre. Les coefficients des constantes et des pentes sont donc maintenant

²⁸ Il s'agit du coefficient de corrélation intra-classe calculé à partir du score agrégé de français et de mathématiques en 5^{ème} année.

rendus aléatoires, ce qui signifie que leurs valeurs sont supposées distribuées selon une fonction de probabilité (Kreft et De Leeuw, 1998). On peut écrire :

$$\beta_{0j} = \gamma_{00} + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j}$$

où γ_{00} représente la constante moyenne pour tous les groupes, γ_{10} la pente moyenne pour tous les groupes, μ_{0j} représente l'écart de chaque groupe à la constante (c'est une variable aléatoire de moyenne nulle et de variance σ_{u0}^2), μ_{1j} représente l'écart de chaque groupe à la relation moyenne (c'est une variable aléatoire de moyenne nulle et de variance σ_{u1}^2). Un paramètre supplémentaire peut être estimé : la covariance entre les constantes et les pentes σ_{u01}^2 . Les termes d'erreurs μ_{0j} et μ_{1j} sont supposés être indépendants de l'erreur au niveau individuel e_{ij} .

En intégrant dans une même équation : $Y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + (\mu_{0j} + \mu_{1j}x_{ij} + e_{ij})$

Entre parenthèses la partie aléatoire du modèle qui distingue fondamentalement le modèle multiniveau du modèle des MCO.

L'étape suivante dans la modélisation multiniveau est d'intégrer des variables de niveau supérieur ou de groupe (variables z) pour expliquer à la fois les constantes et les pentes. Si l'on intègre dans le modèle par exemple une variable de groupe, on obtient :

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_j + \mu_{1j}$$

En intégrant dans une même équation :

$$Y_{ij} = \gamma_{00} + \gamma_{10} + \gamma_{01}z_j + \gamma_{11}z_j x_{ij} + (\mu_{0j} + \mu_{1j}x_{ij} + e_{ij})$$

Le terme $z_j x_{ij}$ correspond à une interaction.

On remarque également que le terme aléatoire μ_{1j} est lié à x_{ij} , cela signifie que le terme d'erreur va varier en fonction de x_{ij} , ce qu'on appelle dans un modèle classique de l'hétéroscédasticité. Il ne s'agit pas d'un phénomène qui parasite l'estimation mais au contraire le modèle multiniveau permet de modéliser l'hétéroscédasticité (Bressoux, 2000).

Les hypothèses du modèle sont les suivantes :

- Tous les termes d'erreurs suivent des lois normales.
- Les aléas de niveau classe sont indépendants de l'aléa de niveau élève, ce qui implique que : $cov(\mu_{0j}, \mu_{1j}) = cov(\mu_{1j}, \mu_{0j}) = 0$
- Les aléas de niveau classe sont indépendants d'une classe à l'autre

Ce modèle peut être étendu à un plus grand nombre de niveaux.

Toutefois, comme précédemment, la sensibilité des résultats selon le type de modèles (MCO *versus* modèle multiniveau –MLN-) a été testée. En ce qui concerne les MCO, il s'agit du modèle avec l'estimation robuste des intervalles de confiance²⁹ qui présente toutefois certaines limites dans l'analyse de données hiérarchiques (Cf. annexe 1.2.2). Les résultats des deux types de modèles ont été mis en perspective toujours avec les données des trois pays du PASEC³⁰. On constate (Cf. tableau 1.2.11) que la modélisation multiniveau amène effectivement des changements, quoique peu nombreux, dans les résultats des estimations. Toutefois, les variations des coefficients restent globalement modérées, même si dans de rares cas elles ne sont pas complètement négligeables (exemple du niveau de vie à Madagascar). On note surtout des évolutions des significativités de différentes variables selon les pays. Le cas du niveau de vie est tout à fait intéressant, car on voit pour Madagascar et pour le Tchad que le modèle multiniveau diminue les coefficients de cette variable pour l'amener à des valeurs proches de 0 tout en leur ôtant leur significativité. Compte tenu des problèmes mentionnés quant à cette variable, ce résultat est donc plutôt cohérent. Il est aussi intéressant de souligner que les changements n'interviennent pas uniquement sur les variables de niveau école et classe mais aussi sur les variables élèves, ce qui montre que l'amélioration de l'estimation concerne potentiellement toutes les variables. On observe cependant très peu de changements pour les variables relatives aux enseignants, aucune n'apparaît significative dans les deux types de modèles. Au niveau de la classe, on voit que la significativité du double flux à Madagascar et au Tchad, modeste avec les MCO, disparaît avec le modèle multiniveau. Inversement, le coefficient du multigrade devient plus important et significatif à Madagascar. Ainsi, c'est avant tout au niveau de la confiance accordée aux estimations des différents coefficients que se trouve l'enjeu entre les deux types de modèles. Or les modèles multiniveaux fournissent les estimations les plus fiables avec des données hiérarchiques comme les données scolaires utilisées ici.

²⁹ On peut déjà considérer que ce modèle est un premier pas vers l'analyse multiniveau puisqu'il corrige l'estimation des intervalles de confiance des coefficients des MCO classiques.

³⁰ Cf. annexe 1.2.7 pour les sorties STATA des différents modèles.

**Tableau 1.2.11 : Comparaison des MCO avec le modèle multiniveau
en 5^{ème} année d'école primaire dans trois pays du PASEC³¹**

Variables	Madagascar		Mali		Tchad	
	MCO	MLN	MCO	MLN	MCO	MLN
L'élève						
est une fille	0,04 (0,04)	0,02 (0,03)	-0,09** (0,05)	-0,09*** (0,03)	0,04 (0,06)	-0,06* (0,03)
Age	-0,08*** (0,02)	-0,04*** (0,01)	-0,00 (0,02)	-0,02** (0,01)	0,00 (0,02)	-0,01 (0,01)
Appartient à une famille de niveau de vie élevé	0,17* (0,09)	0,00 (0,04)	-0,01 (0,07)	0,00 (0,04)	-0,18* (0,10)	-0,04 (0,06)
a redoublé au moins une fois	-0,08 (0,06)	-0,10*** (0,03)	0,01 (0,07)	-0,02 (0,03)	-0,01 (0,06)	-0,01 (0,04)
possède les manuels de français et mathématiques	0,10 (0,07)	0,05 (0,03)	0,04 (0,07)	0,03 (0,03)	0,13 (0,11)	0,01 (0,06)
Le maître						
Années d'ancienneté	0,02 (0,01)	0,02 (0,01)	0,00 (0,01)	0,00 (0,01)	-0,01 (0,01)	-0,01 (0,01)
niveau second cycle secondaire	0,12 (0,13)	0,02 (0,14)	0,07 (0,14)	0,07 (0,15)	0,01 (0,17)	0,10 (0,19)
titulaire du BAC	0,30 (0,24)	0,23 (0,31)	/	/	0,13 (0,22)	0,11 (0,22)
formation professionnelle initiale ≥ 1 an	0,04 (0,14)	0,05 (0,17)	-0,14 (0,13)	-0,13 (0,15)	-0,18 (0,16)	-0,12 (0,16)
La classe						
Double flux	-0,31* (0,16)	-0,36 (0,25)	-0,40*** (0,12)	-0,40*** (0,13)	-0,63* (0,34)	-0,60 (0,50)
Multigrade	0,55 (0,42)	0,68** (0,32)	0,04 (0,25)	0,07 (0,22)	0,07 (0,16)	0,13 (0,17)
Nombre d'élèves	-0,00 (0,00)	-0,00 (0,00)	-0,00 (0,00)	-0,00 (0,00)	-0,01*** (0,00)	-0,01*** (0,00)
L'école						
est en zone rurale	0,02 (0,12)	0,04 (0,13)	-0,11 (0,13)	-0,11 (0,13)	-0,24* (0,14)	-0,22 (0,17)
Score de début d'année (français et mathématiques)	0,43*** (0,06)	0,47*** (0,02)	0,52*** (0,05)	0,43*** (0,02)	0,52*** (0,05)	0,55*** (0,03)
Observations	2063	2063	1761	1761	1235	1235
R ²	0,26	/	0,34	/	0,34	/
Nombre de classes		118		134		108

Erreurs types robustes entre parenthèses

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

³¹ Les scores de début et de fin d'année utilisés dans ces modèles sont standardisés avec une moyenne de 0 et un écart-type de 1.

1.2.6 L'estimation des modèles explicatifs des acquis scolaires : un monde de biais

Une critique majeure adressée aux modèles à valeur ajoutée est les biais potentiels qu'ils comportent et qui remettent en cause la confiance qu'on peut leur accorder. On a cependant pu voir que les modèles ne prenant pas en compte le score initial des élèves étaient potentiellement beaucoup plus biaisés. De plus, cette critique vaut pour la plupart des modèles économétriques et, au-delà, c'est l'importance de ces biais qui est déterminante ainsi que la capacité du modélisateur à les prendre en compte. Un biais peut altérer la précision d'un coefficient de façon minime comme il peut inverser un résultat et donc induire des conclusions complètement erronées. Le tout est de savoir où l'on se situe dans un modèle étant acquis que les modèles à valeur ajoutée comportent par nature des biais.

Les analyses visant à tester la sensibilité des modèles à valeur ajoutée menées ici sur trois bases de données différentes mais comparables du PASEC, permettent de relativiser certaines craintes. La faible variabilité des coefficients, associée à l'introduction d'une variable endogène et très influente comme le score de début d'année, montre que ces biais sont moins nombreux et surtout moins marqués qu'on aurait pu le redouter. En outre, pour limiter les problèmes d'endogénéité posés par la variable de score de début d'année, un modèle basé sur les évolutions de classement des élèves entre le début et la fin de l'année a été proposé. Néanmoins, des biais peuvent persister même s'ils sont *a priori* modérés et il est nécessaire de prendre des précautions dans l'interprétation des coefficients, le modèle n'ayant plus la même signification.

De plus, dans le cadre de données hiérarchiques comme les données scolaires, il est préférable d'avoir recours à des modèles multiniveaux capables de prendre en compte la structure de ces données. En effet, l'hypothèse d'indépendance des résidus des MCO est nécessairement violée avec ce type de données ce qui biaise les estimations.

Enfin, la précision d'un coefficient doit être considérée avec circonspection. De fait, la précision de ce type d'estimations est assez médiocre ; elles permettent de dégager les tendances fortes mais pas des effets modérés. C'est un peu comme si on observait la lune avec des jumelles, la précision serait toute relative mais on pourrait quand même voir les plus gros cratères. Pour accéder à un niveau de précision plus élevé, il devient nécessaire de faire appel à d'autres méthodes d'évaluation comme l'approche expérimentale (Duflo, 2006). Toutefois,

cette dernière catégorie d'évaluation ne permet pas de répondre aux mêmes questions (Glewwe et Kremer, 2006 ; Todd et Wolpin, 2003). Il y a donc plutôt une complémentarité à explorer dans ce domaine.

1.2.7 Quelques perspectives pour l'analyse des déterminants des acquisitions scolaires

Un certain nombre d'enseignements peuvent être tirés de ce chapitre quant à l'analyse des déterminants des acquisitions scolaires, même s'ils concernent essentiellement l'analyse empirique.

La mesure du produit de l'éducation est probablement l'une des premières interrogations qui se pose dans l'analyse des systèmes éducatifs. Il existe désormais un relatif consensus entre chercheurs pour l'utilisation des tests d'acquisitions scolaires comme mesure de la qualité du produit éducatif. Cela n'exclut pas l'utilisation par certaines études d'autres mesures axées sur d'autres dimensions comme la socialisation, mais il faut reconnaître que la mesure des acquis cognitifs s'est peu à peu imposée comme la principale mesure du produit éducatif. Sans doute est-ce d'un point de vue social la mesure la plus légitime. Il est vrai, comme le souligne Hanushek (1979), qu'à l'école primaire les acquis cognitifs demeurent l'objectif prioritaire.

Un autre point délicat est le cadre théorique qui sous-tend les analyses empiriques, en effet la référence à la fonction de production reste omniprésente dans la littérature. Toutefois, l'examen mené montre qu'il s'agit plutôt d'une analogie que d'un véritable cadre théorique. On peut cependant observer que l'interprétation des résultats des études empiriques peut être influencée par le concept de fonction de production, ce qui peut en donner une lecture biaisée favorisant la mise en avant de l'effet d'inputs scolaires, sans en expliciter totalement les modes d'action. La recherche d'un cadre théorique pour l'analyse de la production scolaire reste un chantier relativement récent. Les rares travaux réalisés ont en commun de mettre l'accent sur la nécessité d'une théorie comportementale qui fait défaut jusqu'ici.

L'absence d'un cadre théorique bien établi souligne la dimension essentiellement empirique des travaux réalisés dans ce champ de recherche. De fait, l'écrasante majorité des travaux présents dans la littérature est de nature empirique. D'ailleurs, la richesse de ces travaux et l'insuffisante convergence de leurs résultats, au-delà des interrogations sur le cadre théorique

de référence, ont surtout amené l'attention sur les questions méthodologiques. Si les modèles à valeur ajoutée, notamment les modèles de gains annuels, semblent à privilégier, ils posent des problèmes de biais et d'endogénéité que certains jugent rédhibitoires. Cependant, l'analyse de la sensibilité des modèles à valeur ajoutée, dans le cadre des données PASEC, ne condamne pas aussi formellement l'utilisation de ces modèles puisque les biais mis en évidence ne sont pas aussi importants que supposés. Ils impliquent surtout beaucoup plus de prudence dans les interprétations, notamment sur la confiance qu'on accorde à un coefficient donné. Il est important de le considérer comme un ordre de grandeur, il est d'ailleurs compris dans un intervalle de confiance, plutôt que comme une mesure précise de l'effet d'une variable. On accordera aussi d'autant plus de confiance à un résultat que celui sera stable dans différentes spécifications du modèle (Leamer, 1983). Enfin, la structure hiérarchique des données scolaires remet en cause l'utilisation des MCO classiques et invite, afin d'éviter des biais supplémentaires, à utiliser des modèles capables de prendre en compte cette spécificité des données scolaires.

Chapitre 2 : La problématique enseignante à l'école primaire en Afrique francophone : entre enjeux de scolarisation et qualité de l'éducation

Au début des années 90, la situation de la scolarisation primaire dans beaucoup de pays d'Afrique subsaharienne francophone était particulièrement préoccupante. Des efforts importants devaient donc être réalisés pour accroître significativement la scolarisation primaire, ce qui impliquait le recrutement massif d'enseignants. Toutefois, compte tenu des contraintes budgétaires, les coûts salariaux des enseignants titulaires ne permettaient pas un tel recrutement dans la plupart des pays. Afin d'accroître sensiblement le recrutement d'enseignants, il est devenu nécessaire d'envisager des rémunérations plus faibles et, par conséquent, d'avoir recours à d'autres profils d'enseignants. De nouvelles catégories d'enseignants sont ainsi apparues, soit du fait des communautés (maîtres communautaires ou maîtres de parents), soit, plus tardivement, du fait de la mise en place de politiques éducatives par les pouvoirs publics (volontaires, contractuels...). Les politiques de recrutement des enseignants contractuels ont connu une nette accélération à la fin des années 90. Dans certains pays, ces nouveaux enseignants tendent désormais à être plus nombreux que les enseignants fonctionnaires.

C'est tout d'abord sur le plan quantitatif que se placent les enjeux soulevés par l'objectif de scolarisation primaire universelle (2.1). Les effectifs d'élèves à scolariser pour 2015 et l'incidence que cette situation aura en termes de recrutement des enseignants constituent notre toile de fond. Dans ce contexte, la place grandissante qu'occupent les nouveaux enseignants apparaît comme une caractéristique importante des systèmes éducatifs africains mais encore assez mal renseignée. En outre, les questions de rémunération qui soulignent la spécificité des pays d'Afrique francophone sont au cœur des débats actuels sur les enseignants. Les différents éléments quantitatifs disponibles ainsi que les coûts permettent de réaliser des simulations budgétaires qui jettent une lumière crue sur les contraintes importantes pesant sur les pays et montrent ainsi que le choix de la politique des contractuels était, en quelque sorte, imposé par le contexte économique et budgétaire. Ce choix a permis dans certains pays

comme le Mali ou le Niger une croissance sans précédent des effectifs d'élèves de l'enseignement primaire.

Malgré ces succès quantitatifs, la présence massive de ces nouveaux enseignants suscite au sein de la communauté éducative une interrogation quant à leur incidence sur la qualité de l'éducation. Cette interrogation s'explique par le fait que ces maîtres non-fonctionnaires possèdent des caractéristiques différentes de leurs collègues. En effet, outre leur statut, tous n'ont pas bénéficié d'une formation professionnelle et, quand ils ont suivi une formation, elle a généralement été de courte durée. De plus, ils ont généralement un niveau de rémunération nettement inférieur à celui des enseignants fonctionnaires. Ces différents éléments ne plaident pas *a priori* en faveur de ces enseignants aux yeux des acteurs des systèmes éducatifs. Toutefois, la littérature sur cette question est nettement moins tranchée que les opinions des acteurs (2.2). Des travaux portant sur des pays développés et des pays en développement, en particulier sur les pays africains, ont été mobilisés. Malgré les débats scientifiques autour de la relation entre les caractéristiques des enseignants et les acquisitions des élèves qui ont été relevés en partie dans le chapitre précédent, un certain nombre de convergences peuvent être mises en avant, elles ne vont pas nécessairement dans le sens attendu.

A travers cette interrogation sur la qualité, on voit bien que la question qui se fait jour est en fait celle des profils enseignants qui permettront de réaliser la scolarisation primaire universelle. Le problème est très souvent posé en termes d'arbitrage entre la scolarisation du plus grand nombre d'enfants possible et la qualité de l'enseignement dispensé. Le recrutement d'enseignants non-fonctionnaires, qui est autant le fait des gouvernements que celui des communautés, serait un choix délibéré en faveur de la scolarisation du plus grand nombre au détriment de la qualité de l'éducation. Il s'agit d'une question complexe mais aussi politiquement très sensible du fait qu'elle touche directement au statut de la principale catégorie d'employés de l'Etat et aux attentes des familles en matière d'accès à l'école et de qualité de l'enseignement.

2.1 Extension de la scolarisation et évolution du corps enseignant³²

Le développement de la scolarisation et la composition du corps enseignant sont deux aspects intimement liés dans la plupart des pays d'Afrique francophone. Nous montrerons que les progrès à réaliser en matière de scolarisation impliquent nécessairement une forte croissance des effectifs d'enseignants. Nous verrons qu'au début des années 2000, les enseignants non-fonctionnaires sont déjà très présents dans les systèmes éducatifs, notamment par le biais d'enseignants rémunérés par les familles. A travers l'exemple de quelques pays, nous constaterons que cette évolution de la composition du corps enseignant est liée aux questions de rémunération. Nous essaierons ensuite de quantifier l'impact des enseignants non fonctionnaires sur la scolarisation à l'école primaire.

2.1.1 Les attentes en matière de scolarisation

L'objectif de scolarisation primaire universelle doit tenir compte d'une contrainte démographique importante : l'accroissement du nombre d'enfants à scolariser. Le tableau 2.1.1 présente l'accroissement du nombre d'élèves observé sur la période 1987-2001 pour certains groupes de pays africains et pour l'ensemble du continent, ainsi que les effectifs estimés pour 2015.

Tableau 2.1.1 : Accroissement du nombre d'élèves entre 1987 et 2001 et entre 2001 et 2015

	Elèves scolarisés (en milliers)			Croissance des effectifs passée et requise (en %)	
	1987	2001	2015	1987-2001	2001-2015
CEDEAO	19 646	30 141	55 043	+53,4 %	+82,6 %
CEMAC	3 312	4 976	7 989	+50,2 %	+60,6 %
Afrique	72 711	106 140	176 208	+45,9 %	+66 %

Source : Amelewonou, Brossard et Gacougnolle (2004).

³² Une grande partie de cette section est tirée de Bernard, Tiyab et Vianou (2004).

Selon les projections démographiques de la Division de la Population des Nations-Unies, l'Afrique comptera environ 176 millions d'enfants en âge d'aller à l'école en 2015. Atteindre la scolarisation primaire universelle implique donc que les effectifs d'enfants scolarisés devront croître de plus de 70 millions entre 2001 et 2015. A titre de comparaison, les effectifs ont augmenté d'un peu plus de 33 millions entre 1987 et 2001, soit une croissance d'environ 46 %. Or, pour atteindre la scolarisation primaire universelle en 2015, c'est une croissance de 66 % qui sera nécessaire. On constate que les pays de la CEDEAO devront pour leur part connaître une croissance des effectifs supérieure à 80 %. Bien sûr, derrière ces moyennes se cachent de grandes disparités selon les pays ; toutefois, le défi de la scolarisation primaire universelle apparaît dans toute son ampleur.

Il est bon ici de rappeler que la scolarisation primaire universelle n'est pas seulement une question d'offre scolaire et qu'il faut dans une analyse globale accorder une attention particulière aux questions de demande. Mingat (2007) montre que la disponibilité d'écoles et d'enseignants est une condition nécessaire mais non suffisante pour que tous les enfants aillent à l'école. En se focalisant sur les aspects quantitatifs du recrutement des enseignants dans ce qui suit, on fait comme si ces questions de demande étaient réglées pour les besoins des estimations ; il faut cependant garder à l'esprit qu'il n'en est rien et qu'il s'agit d'une dimension importante pour l'atteinte de la scolarisation primaire universelle.

2.1.2 Les implications de l'accroissement des effectifs d'élèves sur le recrutement d'enseignants

Le corollaire de l'accroissement du nombre d'élèves et de l'amélioration souhaitée des taux d'encadrement³³ est une augmentation du nombre d'enseignants aussi rapide (CEDEAO), voire plus rapide (CEMAC), que celle des effectifs scolarisés. Les estimations de Amelewonou, Brossard et Gacougnolle (2004) permettent de mettre en évidence cette croissance nécessaire des effectifs enseignants dans le secteur public pour atteindre la scolarisation primaire universelle en 2015. Leur modèle est basé sur 52 pays avec comme année de base 2003. Parmi les paramètres retenus pour l'estimation, ils considèrent un ratio élèves-maître de 40, un taux de redoublement de 10 % pour les pays où ce taux est supérieur

³³ Les taux d'encadrement ou ratios élèves/maître qui prévalent aujourd'hui sont jugés excessifs dans de nombreux pays et préjudiciables à la qualité ainsi qu'aux conditions de travail des enseignants.

et le taux réel pour les pays où il est inférieur ; la part des écoles sur financement privé est estimée à 10 % sauf pour les pays où cette part est inférieure : c'est alors la part réelle qui est prise en considération³⁴. Le tableau 2.1.2 montre l'accroissement du nombre d'enseignants publics entre 1987 et 2001, les effectifs estimés pour 2015 de même que l'accroissement requis. D'après ces simulations, pour scolariser près de 180 millions d'enfants africains en 2015, il faudra environ 3,8 millions d'enseignants en poste dans l'enseignement public. On constate que l'accroissement sera plus marqué dans la période 2001-2015 pour les deux groupes de pays (CEDEAO et CEMAC) mais c'est pour l'Afrique centrale que l'accroissement devrait être en moyenne nettement plus rapide (quatre fois plus). Malgré tout, ce sont deux pays de la CEDEAO, le Mali et le Niger, qui présentent les taux d'accroissement annuel moyen les plus élevés pour la période 2001-2015 avec respectivement 10,3 % et 12 %. Cela montre que derrière les moyennes des sous-ensembles se cachent des différences importantes selon les pays. Cela dit, les chiffres globaux donnent une idée générale de l'énorme défi auquel sont confrontés ces pays. D'ailleurs, un rapport de l'Institut statistique de l'UNESCO (2006) qui procède à une nouvelle estimation des besoins en enseignants en 2015 pour l'ensemble des pays du monde souligne que c'est en Afrique subsaharienne que le challenge est le plus grand. Selon ce rapport, 76 pays à travers le monde doivent augmenter leur effectif d'enseignants et environ 60 % de ces besoins en enseignants, évalués à 1,6 million (secteurs public et privé confondus), se situent en Afrique subsaharienne.

³⁴ Les auteurs s'inspirent ici des valeurs fournies par le cadre indicatif de Fast track (Cf. annexe 2.1.1).

Tableau 2.1.2 : Evolution passée et future du nombre d'enseignants du secteur public pour atteindre la scolarisation primaire universelle

	Enseignants publics			Accroissement annuel moyen passé et requis (%)	
	1987	2001	2015	1987-2001	2001-2015
CEDEAO	490 133	694 426	1 262 660	+ 2,5	+ 4,4
Bénin	14 067	17 266	36 036	+ 1,5	+ 5,4
Burkina Faso	5 786	19 007	64 004	+ 8,9	+ 9,1
Cap-Vert	1 892	3 124	1 906	+ 3,6	- 3,5
Côte d'Ivoire	31 717	39 026	71 280	+ 1,5	+ 4,4
Gambie	2 604	4 059	6 435	+ 3,2	+ 3,3
Ghana	63 161	61 229	92 022	- 0,2	+ 3,0
Guinée	7 203	14 267	37 521	+ 5,0	+ 7,2
Guinée-Bissau	3 065	3 343	5 816	+ 0,6	+ 4,0
Libéria	3 680	7 450	18 266	+ 5,2	+ 6,6
Mali	7 855	17 788	69 968	+ 6,0	+ 10,3
Niger	8 019	14 998	73 384	+ 4,6	+ 12,0
Nigéria	309 032	442 322	679 909	+ 2,6	+ 3,1
Sénégal	10 793	20 081	54 252	+ 4,5	+ 7,4
Sierra Leone	13 640	14 577	25 379	+ 0,5	+ 4,0
Togo	7 619	15 889	26 483	+ 5,4	+ 3,7
CEMAC	46 757	60 959	176 171	+ 1,9	+ 7,9
Cameroun	25 185	30 833	83 804	+ 1,5	+ 7,4
République centrafricaine	4 563	6 223	16 979	+ 2,2	+ 7,4
Tchad	6 215	13 819	49 698	+ 5,9	+ 9,6
Congo	7 429	5 185	17 746	- 2,5	+ 9,2
Guinée équatoriale	664	1 440	1 980	+ 5,7	+ 2,3
Gabon	2 701	3 459	5 965	+ 1,8	+ 4,0
Afrique	1 961 282	2 698 878	3 809 563	+ 2,3	+ 2,5

Source : Amelewonou, Brossard et Gacougnolle (2004).

De telles évolutions ont inévitablement des implications notables dans le domaine financier. Du fait des contraintes budgétaires et macroéconomiques ainsi que du niveau de rémunération des enseignants fonctionnaires, beaucoup de pays africains n'ont pas été (et ne sont toujours pas) en mesure de recruter des enseignants fonctionnaires suivant un rythme d'expansion

nécessaire à l'atteinte de l'objectif de scolarisation primaire universelle. Ceci a conduit certains gouvernements, parfois sous la pression de la demande et des initiatives locales, parfois sur le conseil des partenaires techniques et financiers, à modifier ou à diversifier le recrutement des maîtres en introduisant de nouvelles catégories d'enseignants qui n'étaient plus des fonctionnaires. Parallèlement à cette réponse de l'Etat, et le précédent souvent, on a également vu émerger et se développer des réponses des populations par la mise en place d'écoles communautaires recrutant localement leurs enseignants (appelés enseignants communautaires ou maîtres de parents), mais aussi par le recrutement direct d'enseignants pour des écoles publiques insuffisamment dotées.

Ainsi, les dernières années ont vu se développer et s'accélérer le recrutement d'enseignants aux profils nouveaux, non-fonctionnaires, au service de l'Etat ou des communautés, créant par là une situation de dualité des statuts et des niveaux de rémunération au sein des systèmes éducatifs.

2.1.3 La composition du corps enseignant : tendance actuelle et évolution

Aujourd'hui, la politique de recrutement de ces nouveaux enseignants (contractuels, volontaires, maîtres de parents) n'est pas au même stade de développement dans tous les pays. Le tableau 2.1.3 présente la distribution des enseignants selon leur statut dans 12 pays francophones d'Afrique subsaharienne au début des années 2000.

Tableau 2.1.3 : Distribution des enseignants du primaire selon le statut

PAYS	Fonctionnaires	Contractuels de l'Etat	Parents ³⁵	Total
Bénin (2002)	55 %	16 %	29 %	100 %
Burkina Faso (2002)	64 %	24 %	12 %	100 %
Cameroun (2002)	35 %	20 %	45 %	100 %
Congo (2003)	42 %	4 %	54 %	100 %
Côte d'Ivoire (2001)	87 %	0 %	13 %	100 %
Guinée (2003)	31 %	39 %	30 %	100 %
Madagascar (2003)*	46 %	0 %	54 %	100 %
Mali (2000)	71 %	8 %	21 %	100 %
Niger (2003)	46 %	50 %	4 %	100 %
Sénégal (2003)	44 %	41 %	15 %	100 %
Tchad (2002)*	33 %	0 %	67 %	100 %
Togo (2001)	35 %	30 %	35 %	100 %
Moyenne	49 %	19 %	32 %	100 %

* Les maîtres communautaires reçoivent une subvention salariale de la part de l'Etat

Source : A. Mingat (2004)

Il ressort qu'au début des années 2000, les enseignants fonctionnaires ne représentent que 49 % du corps enseignant au niveau de l'enseignement primaire. Les enseignants non-fonctionnaires constituent l'autre moitié du personnel enseignant, avec 19 % de contractuels de l'Etat et plus de 30 % de maîtres payés par les parents. Ceci amène un premier constat : **les nouveaux enseignants sont massivement présents dans ces systèmes éducatifs**. Par ailleurs, on remarque que les maîtres rémunérés par les parents sont plus nombreux que les contractuels de l'Etat. En effet, les politiques de recrutement massif des contractuels par les pouvoirs publics sont relativement récentes (fin des années 90), alors que les communautés ont été plus proactives et ont depuis plus longtemps eu recours à des enseignants recrutés localement dans les écoles communautaires comme dans les écoles publiques sans oublier le développement des écoles privées. En effet, environ 13 %³⁶ des élèves africains de l'enseignement primaire sont scolarisés dans des établissements privés (Pôle de Dakar, 2005).

³⁵ Comprend les enseignants du secteur privé.

³⁶ Cf. annexe 2.1.2 pour les chiffres des pays du tableau précédent.

La tendance globale masque cependant des disparités énormes entre les pays. Au Cameroun, au Congo et au Tchad, les maîtres payés par les parents occupent une place prépondérante qui souligne l'incapacité des systèmes éducatifs, au cours de la dernière décennie, à recruter suffisamment d'enseignants dans le secteur public. Dans l'ensemble, il convient malgré tout de prendre en compte les dynamiques actuelles qui voient l'accroissement des recrutements de contractuels de l'Etat et des collectivités locales qui devraient faire évoluer sensiblement les chiffres observés pour 2000 et 2001. Au Mali, par exemple, en 2004, les enseignants fonctionnaires ne représentent plus que 34,2 % des enseignants au premier cycle de l'enseignement fondamental (Banque mondiale, 2007). En effet, les autorités maliennes (Etat et collectivités locales) ont procédé à un recrutement accéléré d'enseignants contractuels au début des années 2000 (Cf. tableau 2.1.4). Elles ont recruté plus de 3000 enseignants contractuels en 2001 contre seulement un peu plus de 400 fonctionnaires. En 2002, ce sont 5800 enseignants contractuels qui ont été recrutés contre un peu plus de 200 fonctionnaires. Il est d'ailleurs intéressant de noter qu'à partir de 2001, ce sont les contractuels des collectivités qui constituent l'essentiel des recrutements.

**Tableau 2.1.4 :
Recrutement du personnel fonctionnaire et contractuel de 1998 à 2002 au Mali**

	1998	1999	2000	2001	2002	TOTAL
Fonctionnaires	697	664	-	449	239	2 049
Contractuels de l'Etat	1 270	771	1 804	285	259	4 389
Contractuels des collectivités				2 874	5 546	8 420
TOTAL	1 967	1 435	1 804	3 608	6 044	14 858

Source : PASEC (2004b)

L'exemple du Mali illustre bien l'évolution extrêmement rapide du corps enseignant dans certains pays avec l'arrivée massive d'enseignants contractuels de l'Etat et des collectivités locales, catégorie très minoritaire avant 2000. Malheureusement, il n'y a pas pour l'heure de chiffres plus complets et plus récents que ceux présentés dans le tableau 2.1.3 et qui permettraient de mieux rendre compte des dynamiques à l'œuvre.

En outre, ce changement, conforme à la réhabilitation du rôle de l'Etat dans la fourniture des services sociaux de base, pourrait avoir une incidence positive en matière d'équité. A l'heure actuelle, tandis que la communauté internationale prône une école primaire gratuite pour tous, des familles – généralement les plus défavorisées – rémunèrent elles-mêmes les enseignants pour pallier l'insuffisance de maîtres titulaires. Dans une logique d'équité et d'accès à une école primaire gratuite, conformément aux grandes résolutions prises par la communauté internationale, le recrutement des contractuels de l'Etat doit se concevoir, à terme, dans une perspective de substitution et non d'additivité aux maîtres de parents³⁷.

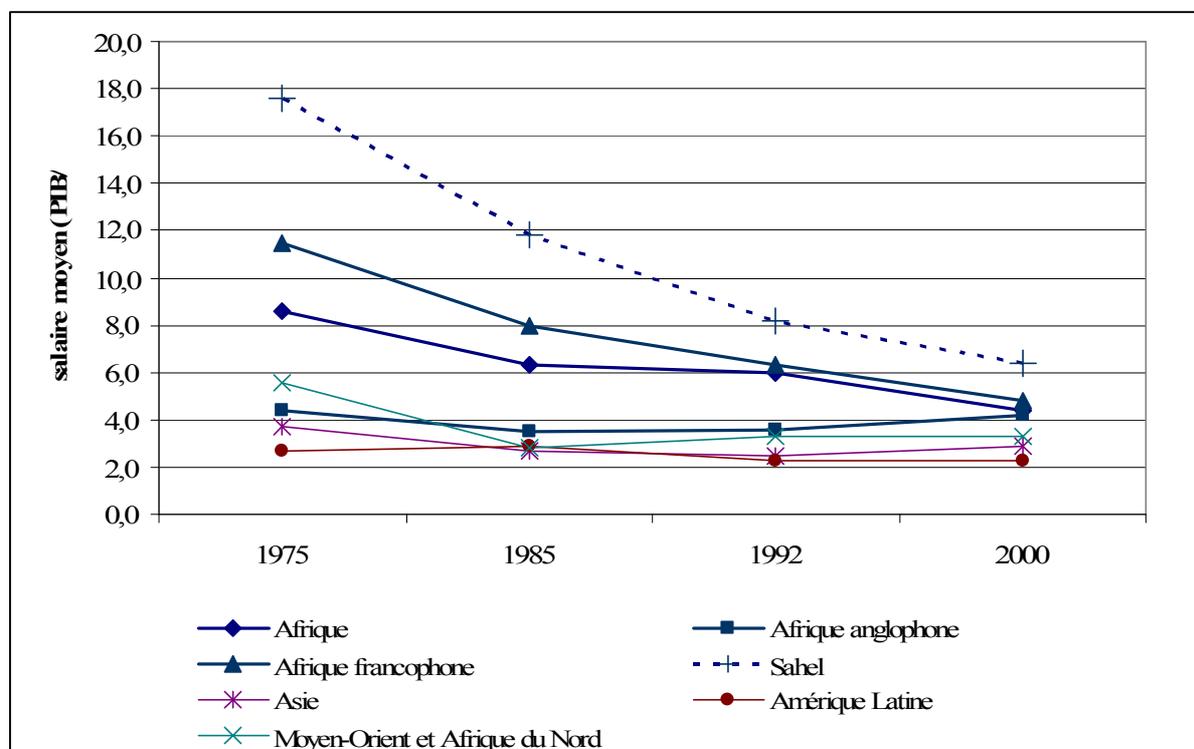
2.1.4 La rémunération des enseignants

Pour mieux comprendre ces évolutions récentes des recrutements et des profils d'enseignants, il est instructif de se pencher sur l'évolution des niveaux de rémunération des enseignants. Un premier point important à considérer est le niveau relatif des salaires en Afrique francophone par rapport aux autres régions du monde. Le graphique 2.1.1 montre que les pays africains, et plus précisément les pays d'Afrique francophone présentaient des salaires moyens relatifs nettement plus élevés que ceux observés dans les autres régions du monde en 1975³⁸ y compris l'Afrique anglophone. La situation des pays sahéliens est extrême avec un salaire moyen proche de 18 fois le PIB par tête contre moins de trois fois en Amérique latine.

³⁷ Cette tendance s'observe déjà dans plusieurs pays où le recrutement des maîtres de parents comme contractuels de l'Etat ou des collectivités se met progressivement en place.

³⁸ Il convient de souligner que seuls les pays avec un PIB par tête inférieur à 2000\$ ont été pris en compte dans les calculs.

**Graphique 2.1.1 : Evolution du salaire moyen des enseignants de l'école primaire
par région et sous-région du monde (PIB/tête)
(pays ayant un PIB par tête inférieur à 2000 \$US en 1993)**



Source : Mingat (2001) repris dans Bruns, Mingat et Rakotomalala (2003)

La situation des pays d'Afrique francophone, principalement des anciennes colonies françaises, au début des années 70 est en large partie imputable à l'héritage colonial. En effet, les salaires des enseignants étaient alors indexés sur ceux de la métropole³⁹, ce qui n'était pas le cas dans les pays sous administration britannique où l'éducation a été confiée à des missionnaires avec le soutien de subventions publiques. Pour Cogneau (2003), le poids de la politique d'assimilation française contribue à expliquer après les indépendances la tendance de long terme à maintenir des salaires des enseignants plus en rapport avec les salaires pratiqués en France qu'avec les capacités financières des pays. Il faut aussi signaler que la

³⁹ Cette situation découle directement de la loi n°50-772 du 30 juin 1950 dite loi Lamine Guèye « fixant les conditions d'attribution des soldes et indemnités des fonctionnaires civils et militaires relevant du ministère de la France d'outre-mer, les conditions de recrutement, de mise en congé ou à la retraite de ces mêmes fonctionnaires ».

parité fixe du franc cfa avec la devise française a renforcé cette tendance dans les pays de la zone cfa.

Néanmoins, entre 1975 et 2000, on observe une convergence des différentes régions du monde avec une baisse du salaire relatif moyen qui passe de 6,6 à 3,7 fois le PIB par tête. En 2000, les écarts se sont largement réduits, mais les salaires relatifs demeurent plus élevés en Afrique francophone et notamment dans les pays du Sahel. Mingat et Suchaut (2000) montrent que l'un des principaux facteurs explicatifs du retard de scolarisation des pays d'Afrique francophone par rapport aux pays d'Afrique anglophone réside dans des coûts unitaires nettement plus élevés, ceux-ci dépendant essentiellement des salaires enseignants. L'objectif de scolarisation primaire universelle, devenu le leitmotiv de toutes les grandes réunions internationales, a contraint les pays à considérer cette dimension comme essentielle puisque le premier obstacle auquel ils sont confrontés est le recrutement en nombre suffisant d'enseignants. Une estimation de Mingat (2004) permet de mieux apprécier la situation inextricable à laquelle sont confrontés nombre de pays africains. Le taux de dépendance correspond ici à la proportion des dépenses courantes d'éducation primaire qui devrait être prise en charge par un financement extérieur faute de financement national suffisant. Ce taux est estimé en fonction de différents niveaux de rémunération des enseignants et sur la base des paramètres du cadre indicatif de l'initiative de mise en œuvre accélérée de l'Education pour tous ou Initiative Fast track⁴⁰ (tableau 2.1.5). Les résultats sont assez inquiétants puisque dans l'hypothèse basse, avec une rémunération de 3,5 fois le PIB par tête, le taux de dépendance moyen est supérieur à 30 %, et dans l'hypothèse haute, c'est-à-dire en maintenant le niveau des salaires actuels des enseignants fonctionnaires, on dépasse les 50 %. Il existe bien sûr des situations assez variables selon les pays. Ces estimations et les ordres de grandeur qu'elles fournissent amènent plusieurs remarques. Tout d'abord, on constate sans surprise l'impact considérable du niveau de rémunération des enseignants sur les budgets éducatifs. Ensuite, on voit que même avec un niveau de rémunération des enseignants correspondant au cadre indicatif de Fast Track, il faudra avoir recours de façon importante au financement extérieur pour les dépenses courantes d'éducation primaire. Cela laisse relativement peu de marges de manoeuvre aux gouvernements sur ces questions de rémunérations des enseignants sauf à remettre en cause l'objectif de scolarisation primaire universelle.

⁴⁰ L'auteur précise qu'il considère 20 % des recettes publiques nationales pour le secteur de l'éducation et 50 % de cette somme pour le financement de l'éducation primaire.

Tableau 2.1.5 : Taux de dépendance en 2015 selon trois hypothèses pour le salaire des enseignants recrutés pour assurer l’achèvement universel du primaire

Pays	Niveau de salaire des enseignants recrutés entre 2001 et 2015 et taux de dépendance en 2015		
	3,5 fois le PIB par	4,5 fois le PIB par	Salaires des
	habitant	habitant	fonctionnaires en 2001
Bénin	28,8	40,6	45,0
Burkina Faso	46,2	55,3	71,9
Cameroun	34,5	45,9	52,5
Congo	8,4	26,0	0,0
Côte-d’Ivoire	18,4	29,4	39,1
Guinée	30,4	43,8	32,0
Madagascar	41,9	53,0	46,7
Mali	47,0	57,2	67,2
Niger	48,7	57,6	76,0
Sénégal	31,8	44,2	54,2
Tchad	41,2	52,8	72,6
Togo	41,4	52,3	64,7
Ensemble	31,4	42,9	52,2

Source : Mingat (2004)

Les éléments qui précèdent amènent un éclairage relativement cru sur les évolutions récentes observées dans les profils de recrutement des enseignants. En effet, comme le souligne Lambert (2004) le recrutement d’enseignants contractuels est probablement la voie la plus simple pour diminuer les coûts salariaux, cela ne signifie pas pour autant que ce choix soit facilement accepté par les acteurs⁴¹.

Le tableau 2.1.6 fournit des éléments d’appréciation sur la diminution des coûts qu’induit le recrutement d’enseignants non-fonctionnaires. Les renseignements sur les niveaux de rémunération des enseignants sont sans ambiguïté. Ils montrent qu’un fonctionnaire coûte, en moyenne, deux fois plus qu’un contractuel de l’Etat et quatre fois plus qu’un maître de parents, orientant ainsi les choix vers les enseignants non fonctionnaires vu le contexte décrit

⁴¹ Cf. par exemple Alternatives Internationales de mars 2007 pour la position des syndicats d’enseignants.

précédemment. Cela dit, on observe une nouvelle fois une certaine variété selon les pays. Ainsi, au Burkina Faso, le salaire des contractuels est à peu près équivalent à celui des fonctionnaires alors qu'au Cameroun et au Mali il est quatre fois inférieur. En outre, il faut relever les niveaux de rémunération extrêmement bas des maîtres directement recrutés par les familles.

Tableau 2.1.6 : Niveau de rémunération des enseignants du primaire selon le statut
(en unités de PIB/habitant)

PAYS	Fonctionnaires	Contractuels de l'Etat	Maîtres de parents
Bénin (2002)	5,2	2,1	1,3
Burkina Faso (2002)	5,8	5,6	2,2
Cameroun (2002)	5,3	1,4	0,8
Congo (2003)	2,4	0,9	0,6
Côte d'Ivoire (2001)	4,8	-	-
Guinée (2000)	3,5	1,1	-
Mali (2000)	5,8	1,5	0,9
Niger (2000)	8,9	3,5	-
Sénégal (2003)	5,7	2,6	-
Tchad (2002)	8,2	-	2,3
Togo (2001)	6,4	3,3	1,3
Moyenne	5,6	2,4	1,3

Source : Mingat (2004)

La coexistence de deux catégories d'enseignants, recrutés par les pouvoirs publics, dont les niveaux de rémunération sont très éloignés, ne va pas sans poser de problème notamment quand on s'inscrit dans une perspective de long terme. Néanmoins, il ne faut pas perdre de vue les bénéfices en matière de scolarisation qu'a impliqués cette politique de recrutement d'enseignants non fonctionnaires.

2.1.5 L'impact sur la scolarisation du recrutement d'enseignants non fonctionnaires

Le recours au recrutement, parfois massif, d'enseignants non-fonctionnaires a permis d'accélérer le rythme de recrutement des enseignants, ce qui n'a évidemment pas été sans incidence sur la scolarisation. Il est possible d'estimer l'impact de l'emploi des non-fonctionnaires sur la scolarisation en comparant le nombre d'élèves effectivement scolarisés avec le nombre des élèves qui l'auraient été si le corps enseignant ne comptait que des enseignants fonctionnaires à budget et à taille moyenne de classe constants. A titre indicatif, des estimations ont été faites dans quelques pays et les résultats sont présentés dans le tableau 2.1.7. Elles nous fournissent des ordres de grandeur concernant l'impact des politiques de recrutement d'enseignants non-fonctionnaires⁴². On constate que des dizaines voire des centaines de milliers d'enfants sont scolarisés du fait du recrutement de ces enseignants. Dans un pays comme le Niger, ce nombre est considérable puisqu'il représente près de 50 % d'enfants scolarisés en plus par le pays en 2002 ; en revanche, on constate que l'impact est faible au Burkina Faso. Ceci s'explique par le fait que la proportion d'enseignants burkinabés non-fonctionnaires est moins élevée que dans les autres pays et que ceux-ci ont un niveau de rémunération proche de celui des enseignants fonctionnaires comme cela a été souligné précédemment.

Tableau 2.1.7 : Estimation du gain de scolarisation imputable au recrutement d'enseignants non-fonctionnaires

	Bénin (2002)	Burkina Faso (2002)	Cameroun (2002)	Mali (2003)	Niger (2002)	Togo (2001)
Nombre total d'élèves scolarisés	805 600	927 000	1 598 900	844 493	824 500	648 000
Nombre d'élèves qui auraient été scolarisés s'il n'y avait que des fonctionnaires	694 830	918 720	1 164 247	612 853	552 441	504 040
Gain de scolarisation imputable au recrutement de maîtres non-fonctionnaires	110 770	8 280	434 653	231 640	272 059	143 960
En % d'enfants scolarisés en plus	16 %	1 %	37 %	38 %	49 %	29 %

Source : Bernard, Tiyaab et Vianou (2004)

⁴² Les estimations ne prennent pas en compte ici le fait que certains pays ont bénéficié d'un appui extérieur important pour le recrutement d'enseignants contractuels, appui qui n'aurait pas existé sans l'adoption de cette politique. Cela signifie que les chiffres sont sous-estimés.

Par ailleurs, il faut considérer que ces estimations ont été réalisées sur les premières années de la mise en place des nouvelles politiques de recrutement de contractuels de l'Etat. Dans certains pays, les tendances observées sur l'accélération du recrutement des enseignants contractuels laissent penser que l'impact sur la scolarisation va en augmentant. On manque encore de données suffisamment récentes pour mieux apprécier les évolutions.

2.1.6 Au-delà de l'extension de la scolarisation, la question de la qualité de l'éducation

Dans cette section, nous avons vu que le défi de la scolarisation primaire universelle en 2015 implique une croissance de la scolarisation bien plus marquée que ce que le continent africain a pu connaître jusqu'ici. Cette croissance passe nécessairement par un recrutement massif d'enseignants qui ne pourra se faire dans certains pays, faute de ressources, sans financements extérieurs. A ce jour, pour répondre à ce défi, bon nombre de pays africains ont choisi de recruter des enseignants non-fonctionnaires avec des niveaux de rémunération sensiblement plus faibles. Dans 12 pays francophones d'Afrique subsaharienne, déjà au début des années 2000, au tout début de la mise en place des politiques de recrutement des enseignants contractuels, les enseignants non-fonctionnaires représentaient près de la moitié des effectifs d'enseignants en incluant les enseignants du privé et ceux recrutés et payés par les parents. Avec l'intensification du recrutement d'enseignants contractuels, ils sont devenus très largement majoritaires dans certains pays, comme nous avons pu le voir au Mali. Si la présence de ces enseignants non-fonctionnaires est effectivement massive, leur effet sur la scolarisation est aussi visible. Les estimations effectuées montrent des impacts considérables. Dans des pays comme le Cameroun et le Niger, il s'agit de plusieurs centaines de milliers d'enfants scolarisés.

S'il n'y a aucun doute que la politique de recrutement de ces nouveaux enseignants permette de scolariser plus d'enfants, l'effet sur les apprentissages reste cependant largement débattu au sein de la communauté éducative. En effet, ces enseignants outre leur statut ont aussi des caractéristiques différentes des autres enseignants, notamment en termes de formation professionnelle et d'ancienneté. La question des relations entre les caractéristiques des enseignants et les acquisitions des élèves a fait l'objet de nombreux travaux empiriques, notamment dans le cadre des fonctions de production éducatives, et aussi de nombreux débats. Une présentation des principaux résultats de la recherche sur ces aspects, pas seulement sur le

continent africain, nous permettra de mieux cerner les problématiques et les enjeux en matière de qualité de l'éducation autour de ces nouveaux enseignants.

2.2 Caractéristiques des enseignants et acquisitions scolaires : des convergences dans les résultats des recherches

Plusieurs particularités sont à prendre en considération quand on s'intéresse aux principaux résultats relatifs aux déterminants des acquis scolaires et notamment aux caractéristiques des enseignants. Tout d'abord, on observe une disproportion d'études entre les pays développés et les pays pauvres. Au sein des pays les plus riches, c'est la recherche aux Etats-Unis qui domine de loin la littérature existante (Leclercq, 2005). En outre, la recherche sur les déterminants des acquisitions scolaires ne se limite pas aux travaux des économistes et une multitude de travaux se développent qui ne sont pas toujours pris en compte malgré la qualité tout à fait remarquable de certains d'entre eux. Ensuite, on observe une prédominance des travaux anglophones en la matière qui ignorent souvent une bonne part des travaux francophones pourtant assez développés. Enfin, une difficulté est liée à l'inégale qualité sur le plan méthodologique des études. Les méta-analyses discriminant peu les études prises en compte ont amené des résultats controversés (Hanushek, 1995, 1997, 2003 ; Kremer, 1995 ; Krueger, 2003) qui incitent à la prudence. Il est donc préférable de porter une attention marquée à la qualité des études prises en compte tout en ne perdant pas de vue que la répétition de résultats identiques tend à renforcer la confiance de la communauté scientifique dans ces résultats.

Si on ne tient pas compte de ces différentes particularités, on risque d'aboutir à une vision biaisée des travaux sur les déterminants des acquisitions scolaires. En ce qui concerne les différents travaux disponibles, notre centre d'intérêt étant les pays pauvres, et plus particulièrement l'Afrique, nous accorderons une importance spécifique aux travaux réalisés dans ce contexte qui sont plus nombreux que ne le laissent penser les revues de littérature disponibles. Cela dit, il faut se donner quelques critères d'appréciation de la qualité des études pour ne pas accorder trop de crédit à des études présentant de trop grandes faiblesses méthodologiques. Une importance particulière est accordée, d'une part, à la disponibilité de données individuelles au niveau de l'élève sur de grands échantillons et, d'autre part, à un nombre d'observations au niveau de l'enseignant et/ou de l'école se situant dans les grands

nombres. Pour la modélisation, les résultats du chapitre précédent montrent la nécessité de prendre en compte la dimension cumulative du processus d'enseignement pour bien rendre compte des effets des caractéristiques des enseignants⁴³, ce qui implique l'utilisation de plusieurs tests d'évaluation des acquis. Cependant, les problèmes d'estimation des données hiérarchiques doivent aussi être pris en considération (Cf. annexe 1.2.2). Ces différents critères sont à considérer pour discuter des différentes études retenues et de la fiabilité de leurs résultats sur la relation entre les caractéristiques des enseignants et les acquisitions scolaires. En outre, nous avons cherché à présenter des travaux relatifs aux principales caractéristiques des enseignants. Les travaux discutés ci-après cherchent à représenter en quelque sorte le meilleur compromis par rapport à ces différentes conditions.

2.2.1 Caractéristiques des enseignants et acquisitions scolaires dans les pays développés

Aux Etats-Unis, la multitude des travaux sur les facteurs influençant les acquis scolaires permet de sélectionner des études de très grande qualité basées sur des données exceptionnelles. Nous nous focaliserons ici sur deux études, Rivkin, Hanushek et Kain (2005) et Krueger (1999), très souvent citées dans la littérature notamment en raison de la prise en compte des limites méthodologiques évoquées au chapitre 1, ce qui ne signifie pas bien sûr qu'elles sont exemptes de toute critique.

Rivkin, Hanushek et Kain (2005) utilisent une base de données extrêmement riche recueillie par le Texas School Project de l'Université du Texas à Dallas. Le Texas Assessment of Academic Skills (TAAS) a été administré chaque année depuis 1993 aux élèves inscrits dans les classes de 3^{ème} année (grade 3) jusqu'à la 8^{ème} année (grade 8). Les données regroupent tous les élèves inscrits dans les écoles publiques du Texas. Les auteurs ont utilisé les données de trois cohortes, chacune comprenant plus de 200 000 élèves dans environ 3 000 établissements publics du primaire (elementary schools) et du collège ou école moyenne (middle schools). De tels effectifs permettent d'obtenir une précision bien supérieure à celle des enquêtes sur échantillon limité. Les tests ont été élaborés pour évaluer la maîtrise des disciplines spécifiquement pour chaque niveau, ce qui implique que le test de

⁴³ Cf. notamment le cas du niveau académique de l'enseignant à Madagascar dans la section 2 du chapitre 1.

3^{ème} année est complètement différent de celui de 4^{ème} année. Ils bénéficient donc de données de panel permettant de prendre en compte la valeur ajoutée par chaque année d'enseignement.

Les auteurs constatent que les enseignants titulaires d'un diplôme de master n'apparaissent pas plus performants que leurs collègues. L'effet de l'expérience est particulièrement important la première année. Ainsi, les enseignants qui n'ont pas d'expérience ont de moins bons résultats que leurs collègues ; cela reste vrai dans une moindre mesure pour deux à trois années d'ancienneté. Toutefois, il semble qu'il n'y ait plus de gains au-delà de 3 ans. Globalement, les effets des caractéristiques des enseignants sont modérés et concentrés sur les plus jeunes élèves dans les premières classes de l'école primaire. Il ressort également que les caractéristiques des enseignants prises en compte expliquent peu de chose de la variété de la qualité de l'enseignant observée.

Plusieurs problèmes doivent cependant être relevés dans cette étude. Tout d'abord, il n'y a pas d'information permettant de relier directement les élèves à leurs enseignants. Ainsi, on sait quels sont les élèves de 3^{ème} année et les enseignants qui enseignent en 3^{ème} année pour chaque école, mais pas quel maître enseigne à quels élèves. Cela porte atteinte à la précision de la mesure des effets des enseignants malgré la prise en compte de variables visant à limiter ce problème comme le taux de mobilité des enseignants par établissement. Ensuite, le modèle utilisé cherche à expliquer la différence des acquis scolaires entre deux classes en faisant la différence entre le score de l'année n en classe c et celui de l'année $n+1$ en classe $c+1$. Là aussi, se pose un problème dans la mesure où les tests sont différents si bien qu'ils ne correspondent pas à une même échelle, la différence ainsi effectuée entre les scores aux tests ne traduit pas nécessairement un accroissement des acquisitions des élèves (Mc Caffrey et al., 2003)⁴⁴. Pour procéder de la sorte, il aurait fallu ramener les différents tests à une échelle commune, ce qui implique une construction spécifique des tests permettant un ancrage (Laveault et Grégoire, 1997 ; Demeuse, 2004).

⁴⁴ Il se peut par exemple que le test du niveau $N+1$ ait une structure différente du test de niveau N , en accordant par exemple plus de poids à une dimension particulière comme la grammaire, il y aurait alors un écart plus grand avec les élèves faibles en grammaire relativement aux autres dimensions (orthographe, compréhension de texte, etc.) sans que cela traduise réellement une différence de progression. Un problème similaire se pose si la difficulté des tests et de leurs différentes composantes n'est pas identique. Si les exercices de grammaire sont plus difficiles dans un test que dans l'autre, on sera à nouveau confronté à un biais.

Krueger (1999) s'intéresse à l'estimation des fonctions de production éducative dans le cadre d'une approche expérimentale. Le projet STAR (Tennessee Student/Teacher Achievement Ratio experiment) cherche à évaluer l'impact de la taille de classe sur les acquisitions des élèves. Le principe de départ de l'expérimentation est d'affecter aléatoirement les élèves dans trois types de classe différents : (i) des classes à petits effectifs (entre 13 et 17 élèves), (ii) des classes standard (entre 22 et 25 élèves), et (iii) des classes standard avec un enseignant supplémentaire pour aider. Les données ont été collectées à partir de la dernière année de maternelle jusqu'à la 3^{ème} année, c'est-à-dire durant quatre ans. La maternelle n'était pas obligatoire au début du projet en 1985-86, 2200 élèves sont entrés dans le projet en 1^{ère} année du primaire, respectivement 1600 et 1200 élèves ont intégré le projet en 2^{ème} et 3^{ème} années. Les nouveaux entrants ont été affectés aléatoirement dans les classes. Au total, 11 600 enfants ont été impliqués dans l'expérimentation pour seulement 80 écoles⁴⁵. Ces données, bien que très riches (l'auteur présente cette expérimentation comme la plus grande du genre menée aux Etats-Unis), sont touchées par les limitations classiques des enquêtes longitudinales (perte d'élèves, mobilité des élèves entre les types de classes, difficulté à contrôler les tailles réelles de classe, etc.) qui peuvent induire des biais dans l'estimation. L'auteur, en fonction des biais identifiés, propose des solutions pour les estimations afin de les rendre aussi fiables que possible. Si l'objectif de cette expérimentation porte sur les tailles de classe, les estimations réalisées fournissent aussi des résultats intéressants sur les caractéristiques des enseignants. Sur ce point, Krueger rejoint globalement les conclusions de Rivkin et al. (2005). Il ne constate pas d'effet significatif du diplôme de master pour l'enseignant et il observe un petit effet positif de l'ancienneté. Les élèves qui ont un enseignant avec 20 ans d'ancienneté ont des résultats en moyenne 3 % plus élevés que ceux qui ont un enseignant sans ancienneté. Il conclut que conformément à la littérature antérieure, les données STAR suggèrent que les caractéristiques mesurées des enseignants expliquent relativement peu des résultats aux tests des élèves.

Toutefois, certaines limites sont pointées par Hanushek (2003). L'absence de test initial des élèves au début de l'expérimentation ne permet pas de vérifier la qualité de l'affectation aléatoire réalisée. Plus important, alors que les résultats dépendent beaucoup des enseignants, on ne sait que peu de choses sur la façon dont ils ont été répartis concrètement dans les classes et qui aurait dû être une affectation aléatoire. Enfin, le choix des écoles n'a rien d'aléatoire

⁴⁵ Le coût du projet STAR est de 12 millions de dollars pour 4 ans.

puisqu'elles étaient volontaires et devaient être de taille suffisante pour accueillir au moins chacun des trois types de classe de l'expérimentation, ce qui excluait de fait les écoles de plus petite taille. Enfin, les administrateurs et les enseignants sachant que cette expérimentation pouvait déboucher sur de nouvelles ressources ont pu être influencés dans leurs comportements. Autant d'éléments qui limitent la portée des conclusions sur les données STAR et qui montrent qu'une généralisation pourrait être abusive.

On peut être frappé que ces deux études, réalisées par des chercheurs de renommée internationale et basées sur des données d'une richesse exceptionnelle, présentent des limites qu'on ne peut pas considérer comme anecdotiques. Elles permettent néanmoins de dégager des résultats relativement proches sur les caractéristiques des enseignants et qui, comme le souligne Krueger (1999), rejoignent ceux de la littérature existante aux Etats-Unis.

Dans les deux études précédentes, il n'est pas fait mention de la formation professionnelle des enseignants. C'est relativement fréquent dans les pays développés où la totalité des enseignants ont généralement suivi une formation professionnelle initiale identique, ce qui ne permet pas aux chercheurs d'effectuer les comparaisons nécessaires aux analyses. L'étude de Bressoux, Kramarz et Prost (2006) constitue ainsi une exception particulièrement intéressante dans le cadre des pays riches. Les auteurs cherchent à estimer l'effet de la formation professionnelle initiale des enseignants de l'école élémentaire sur les acquis des élèves de 3^{ème} année (CE2) en France. Ils profitent d'une particularité du système français qui permet à de jeunes diplômés de commencer à enseigner sans avoir suivi de formation professionnelle. Ils distinguent trois catégories d'enseignants : (i) ceux qui sont expérimentés, (ii) les novices sans formation professionnelle et (iii) les novices avec formation professionnelle. Les auteurs trouvent que les plus anciens sont dans des classes avec de meilleurs élèves et de meilleures conditions, alors que les novices formés ou non sont affectés dans des classes similaires. Ils se concentrent donc sur les enseignants novices pour éviter le problème de biais dû à l'ancienneté. Ils utilisent une enquête spécifique, suivant un plan quasi expérimental, réalisée en 1991-1992 auprès d'élèves et de maîtres de 3^{ème} année de l'enseignement primaire (Bressoux, 1996). L'échantillon couvre les élèves et les maîtres de 3^{ème} année dans 12 départements français. Tous les enseignants débutants ont été intégrés à l'échantillon. Au total, 3 842 élèves et 198 enseignants composent l'échantillon dont 102 enseignants débutants. Pour leurs analyses, les auteurs ne retiennent que les novices soit 66 débutants formés et 36 débutants non formés. Les auteurs n'observent pas de différence significative

entre les enseignants novices en français, alors qu'un écart d'environ 20 % d'écart-type est relevé en mathématiques. Ils notent également que le champ de spécialisation à l'université a un effet significatif. Ceux qui ont suivi une formation en science (12 %) sont associés à de meilleurs résultats de leurs élèves en mathématiques mais pas en français. Cependant, parmi les enseignants novices, ceux n'ayant pas précisé leur spécialisation semblent également mieux faire apprendre leurs élèves aussi bien en français qu'en mathématiques, ce qui relativise le constat précédent. Par ailleurs, les auteurs remarquent que les enseignants selon leur spécialisation ne sont pas affectés aléatoirement dans les classes. Les enseignants avec une spécialisation en science ou inconnue et qui enregistrent les meilleurs résultats sont aussi ceux qui ont les meilleurs élèves. Bressoux, Kramarz et Prost constatent également que la formation semble plutôt aider les enseignants à faire plus progresser les élèves dans les classes d'assez bon niveau mais pas nécessairement dans les classes de faible niveau.

Ce travail souffre néanmoins d'une limite que les auteurs discutent. En effet, les enseignants novices ne sont pas choisis aléatoirement pour suivre la formation. Ceux qui suivent la formation ont réussi un concours alors que ceux qui n'ont pas eu de formation ont échoué au concours et ont été classés sur liste d'attente. Cependant, les auteurs remarquent que l'année considérée (1991) est exceptionnelle car nettement moins de personnes ont été sélectionnées au concours, ce qui fait que ceux qui sont sur liste d'attente cette année-là, et qui enseignent sans formation, auraient été pris s'ils avaient passé le concours une autre année. Malgré tout, on ne sait pas exactement si leurs résultats au concours étaient réellement comparables à ceux des enseignants formés qu'ils côtoient et qui ont passé le concours une année plus tôt ; rien ne garantit en effet qu'entre deux cohortes le niveau est similaire. On ne peut donc exclure un biais même s'il est probablement modéré. Un autre point troublant porte sur les différences entre les disciplines qui relativisent quand même l'incidence globale de la formation. En outre, l'écart de 20 points d'écart-type observé en mathématiques apparaît somme toute modéré quand on considère qu'il s'agit de la plus-value apportée par la formation professionnelle puisqu'on compare des enseignants formés à des enseignants sans aucune formation.

Une autre catégorie de formation doit être considérée, il s'agit de la formation en cours d'emploi ou formation continue des enseignants. Angrist et Lavy (2001) constatent que peu d'études se sont intéressées à la formation sur le tas et à la formation en cours d'emploi ou formation continue des enseignants ; alors que d'autres dimensions comme la formation

académique du maître ou l'expérience professionnelle sont plus souvent étudiées. Ils se réfèrent aux travaux de Farrell et Oliveira (1993) sur les pays en développement qui concluent que la formation avant l'emploi est nécessaire pour maîtriser les contenus disciplinaires mais que la formation en cours d'emploi est essentielle pour acquérir les compétences pédagogiques. Ces dernières étant plus faciles à acquérir quand la personne dispose d'une expérience de l'enseignement.

Angrist et Lavy évaluent un programme de formation continue des enseignants (le programme 30 Towns) mis en œuvre à partir de 1995 dans les écoles élémentaires de Jérusalem et d'autres villes israéliennes⁴⁶. Cette formation supplémentaire a été réalisée sur une base hebdomadaire par des instructeurs extérieurs à l'école qui se sont focalisés sur l'amélioration des techniques d'enseignement en hébreu, en mathématiques et en anglais. Dix écoles élémentaires ont été concernées à Jérusalem, sept laïques et trois religieuses mais nous considérerons uniquement les écoles laïques ici⁴⁷. Ce choix n'a pas été fait sur une base aléatoire et les auteurs ne précisent pas les critères de sélection des écoles. Les auteurs ont choisi six écoles de contrôle comme base de comparaison. Par ailleurs, les données utilisées concernent les scores aux tests des élèves qui ont terminé leur 4^{ème} année en 1994 (avant le début du programme), puis leur 5^{ème} année en 1995 et leur 6^{ème} année en 1996. Cela correspond à un peu plus de 800 élèves au total répartis à peu près également entre écoles du programme et écoles de contrôle. Les auteurs notent que les déperditions d'élèves sont plus marquées dans les écoles de contrôle mais elles seraient dues à des classes entières qui n'auraient pas passé les tests. En outre, les parents des écoles du programme ont un niveau de formation inférieur à celui des autres parents.

Pour surmonter les biais potentiels liés aux différences initiales entre les élèves des écoles du programme et ceux des écoles de contrôle, les auteurs font appel à plusieurs types d'analyse statistique. Si les auteurs suggèrent qu'on peut considérer un effet global d'au moins 0,25 écart-type de la formation pour les écoles laïques, les résultats des différentes estimations présentées sont plus nuancés. On observe notamment des différences notables entre l'hébreu et les mathématiques où les effets sont moins marqués et peu significatifs si on considère les erreurs types qui prennent en compte la structure hiérarchique des données tirées des travaux

⁴⁶ Environ 960 000 \$ par an ont été consacrés aux écoles élémentaires.

⁴⁷ Pour les écoles religieuses, différents problèmes de données (démarrage tardif du programme, faible nombre d'écoles, données uniquement sur l'évaluation des mathématiques, etc.) font que les résultats sont peu concluants.

de Moulton (1986). La dernière estimation de l'article faite sur la base d'un appariement des élèves en fonction de leurs scores en 1994 est une bonne illustration de la différence entre disciplines. Les auteurs obtiennent un coefficient de 0,25 en mathématiques avec une erreur type de 0,16, et en hébreu le coefficient est de 0,40 avec une erreur type de 0,15. Il faut noter que l'appariement au niveau élève laisse sans réponse d'éventuels biais contextuels puisque les effets de groupe ne sont pas pris en compte.

Les auteurs ne donnent pas d'autres informations sur les caractéristiques des enseignants et ils ne fournissent pas non plus l'importance en termes de variance expliquée du programme 30 towns.

Outre le fait que l'étude ne porte au final que sur 13 écoles laïques, les différents traitements statistiques mis en oeuvre par Angrist et Lavy présentent chacun des limites claires et ne permettent donc pas de corriger entièrement les biais potentiels liés aux différences initiales entre écoles et élèves. Ils fournissent des ordres de grandeur où pourrait se situer l'effet du programme. Cependant, on peut s'interroger sur un éventuel effet Hawthorne lié à cette formation dont l'hypothèse est difficile à écarter dans le contexte de ce programme. On constate aussi que les écoles du programme ne bénéficiaient pas initialement de formation des maîtres en hébreu alors que les autres écoles avaient plus de 5 heures de formation par semaine. Par contre, le temps de formation était sensiblement équivalent en mathématiques. L'effet du programme plus net en hébreu qu'en mathématiques pourrait donc être un effet de rattrapage. Par ailleurs, comme le soulignent Jacob et Legfren (2004), la formation des enseignants n'était pas la seule composante du programme 30 towns. La mise en place d'un centre d'apprentissage pour aider les élèves en difficulté après l'école et l'appui spécial apporté aux élèves immigrants et à leur famille ont probablement eu un effet qui n'est pas différentiable de celui de la formation. Ces auteurs, sur la base d'une étude menée à Chicago à la fin des années 1990 selon une approche quasi expérimentale, se montrent beaucoup plus pessimistes sur l'impact éventuel de la formation continue sur les performances des élèves. Les écoles publiques de Chicago où moins de 15 % des élèves atteignaient les normes nationales en anglais étaient mises en probation (soit 71 écoles sur 489) et recevaient une aide financière pour la formation des enseignants. Les auteurs profitent de la discontinuité des données consécutive au mode de sélection des écoles pour analyser l'impact des formations dispensées. La formation apparaît ainsi complètement inefficace pour améliorer les acquis scolaires des élèves. Encore une fois, nous sommes confrontés à des résultats contrastés à l'image de la littérature sur la question de l'effet de la formation des enseignants (Jacob et Legfren, 2004).

Les études qui viennent d'être présentées, sans prétention d'exhaustivité, ont le mérite de fournir des résultats relatifs aux principales caractéristiques des enseignants, à savoir le niveau de formation académique, la formation professionnelle et l'ancienneté, dans les pays développés. En ce qui concerne, la formation académique, les deux études sur les Etats-Unis aboutissent à la même conclusion : la détention d'un diplôme de master n'a pas d'influence sur les progrès des élèves. Il convient ici d'attirer l'attention sur le fait que la plupart des enseignants aux Etats-Unis ont fréquenté l'université, ce qui laisse peu de variabilité pour l'analyse. L'ancienneté de l'enseignant tend à avoir un effet significatif, quoique modéré, mais essentiellement pour les premières années d'exercice du métier et dans les premières classes du primaire. Enfin, les résultats sur la formation professionnelle sont mitigés. L'étude menée en France par Bressoux et al. (2006) sur la formation professionnelle initiale laisse entrevoir au mieux un léger effet positif. Le résultat n'étant pas avéré dans les deux disciplines considérées et certains problèmes de biais n'étant pas complètement écartés, il convient d'être prudent. Pour ce qui est de la formation continue, les travaux de Angrist et Lavy (2001) en Israël et ceux de Jacob et Legfren (2004) aux Etats-Unis ne poussent pas à l'optimisme quant à l'efficacité de ce type d'intervention. Enfin, un constat qui se dessine dans la plupart de ces études est celui de la part modérée des caractéristiques observées des enseignants dans l'explication des performances des élèves alors que, paradoxalement, l'effet de l'enseignant semble marqué.

2.2.2 Caractéristiques des enseignants et acquisitions scolaires dans les pays pauvres

Il a été signalé que les recherches menées dans les pays les plus pauvres sont nettement moins nombreuses que celles réalisées dans les pays développés et notamment aux Etats-Unis. De plus, il faut souligner qu'on ne dispose pas de bases de données comparables à celles du projet STAR ou du TAAS dans les pays pauvres. On peut comprendre aisément que dans des pays où les ressources sont rares mais aussi insuffisantes, on ne consacre pas des montants aussi importants que dans les pays riches à ce type d'études⁴⁸. Toutefois, une accélération notable

⁴⁸ Les 12 millions de dollars du projet STAR correspondent à environ 13 % du budget total de l'éducation au Niger en 1986 c'est-à-dire au démarrage du projet.

des travaux sur les pays pauvres est à noter au cours de la dernière décennie et une quantité appréciable de travaux sont désormais disponibles y compris pour le continent africain.

Aux Philippines, Tan, Lane et Coustère (1997) utilisent des données de 1990 et 1991 pour étudier les effets des variables relatives à l'enseignant et à l'école sur les acquis en philippin et en mathématiques de 2 293 élèves de 1^{ère} année répartis dans 110 écoles choisies aléatoirement. Les auteurs n'ont trouvé d'effet significatif ni pour les enseignants titulaires d'un master (le coefficient est négatif), ni pour l'expérience professionnelle. En fait, la seule variable significative concernant l'enseignant est le score de l'enseignant au test : il se traduit par un effet positif modéré sur les acquis des élèves dans les deux disciplines considérées. Par contre, les auteurs ne prennent pas en compte la formation professionnelle de l'enseignant dans leur modèle. L'ensemble de ces résultats amène les auteurs à conclure que le renforcement des qualifications des enseignants est l'option la moins coût-efficace parmi celles considérées dans l'étude. Les auteurs fournissent également les différentes valeurs prises par le R² ajusté selon la spécification du modèle (Cf. tableau 2.2.1). Cela permet d'observer que les variables relatives aux classes, aux enseignants et aux écoles expliquent entre 7 % et 12 % de la variance selon la discipline, ce qui laisse entre 13 % et 16 % de la variance expliquée par le modèle avec les variables muettes écoles sans explication. On peut ainsi observer que ce qu'on ne sait pas expliquer au niveau de l'établissement est aussi important voire un peu plus important que ce que l'on sait expliquer. Ce résultat n'est pas sans rappeler ce qui a été observé précédemment dans les pays développés et notamment aux Etats-Unis.

**Tableau 2.2.1 : Pouvoir explicatif du modèle de Tan et al. (1997)
selon les différentes spécifications**

Spécifications de la régression	valeurs du R ² ajusté	
	Mathématiques	Philippin
Variables individuelles élève	0.22	0.37
Variables individuelles et variables muettes écoles	0.50	0.57
Variables individuelles, classe, enseignant et école	0.34	0.44

A partir d'une enquête du PASEC à Madagascar en 1997-98 auprès d'un échantillon représentatif des élèves de 2^{ème} et 5^{ème} années de l'école primaire testés en français et mathématiques, Lassibille et Tan (2003) cherchent à mesurer l'effet des écoles privées sur les

acquisitions des élèves. Au total, 90 écoles publiques et 25 écoles privées ont participé à l'enquête, ce qui représente en tout 1993 élèves en 2^{ème} année et 1900 en 5^{ème} année. Le protocole d'enquête de l'étude ayant été conçu pour avoir un échantillon représentatif des élèves et non pas comparer les deux types d'écoles, les auteurs ont dû prendre en compte les biais de sélection potentiels. De plus, ils intègrent dans leurs modèles trois variables relatives aux enseignants : nombre d'années d'ancienneté, nombre d'années de formation académique, nombre de mois de formation professionnelle initiale. Les coefficients attachés à ces variables sont proches de 0 et jamais significatifs à l'exception de l'ancienneté professionnelle en 5^{ème} année où 10 ans d'ancienneté se traduisent par 7 % d'écart-type en plus pour le score des élèves, un chiffre qui apparaît très modeste. On observe également que l'introduction des variables relatives à la classe et à l'enseignant entraîne une augmentation du R² de seulement 1 % en 2^{ème} année et 2 % en 5^{ème} année. Le moins que l'on puisse dire, c'est que les variables qui caractérisent le maître n'apparaissent pas déterminantes dans cette étude.

Michaelowa (2001a) se montre beaucoup plus optimiste sur l'influence des caractéristiques de l'enseignant sur les acquisitions des élèves à l'école primaire en Afrique francophone. L'auteure utilise les données PASEC de cinq pays (Burkina Faso, Cameroun, Côte d'Ivoire, Madagascar et Sénégal), soit plus de 500 écoles et plus de 10 000 élèves, en 5^{ème} année de l'école primaire en français et mathématiques. Elle emploie un modèle hiérarchique à trois niveaux (élève, école et pays), tiré des travaux de Bryk et Raudenbush (1992), pour analyser l'effet d'un ensemble de facteurs. En outre, elle introduit de nouvelles spécifications des variables de formation et d'ancienneté des enseignants. Les variables de formation (académique et professionnelle) sont continues et basées sur la durée respective des formations. L'hypothèse d'une influence non linéaire de l'ancienneté suggérée par Bernard (1999a) est testée de façon concluante à travers l'introduction d'une forme quadratique. Il s'en dégage des effets positifs pour ces variables. L'ancienneté optimale qu'on déduit de la forme quadratique serait à environ 22 ans. D'autres variables comme l'absentéisme du maître, sa motivation et ses autres activités, dont les cours particuliers, ont été prises en compte dans les analyses. Si l'absentéisme apparaît logiquement avoir un effet négatif, il semble que le fait que l'enseignant ait des activités hors de l'école soit associé avec de meilleurs résultats des élèves. Cela pourrait traduire un plus grand dynamisme de l'enseignant plutôt qu'un désengagement comme le souligne l'auteure. Considérant le genre de l'enseignant, Michaelowa remarque que les filles semblent mieux apprendre lorsque c'est une femme qui

enseigne. Enfin, le statut de l'enseignant apparaît avoir un effet très marqué et met en lumière une plus grande efficacité des enseignants non-fonctionnaires.

Toutefois, plusieurs problèmes sont à considérer par rapport à ces résultats. Tout d'abord, le niveau initial des élèves n'a pas été contrôlé dans le modèle. On ne peut donc exclure, par exemple, que les résultats obtenus sur les variables de formation et d'ancienneté soient biaisés si les enseignants les mieux formés et les plus expérimentés sont plus souvent affectés dans de bons établissements. Leurs caractéristiques capteraient alors les effets du contexte scolaire dans lequel ils évoluent plutôt que de refléter leur efficacité propre. De plus, l'analyse sur les données fusionnées de seulement cinq pays peut cacher pour une variable donnée des effets spécifiques à un ou deux pays plutôt qu'une tendance générale à tous les pays. La prise en compte d'effets fixes pays aurait pu atténuer ce problème. Enfin, l'absence de traitement des valeurs manquantes pour les variables utilisées dans les modèles⁴⁹ aboutit à perdre environ la moitié de l'échantillon initial d'élèves, ce qui peut induire des biais importants.

Michaelowa et Wechtler (2006) ont tenu compte des problèmes évoqués précédemment. Ainsi, les auteures examinent la sensibilité des résultats à la prise en compte du score initial et au type de modèle pour huit pays ayant participé à une évaluation du PASEC. En outre, elles réalisent des analyses similaires sur les pays anglophones ayant participé au programme SACMEQ⁵⁰. Les auteures utilisent deux types de modèles, un modèle hiérarchique à deux niveaux⁵¹ et un modèle de régression pondérée (survey regression), et incluent dans leurs modèles des effets fixes pays. En outre, les analyses sur données PASEC sont réalisées sur les 2^{ème} et 5^{ème} années pour chaque discipline⁵² (français et mathématiques). Les résultats relatifs aux principales caractéristiques des enseignants dans les modèles intégrant le score initial de l'élève sont présentés dans le tableau 2.2.2. Ces résultats amènent à beaucoup plus de circonspection quant à l'influence des caractéristiques des enseignants. On observe des coefficients relativement faibles et rarement significatifs.

⁴⁹ Par exemple, si un enseignant n'a pas donné son ancienneté dans un questionnaire, on va perdre l'ensemble des observations de sa classe. Il en va ainsi pour toutes les variables introduites dans le modèle ce qui explique la forte déperdition observée.

⁵⁰ Southern Africa Consortium for Monitoring Education Quality

⁵¹ L'estimation du modèle à deux niveaux est basée sur les moindres carrés généralisés réalisables (feasible generalized least squares) [xtreg option re dans le logiciel STATA].

⁵² Les auteures considèrent comme variable expliquée le pourcentage de réponses correctes (0-100 %) avec une moyenne de 37 % en français et de 39 % en mathématiques et un écart-type de 20.5 % pour la 5^{ème} année; une moyenne de 47.2 % en français et 45 % en mathématiques avec un écart-type respectivement de 27 % et 26 % pour la 2^{ème} année.

**Tableau 2.2.2 : Résultats relatifs aux variables maître des modèles
de Michaelowa et Wechtler (2006)**

Variables	2 ^{ème} année		5 ^{ème} année	
	mathématiques	français	mathématiques	français
Formation académique (0=primaire non achevé- 6 : au moins trois années de supérieur)	-0,01/0,00	-0,19/-0,25	0,34/0,42	0,03/0,02
L'enseignant a suivi une formation professionnelle initiale	-0,27/-0,40	0,77/0,99	-0,64/-0,69	0,06/0,09
Nombre moyen de formations continues par an au cours des 5 dernières années	0,49/0,52	0,77/0,99	0,37/0,40	0,64*/0,67**
Années d'ancienneté	-0,08/-0,07	-0,04/-0,04	0,09**/0,10**	0,11***/0,11**
Enseignant non-fonctionnaire	-1,03/-1,24	-1,69/-2,27*	0,70/0,98	1,14/1,3
Le maître est une femme	1,85**/1,87**	0,65 / 0,76	-1,54**/-1,99**	-0,10/0,28

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

En 2^{ème} année, aucune des variables n'est significative dans les deux disciplines étudiées et seulement deux variables se voient associer des coefficients significatifs. Ainsi, le fait d'avoir une femme comme maître a un effet positif en mathématiques uniquement, alors qu'on observe un résultat inverse en 5^{ème} année. Quant au statut de l'enseignant, le coefficient est significatif dans un seul modèle et au seuil de 10 %, ce qui ne permet pas d'accorder une grande confiance à ce résultat. En 5^{ème} année, seule l'ancienneté est significative dans les deux disciplines, mais avec un coefficient très modéré. Il est vrai que les auteurs n'ont pas estimé une forme quadratique de cette variable.

La formation académique de l'enseignant n'apparaît pas avoir ici un effet sensible sur les acquisitions scolaires puisque la variable est associée à des coefficients très modestes, et même parfois négatifs, jamais significatifs. La formation professionnelle initiale n'est pas mieux lotie. Elle ne présente jamais d'effet significatif dans les modèles estimés. On observe aussi des coefficients négatifs en mathématiques pour la 5^{ème} année. La formation continue, quant à elle, présente un effet positif, quoique modéré, et significatif en français en 5^{ème} année. Dans l'ensemble, on ne peut que constater l'influence limitée des variables de formation dans les analyses.

Ces résultats sur les caractéristiques des enseignants qui montrent globalement une influence pour le moins mitigée rejoignent ceux de Bernard, Tiyab et Vianou (2004) toujours sur les données du PASEC. Ceux-ci réalisent une analyse individuelle pour les différents pays où ont été menées des évaluations PASEC. Ils n'observent pas d'effet systématique des différentes caractéristiques des enseignants sur les acquis des élèves. Mingat et Suchaut (2000) utilisant les résultats de 15 études empiriques réalisées en Afrique francophone aboutissent à des conclusions similaires quant au rôle modéré des caractéristiques observées des enseignants dans le processus d'acquisition.

Ces résultats assez peu concluants ne sont pas sans rappeler, pour des variables comme la formation académique ou la formation professionnelle, ceux observés dans les pays développés. Toutefois, il convient de souligner certaines limites des données sur lesquelles se fondent les études qui viennent d'être présentées. Ces études sont basées sur des échantillons représentatifs des élèves ; elles ne sont pas conçues pour traiter spécifiquement l'incidence de telle ou telle caractéristique de l'enseignant sur les acquis scolaires. Cela peut avoir différents types de conséquences comme des mesures imprécises de certaines variables ou des problèmes de biais. La formation professionnelle initiale par exemple est mesurée sur la base de la durée. On distingue généralement les formations longues (un an et plus) des formations courtes voire de l'absence de formation. Toutefois, un enseignant qui a reçu une formation professionnelle initiale d'une durée d'un an dans une période récente, disons deux ou trois années plus tôt, aura probablement suivi une toute autre formation que l'enseignant qui a connu une même durée de formation initiale mais il y a 20 ans de cela. En effet, on conçoit facilement que les contenus et les méthodes de formation évoluent. La variable de formation utilisée habituellement compile en fait les différentes formations existantes ou ayant existé qui ont la même durée et nous donne dans le modèle l'effet moyen de ces formations. L'absence d'un effet positif et significatif n'est pas pour autant rassurante puisqu'elle traduit qu'en moyenne les formations dispensées n'ont pas d'incidence sur les apprentissages des élèves. Toutefois, l'analyste n'est pas en mesure de dire si certaines de ces formations se sont avérées plus efficaces que d'autres. On peut prolonger ce type de questionnement à la formation académique en se demandant si, en moyenne, être titulaire d'un diplôme comme le Bac, par exemple, obtenu en 2005 est équivalent à être titulaire du même diplôme mais obtenu en 1980. On soupçonne, même si c'est moins flagrant que pour la formation professionnelle, que cela engendre une relative imprécision des mesures. En ce qui concerne des variables comme

le genre ou le statut de l'enseignant, des biais liés à l'affectation dans les écoles doivent être envisagés. Par exemple, il est relativement classique de voir les enseignants non-fonctionnaires affectés majoritairement dans des zones ou des écoles réputées difficiles où les autres enseignants rechignent à se rendre. Ce type de biais est difficile à contrôler correctement *a posteriori* dans les modèles statistiques surtout s'il renvoie à des caractéristiques non observées de l'environnement scolaire.

Il est donc utile d'avoir recours à des protocoles d'enquête spécifiques pour analyser ces questions de façon plus précise. Ce type de travaux demeurent relativement peu fréquents dans les pays pauvres et plus particulièrement en Afrique même si les choses évoluent progressivement. Ainsi, le PASEC a mené différentes études dites thématiques pour traiter des questions spécifiques comme l'impact du statut et de la formation des enseignants. La principale originalité de ces études est que l'échantillonnage s'est fait sur la base des caractéristiques des enseignants afin de permettre des comparaisons légitimes entre catégories d'enseignants plutôt que de viser la représentativité des élèves.

En Guinée, deux évaluations du programme de formation initiale des maîtres de Guinée (FIMG) ont été menées par le PASEC en 1999-2000 et 2004-2005. Ce programme, avec des formations d'une durée plus courte⁵³ axées sur la professionnalisation, visait à former un plus grand nombre d'enseignants (2000 par an contre un peu plus de 700 auparavant) pour répondre au défi de la scolarisation primaire universelle. Les nouvelles formations s'accompagnaient également d'une contractualisation des nouveaux enseignants recrutés. Les évaluations menées ont cherché à comparer les enseignants FIMG avec les autres enseignants. La première évaluation qui concernait seulement les deux premières cohortes a conclu à des résultats très proches entre les différentes catégories d'enseignants, légèrement inférieurs en 2^{ème} année pour les enseignants contractuels et quasiment identiques en 5^{ème} année (PASEC, 2003). Ce résultat est d'autant plus remarquable que les enseignants FIMG affichaient au plus une année d'expérience professionnelle en début d'année scolaire et il n'était pas possible de contrôler l'ancienneté professionnelle dans les modèles. Cette évaluation ayant été menée au début du processus, une seconde étude a été réalisée en 2004-2005. Pour cette nouvelle étude,

⁵³ En fait, deux catégories de formation sont à prendre en compte. La première comprenait 3 mois de formation, suivis d'une année scolaire avec une classe à charge, mais avec un accompagnement pédagogique, et enfin 3 nouveaux mois de formation. La seconde catégorie correspondait à 9 mois de formation puis une année scolaire avec une classe à charge mais avec un accompagnement pédagogique. Cela permettait de former deux cohortes au cours d'une année civile.

l'échantillon a été agrandi pour avoir une meilleure précision des estimations. Au total, ce sont 323 enseignants qui ont participé à l'étude, soit environ 50 % de plus que pour l'enquête précédente.

Tableau 2.2.3 : Les enseignants de l'échantillon PASEC Guinée

Formation professionnelle	2 ^{ème} année	5 ^{ème} année	Total
Classique (ENI)	46	49	95
FIMG 9-9	50	50	100
FIMG 3-9-3	40	40	80
FIME 9-9-3	19	13	32
Autres	4	12	16
Total	159	164	323

Source : PASEC (2006b)

Les conclusions confortent celles de l'étude précédente et montrent que les enseignants FIMG tendent à être plus efficaces en 2^{ème} année tandis que les différences en leur faveur ne sont pas statistiquement significatives en 5^{ème} année (PASEC, 2006b). Bien que les éventuels problèmes de biais liés à l'affectation des enseignants ne sont pas discutés dans le rapport du PASEC, le protocole d'enquête suivi permet d'être relativement confiant dans les tendances présentées. Il reste que le statut et la formation sont inextricablement liés dans le contexte guinéen et qu'on ne saurait distinguer l'effet de l'un ou de l'autre.

Le statut de l'enseignant est précisément au centre de deux autres études thématiques du PASEC au Mali (PASEC, 2004b) et au Niger (PASEC, 2004a) durant l'année scolaire 2000-2001. Ces études traitaient de l'incidence des enseignants contractuels sur les acquisitions scolaires des élèves. Le protocole d'enquête retenu était basé sur l'appariement des enseignants contractuels et fonctionnaires. Pour chacun des niveaux enquêtés (2^{ème} et 5^{ème} années), à chaque enseignant contractuel enquêté, un enseignant fonctionnaire d'une école voisine (aussi proche que possible) était également enquêté. En outre, le budget étant limité, il a été décidé d'enquêter la classe de l'autre niveau de l'école choisie quel que soit le profil de l'enseignant. Au final, les échantillons regroupent des enseignants fonctionnaires et contractuels qui travaillent dans des conditions comparables (Cf. tableau 2.2.4).

**Tableau 2.2.4 : Répartition selon le statut des enseignants
des échantillons PASEC Mali et Niger**

	Mali			Contractuel	Niger	
	Contractuel	Fonctionnaire	Total		Fonctionnaire	Total
2^{ème} année	83	49	132	66	61	127
5^{ème} année	59	76	135	30	97	127
Total	142	125	267	96	158	254

On observe que malgré le protocole d'enquête identique, on ne retrouve pas la même répartition des enseignants aux deux niveaux. En effet, dans les deux pays, les enseignants contractuels sont nettement moins nombreux en 5^{ème} année, traduisant ainsi des préférences dans la répartition des enseignants. Les classes les moins prestigieuses, c'est-à-dire les classes de début de cycle, sont affectées en priorité aux enseignants contractuels.

Bernard et al. (2004) présentent les principaux résultats de ces études (Cf. tableau 2.2.5). Au Mali, les enseignants contractuels tendent à obtenir de meilleurs résultats que leurs collègues pour les deux niveaux considérés. Au Niger, le résultat est moins net puisqu'il n'y a pas de différence significative en 2^{ème} année alors qu'en 5^{ème} année les enseignants fonctionnaires semblent se montrer plus efficaces. Toutefois, si le protocole d'enquête tend à se rapprocher d'une approche quasi expérimentale, les analyses menées n'ont que peu discuté la qualité de l'appariement réalisé ainsi que les problèmes de biais éventuels.

**Tableau 2.2.5 : Efficacité pédagogique des enseignants non-fonctionnaires
en comparaison aux fonctionnaires au Mali et au Niger**

(en % d'écart-type)

Pays	2 ^{ème} année	5 ^{ème} année
Mali	+ 0.24**	+0.22**
Niger	-0.08	-0.27**

** significatif au seuil de 5 % (intervalles de confiance robustes)

Source : Bernard et al. (2004)

Bourdon, Frölich et Michaelowa (2006) utilisent la base de données PASEC du Niger mais procèdent à un appariement des classes en fonction du statut de l'enseignant. Ils tentent ainsi

de sélectionner des classes de contrôle avec un enseignant fonctionnaire ayant des caractéristiques similaires aux classes qui sont dirigées par un enseignant contractuel. *In fine*, ils n'observent pas de différences significatives dans les acquisitions scolaires selon le statut de l'enseignant, ni en 2^{ème} année ni en 5^{ème} année.

Plusieurs problèmes se posent dans l'évaluation des différences de performance entre contractuels et fonctionnaires malgré les données spécialement collectées pour cela. Tout d'abord, il y a un manque de perspective temporelle puisque les études ont été menées peu de temps après la mise en place de la politique des enseignants contractuels au Mali comme au Niger. Il sera donc nécessaire de procéder à de nouvelles évaluations pour compléter ces premiers résultats. Ensuite, il est très difficile de différencier les enseignants selon la seule caractéristique du statut. En général, la formation professionnelle et l'expérience sont extrêmement liées au statut, ces contractuels de la première génération étant le plus souvent moins expérimentés et moins formés. On observe également que des différences de motivation assez marquées peuvent distinguer les enseignants contractuels des fonctionnaires (PASEC, 2004b). Au total, il reste encore relativement délicat de mesurer avec précision les différences de performance entre enseignants imputables au seul statut. Les mesures disponibles englobent différentes dimensions, mais elles n'interdisent pas de conclure sur l'effet global de cette politique des enseignants contractuels sur les acquis scolaires pour ce qui est du court terme.

Ces travaux spécifiques permettent de confirmer les tendances générales observées jusqu'ici dans les pays africains mais aussi, c'est à souligner, dans les autres régions du monde. Les caractéristiques observables de l'enseignant n'ont qu'une influence relativement modérée sur les acquisitions scolaires. Ce type de constat est généralement assez mal accepté par le corps enseignant qui l'assimile à une tentative de minimiser le rôle de l'enseignant dans le processus d'apprentissage. Il serait toutefois hasardeux de considérer que l'influence de l'enseignant se limite aux caractéristiques que nous venons d'évoquer. C'est précisément ce que mettent en lumière Murnane et al. (2005) à partir d'une étude sur les classes de 5^{ème} année des écoles publiques et privées de Bogota en Colombie. A partir d'une stratification prenant en compte les secteurs public et privé ainsi que les caractéristiques socioéconomiques des écoles, les auteurs choisissent aléatoirement les écoles. L'échantillon

final comprend 73 écoles⁵⁴, 97 classes de 5^{ème} année et un peu plus de 3000 élèves⁵⁵. Enfin, une spécificité importante de l'étude est que 13 enseignants de mathématiques, sur les 77 de l'enquête, enseignaient à plusieurs classes, ce qui permet aux auteurs de réaliser une analyse originale visant à différencier l'effet de la classe et l'effet de l'enseignant. Il ressort que certains enseignants sont plus efficaces que d'autres mais qu'il y a peu de preuves qu'un même enseignant soit inégalement efficace selon la classe où il exerce. Ces résultats sont à prendre avec précaution puisqu'ils ne reposent que sur un nombre restreint d'individus. Les auteurs constatent également pour l'ensemble de l'échantillon que la formation académique et l'expérience professionnelle n'expliquent qu'une part modeste des différences d'efficacité des enseignants. En effet, sur les 35 % de variance expliquée par le modèle avec les effets fixes enseignants, seuls 20,6 % sont expliqués par le modèle prenant en compte entre autres les caractéristiques des enseignants, ce qui laisse près de 15 % de variance non expliquée. Cela rejoint d'autres résultats déjà évoqués, et notamment Rivkin et al. (2005), qui soulignent que les caractéristiques observables de l'enseignant n'expliquent qu'une part modeste de l'influence globale de l'enseignant sur les acquisitions scolaires des élèves.

Les différentes études qui ont été présentées permettent de faire le point sur les relations entre les caractéristiques des enseignants et les acquis scolaires dans les pays pauvres avec un accent particulier sur les pays africains. Si on ne dispose pas de bases de données aussi riches que celles qui sont utilisées au Nord notamment aux Etats-Unis, un nombre d'études appréciable permet d'accumuler des résultats et de dégager des tendances. En ce qui concerne la formation académique, malgré la grande variété des situations des enseignants sur ce plan, les études ne permettent pas de dégager une incidence notable. Des conclusions à peu près similaires émergent pour la formation professionnelle. Les études sur la Guinée, si elles mettent en lumière une relative efficacité de la formation FIMG, montrent aussi que les écarts sont modérés et variables selon les niveaux et/ou les disciplines. L'influence de l'ancienneté n'apparaît pas très marquée. Michaelowa et Wechtler (2006) et Lassibille et Tan (2003) ne montrent qu'un effet limité en 5^{ème} année de l'école primaire. Enfin, le statut de l'enseignant (contractuel *versus* fonctionnaire) présente des effets variables, selon les études relatives au Mali et au Niger, mais peu ou pas marqués. L'aspect qui ressort ici est à nouveau le rôle

⁵⁴ 102 écoles avaient été initialement sélectionnées.

⁵⁵ Les auteurs disposent également du score des élèves en fin de 4^{ème} année en langues et mathématiques.

relativement modéré que jouent les caractéristiques observables de l'enseignant bien que le rôle de celui-ci semble tout à fait majeur.

2.3 Les caractéristiques des enseignants au cœur des enjeux de la scolarisation primaire universelle

Il apparaît clairement que l'un des enjeux majeurs pour l'atteinte de l'objectif de scolarisation primaire universelle dans les pays d'Afrique francophone est le recrutement d'enseignants en nombre suffisant. Or, pour y parvenir, la plupart des pays devront accroître le rythme de recrutement de façon très importante. Ainsi, pour les pays de la CEMAC, il a été estimé que le taux d'accroissement annuel moyen devrait quadrupler et passer de 1,9 % pour la période 1987-2001 à 7,9 % pour la période 2001-2015. Pour la CEDEAO ce taux devrait passer de 2,5 % à 4,4 %. Même s'il existe une grande variété selon les pays, pour beaucoup le défi est immense d'autant que les ressources dont ils disposent sont limitées voire insuffisantes. Ainsi, les estimations réalisées par Mingat (2004) montrent pour 12 pays francophones que, même en diminuant sensiblement le coût salarial moyen, une part conséquente (environ 30 %) des dépenses courantes d'éducation devrait être financée par l'aide extérieure. Dès lors, les marges de manœuvre interne et externe des pays se trouvent extrêmement réduites. C'est dans ce contexte qu'ont émergé au début des années 2000 de nouvelles politiques de recrutement d'enseignants contractuels. A vrai dire, l'insuffisance des enseignants avait depuis longtemps fait émerger une nouvelle catégorie d'enseignants directement recrutés par les parents et nettement moins rémunérés⁵⁶.

Ces politiques ont eu des résultats spectaculaires en matière de scolarisation, des pays comme le Mali et le Niger ont pu scolariser plusieurs centaines de milliers d'enfants en plus, tandis qu'elles soulevaient d'intenses polémiques quant à la qualité de l'enseignement dispensé. Ces nouveaux enseignants diffèrent par le statut, la rémunération, la formation et l'ancienneté de leurs collègues fonctionnaires. Autant de caractéristiques jugées essentielles pour la qualité de l'enseignement par les acteurs et les partenaires de l'éducation. Cependant, au regard des résultats de recherche, au Nord comme au Sud, cette évidence de la relation entre les caractéristiques des enseignants et les apprentissages des élèves semble plus nuancée et le

⁵⁶ « un système dual qui taisait son nom prédominait », Alternatives internationales, mars 2007.

débat scientifique n'est pas clos. Même si quelques chercheurs trouvent les résultats peu conclusifs (Leclercq, 2005), l'angle de lecture adopté dans ce chapitre a permis de mettre en évidence certains constats. Ainsi, le niveau académique, la formation professionnelle voire l'ancienneté sont loin d'apparaître comme des déterminants majeurs de la qualité de l'éducation dans les travaux existants. On observe en fait une relative convergence de ces travaux quant au fait que les caractéristiques observables de l'enseignant n'expliquent qu'une faible partie de l'effet total de l'enseignant sur les acquis des élèves. Ce qui laisse entrevoir que les enjeux pour les systèmes éducatifs se situent plutôt au niveau du recrutement des « bons » enseignants et des incitations à fournir un travail de qualité. En ce qui concerne plus spécifiquement le statut, les travaux réalisés par le PASEC dans les pays africains, et approfondis par d'autres chercheurs, montrent à peu près la même tendance, c'est-à-dire des effets modérés, pas nécessairement en défaveur des enseignants contractuels, voire inexistantes. Ainsi, il n'apparaît pas avéré sur le plan empirique que ces politiques de recrutement d'enseignants contractuels aient eu une incidence marquée sur la qualité de l'enseignement primaire dans les pays concernés dans un sens ou dans un autre.

Toutefois, les données dont nous disposons sont encore peu nombreuses et relativement récentes par rapport à la mise en œuvre de ces politiques. Elles ne permettent donc pas d'anticiper de façon satisfaisante les effets sur le long terme de ces politiques. Enfin, elles n'ont pas encore été suffisamment exploitées par les chercheurs pour pouvoir affiner les conclusions actuelles. C'est d'ailleurs dans cette perspective que le chapitre suivant sera consacré à une analyse empirique sur le cas du Mali à partir des données du PASEC.

Chapitre 3 :
Statut des enseignants et acquisitions à l'école primaire au Mali :
faux problème et vraies questions

L'accélération du recrutement des contractuels est un fait marquant du début des années 2000 dans nombre de pays francophones, dont le Mali, comme cela a été souligné dans le chapitre précédent. Il est vrai que les enjeux sont de taille et que les gouvernements, face au défi de la scolarisation primaire universelle, ont dû mettre en place des mesures qui ne font pas l'unanimité. Bien que, très rapidement, ces nouveaux enseignants aient contribué de façon déterminante à l'extension de la scolarisation, les critiques n'en sont pas moins restées acerbes, y compris du côté des responsables politiques qui ont parfois vécu cette situation comme imposée de l'extérieur.

Plus généralement, si le débat est vif sur la question des enseignants contractuels, c'est en partie parce que cette mesure rompt avec la vision traditionnelle de l'éducation dans les pays francophones fortement attachés au statut et au prestige de leurs enseignants. Deux catégories d'arguments peuvent être avancées pour contester ou renforcer la logique de la politique des contractuels. D'un côté, ses détracteurs diront qu'avec une moindre formation professionnelle qui ne leur donnerait pas les moyens d'un enseignement de qualité, et un salaire moins élevé qui ne les encouragerait pas à fournir les efforts nécessaires, il ne faut pas s'attendre à de bons résultats, pire, il faut anticiper une dégradation de la qualité de l'enseignement. D'un autre côté, le faible impact de la formation initiale des maîtres dans les études menées ne semble pas de nature à discréditer d'emblée ces enseignants, sans oublier que dans certains pays comme la Guinée, ils bénéficient d'une formation à part entière⁵⁷. De plus, ces nouveaux statuts pourraient présenter un double avantage pour les systèmes éducatifs. Tout d'abord, ils peuvent agir comme une incitation forte pour les enseignants, si des incitations financières sont prévues, mais également si les manquements graves (comme l'absentéisme régulier et non justifié ou encore les comportements prohibés comme la violence ou le harcèlement sexuel) étaient sanctionnés par une rupture ou un non-renouvellement de contrat. Ensuite, le

⁵⁷ D'ailleurs, la situation a évolué rapidement et désormais, dans la plupart des pays, dont le Mali, les enseignants contractuels suivent tous une formation professionnelle initiale.

renouvellement des contrats offre une opportunité à l'individu de changer de voie s'il considère qu'il n'est pas fait pour ce métier⁵⁸. Enfin, l'administration scolaire pourrait sélectionner les meilleurs enseignants en ne renouvelant pas ceux qui présentent des lacunes trop importantes dans l'exercice du métier. On voit que les arguments sont nombreux et contradictoires.

La question en suspens est de savoir quel est l'impact de cette politique de contractualisation de la fonction enseignante sur la qualité des apprentissages. Il est alors préférable de s'appuyer sur des travaux de recherche pour s'éloigner des visions partisans, si prégnantes sur ces questions, et nourrir le débat avec de nouveaux éléments. Ainsi, les données d'une enquête du PASEC portant spécifiquement sur les enseignants contractuels au Mali sont mobilisées dans ce chapitre. Après la présentation du protocole d'enquête et des données utilisées (3.1), les résultats des traitements économétriques sont discutés (3.2). Les résultats obtenus invitent à s'interroger sur les principaux leviers d'amélioration de la qualité des apprentissages au Mali (3.3).

3.1 Protocole d'enquête et données disponibles

Dans le cadre du programme PASEC, le ministère de l'Education du Mali a demandé en 2000 à évaluer l'impact de sa politique de recrutement d'enseignants contractuels, nouvellement mise en place, sur les acquis des élèves à l'école primaire. Le PASEC dispose de sa propre méthode pour l'analyse des facteurs qui influencent les apprentissages des élèves. Celle-ci, basée sur le modèle à valeur ajoutée, repose, d'une part, sur la mesure des acquis scolaires en début et en fin d'année scolaire, à travers des tests de français et de mathématiques et, d'autre part, sur la collecte d'informations sur les élèves, les enseignants, les classes et l'école⁵⁹. Les classes de 2^{ème} et 5^{ème} années sont ciblées pour rendre compte des différences entre le début et la fin du cycle primaire⁶⁰. La collecte des données se fait classiquement sur la base d'un échantillon représentatif d'élèves construit selon les mêmes critères que ceux qui sont utilisés

⁵⁸ Alors que renoncer à un emploi de fonctionnaire à vie dans des contextes économiques aussi difficiles que ceux des pays africains n'apparaît certainement pas comme un choix envisageable très facilement.

⁵⁹ Cf. PASEC (2000).

⁶⁰ La première année ne permet pas encore d'assurer une évaluation des compétences, le résultat n'en serait que la description de la condition sociale et familiale. La dernière année, 6^{ème}, est à éviter du fait des enjeux et biais que pourraient introduire dans l'évaluation la préparation des examens et le processus de sélection vers l'enseignement secondaire.

par l'ensemble des programmes internationaux d'évaluation des acquis. Toutefois, dans le cas d'espèce, ce sont les enseignants et plus précisément leur incidence sur les apprentissages qui étaient l'objet d'étude. Il fallait donc comparer les progrès des élèves scolarisés avec les enseignants contractuels avec ceux des élèves ayant des maîtres titulaires, *ceteris paribus*.

3.1.1 Le protocole d'enquête

On peut d'emblée souligner que la comparaison *ceteris paribus* relève ici de l'incantation. En effet, il n'est pas permis de faire l'hypothèse que les enseignants sont affectés dans les classes de façon aléatoire. D'une part, nous savons que les enseignants contractuels sont en fonction depuis peu de temps au moment de l'étude et ils répondent à une volonté d'extension de la scolarisation. D'autre part, en règle générale, on observe que les enseignants débutants sont envoyés dans les zones où ceux plus expérimentés ne souhaitent pas rester ou aller. Sans précaution, on risque donc de comparer en moyenne des enseignants débutants et contractuels qui enseignent dans des zones spécifiques avec des enseignants titulaires plus expérimentés qui exercent dans des environnements différents et supposés plus favorables. Il n'aurait donc pas été approprié de choisir un échantillon d'enseignants titulaires et contractuels sans tenir compte du risque important de biais contextuels lié à l'affectation des maîtres. Pour limiter ces biais, il a été retenu d'utiliser une procédure d'appariement. Tout d'abord, le choix des enseignants de l'échantillon a été effectué en fonction de la répartition géographique des enseignants contractuels. Pour cela, dans un premier temps, le PASEC a procédé à une stratification proportionnelle des enseignants contractuels par région et à une répartition par unité administrative la plus petite, c'est-à-dire au Mali, la cellule d'animation pédagogique ou CAP (Cf. annexe 3.1.1). Ensuite, sur la base de cette répartition géographique, le principe retenu était, dans un premier temps, de choisir aléatoirement un enseignant contractuel et son école et, dans un second temps, un enseignant titulaire dans une école voisine aussi proche que possible en termes de distance, ceci dans l'idée de couvrir des contextes socio-économiques identiques. Idéalement, il aurait fallu choisir aléatoirement un enseignant contractuel puis un enseignant titulaire (son binôme en quelque sorte) et ainsi de suite pour chaque CAP à partir d'une base de données avant l'envoi des enquêteurs sur le terrain. Malheureusement, ce type d'informations n'était pas disponible au niveau central. D'une part, il n'existait pas de fichier individuel des enseignants complet et, d'autre part, compte tenu de l'importance des mouvements du personnel enseignant entre les années scolaires, il aurait été

très difficile de localiser précisément les enseignants en début d'année scolaire. En effet, il est quasiment impossible en début d'année de connaître les classes dans lesquelles les enseignants sont affectés et souvent bien difficile de connaître avec certitude l'école. Autant d'éléments, certains évoqueront les réalités du terrain, qui ont fait que les enseignants ont dû être choisis par les enquêteurs une fois sur le terrain. Même s'il aurait été préférable de choisir au préalable les enseignants, cette méthode assurait de sélectionner des maîtres dans les deux statuts pour des contextes de demande d'éducation comparables. Ainsi, chaque enquêteur avait les profils des enseignants à enquêter (toujours par binôme), par exemple un maître contractuel enseignant en 2^{ème} année et son binôme titulaire dans une école proche, et il devait faire le choix (aléatoirement) sur la base des informations disponibles au niveau local. Outre le statut, les enseignants différaient par leur ancienneté, les contractuels débutaient dans le métier, et le profil de formation, une grande partie des contractuels avait une formation de courte durée voire pas de formation du tout pour certains. Afin d'éviter un possible biais lié à ces différences, les consignes aux enquêteurs ont essayé d'en tenir compte. D'une part, pour les contractuels, on a distingué ceux qui avaient une formation professionnelle longue, d'autre part, en ce qui concerne les titulaires, il a été demandé aux enquêteurs de privilégier ceux qui avaient moins de cinq ans d'ancienneté quand l'information était disponible et le choix possible. Il s'agissait toutefois d'une consigne de second degré qui passait après celle de proximité géographique de l'école de celle du contractuel. Enfin, il a été arrêté pour étudier d'éventuelles différences entre début et fin de cycle de prendre en compte à la fois des classes de 2^{ème} année et des classes de 5^{ème} année. Or, se posait alors un problème de coût. En effet, si l'on n'enquête dans les écoles qu'une seule classe (celle pour laquelle l'école a été choisie en fonction des caractéristiques de l'enseignant), il faudrait un grand nombre d'écoles pour avoir suffisamment d'enseignants de chaque catégorie. Si par exemple, on souhaite pouvoir comparer 50 enseignants de chaque catégorie (contractuel/titulaire) et ce, pour les deux classes, il faudrait se rendre dans 200 écoles. Or les coûts de déplacement sont les plus importants dans ce type d'enquête. Le budget ne le permettait pas ; en revanche, une autre solution était possible : enquêter les deux classes dans le même établissement. Le budget permettait de se rendre dans environ 140 écoles⁶¹. La moitié de ces écoles a été choisie sur la base des caractéristiques de l'enseignant de 2^{ème} année tandis que l'autre moitié fut sélectionnée d'après les caractéristiques du maître de 5^{ème} année. Concrètement, pour chaque

⁶¹ Après avoir été complété par des financements complémentaires de l'Agence Intergouvernementale de la Francophonie et la direction de l'évaluation et de la prospective du ministère de l'Éducation nationale français.

niveau, on choisissait 35 enseignants contractuels et 35 titulaires, les 70 autres enseignants dépendaient du choix opéré pour l'autre niveau. On pouvait donc espérer avoir pour chacun des deux niveaux environ 140 enseignants (un peu moins car on sait qu'il y a toujours des déperditions, mais aussi des écoles à cycle incomplet). Cette façon de procéder a été imposée par les contraintes budgétaires, mais elle permettait d'obtenir une taille d'échantillon supérieure à ce qu'on aurait eu en enquêtant une seule classe par école⁶². Par ailleurs, cela permettait de renforcer la comparabilité des contextes puisque des enseignants de statuts différents peuvent ainsi être enquêtés, dans une même école, à deux niveaux différents.

Le protocole retenu dans cette étude est loin d'être parfait et il est grandement tributaire des informations disponibles au niveau local. Le caractère non aléatoire de l'affectation des enseignants est le problème méthodologique le plus sérieux à considérer ici. La démarche suivie a essayé d'y répondre au mieux en fonction de l'information disponible. Cela dit, si on peut espérer avoir ainsi limité sérieusement les biais possibles, on ne peut garantir de les avoir éliminés. Une attention particulière doit être portée à cet aspect dans les analyses.

⁶² Sous réserve que les capacités organisationnelles aient permis de gérer un plus grand nombre d'écoles et donc d'enquêteurs.

3.1.2 Les données

Les tableaux 3.1.1 et 3.1.2 présentent, respectivement pour la 2^{ème} et la 5^{ème} année, les données qui ont finalement pu être exploitées pour l'analyse ainsi que certaines variables d'intérêt en fonction du statut de l'enseignant. Il faut signaler qu'un nombre limité d'observations qui présentaient des évolutions extrêmes entre le score de début d'année et celui de fin d'année a été écarté afin de ne pas perturber les analyses (Cf. annexe 3.1.2).

En ce qui concerne la 2^{ème} année (Cf. tableau 3.1.1), on observe que les enseignants contractuels y sont beaucoup plus nombreux que les enseignants fonctionnaires (79 contre 49). Compte tenu du protocole identique pour les deux catégories, cela tend à indiquer que les contractuels sont plus souvent affectés en classe de 2^{ème} année que les titulaires dans les écoles de notre échantillon. Cela n'est pas sans rappeler les pratiques répandues dans les pays francophones d'Afrique de l'Ouest visant à affecter les enseignants les moins expérimentés et/ou les moins formés dans les classes de début de cycle (Bernard, 1999b). On constate également que les déperditions d'élèves, entre le début et la fin de l'année, sont assez nombreuses : près de 12 % de l'échantillon, et légèrement plus marquées chez les élèves des enseignants contractuels (13,1 %) que chez ceux des titulaires (10 %). Par ailleurs, les résultats des élèves aux tests de début et de fin d'année montrent une tendance très nette en faveur des élèves des enseignants non-fonctionnaires. En effet, alors que ces derniers ont un résultat significativement inférieur à leurs camarades en début d'année (22,8 contre 26,1 sur 100), ils obtiennent un score moyen nettement supérieur en fin d'année⁶³ (41,8 contre 36,2 sur 100) soulignant une meilleure progression en cours d'année. Bien sûr, à ce stade, il s'agit simplement de poser un constat et non pas de chercher une causalité. En outre, les coefficients de corrélation intraclasse, qui indiquent le degré de ressemblance des élèves par rapport au score au sein d'une même classe, sont assez élevés (entre 0,4 à 0,5) mais proches pour les deux catégories d'enseignants. Il existe donc des différences assez marquées entre classes mais qui ne recoupent pas le statut de l'enseignant. D'ailleurs, dans l'ensemble, les caractéristiques des élèves sont très semblables quel que soit le type d'enseignant. Une exception notable réside toutefois dans la proportion d'enfants vivant avec leurs deux parents qui est significativement plus élevée chez les élèves des enseignants titulaires.

⁶³ Cf. annexe 3.1.3 pour les tests de différence de moyennes.

Tableau 3.1.1 : Caractéristiques des élèves, des enseignants et des écoles de l'échantillon de 2^{ème} année

	Contractuels	Fonctionnaires	Ensemble
Caractéristiques des élèves			
Nombre d'élèves présents en début et fin d'année	1006	656	1662
% d'élèves n'ayant pas passé le test en fin d'année	13,1 %	10 %	11,9 %
Score moyen global au pré-test (sur 100)	22,8	26,1	24,1
Score moyen global au post-test (sur 100)	41,8	36,2	39,6
Coefficient de corrélation intraclasse en début d'année (score global)	0,44	0,54	0,48
Coefficient de corrélation intraclasse en fin d'année (score global)	0,43	0,46	0,45
% de filles	46,9 %	47,6 %	47,2 %
Age moyen des élèves	8,2	8,3	8,2
% d'élèves vivant avec leurs deux parents	68,3 %	79,6 %	72,7 %
Moyenne de l'indice de richesse	3	3,1	3
% d'élèves appartenant au quintile le plus pauvre	18,8 %	17,7 %	18,4 %
% d'élèves appartenant au quintile le plus riche	20,7 %	21,6 %	21,1 %
% d'élèves ayant redoublé leur 1 ^{ère} année	19,3 %	16,6 %	18,2 %
% d'élèves ayant redoublé leur 2 ^{ème} année	16,4 %	15,2 %	15,9 %
% d'élèves possédant les manuels de français et de mathématiques	6,2 %	7,5 %	6,7 %
Caractéristiques des enseignants			
Nombre d'enseignants	79	49	128
% d'enseignants ayant le niveau 2 nd cycle secondaire	79,7 %	81,6 %	80,5 %
% d'enseignants ayant le Bac et plus	7,6 %	8,2 %	7,8 %
% d'enseignants ayant suivi une formation professionnelle d'un an ou plus	7,6 %	100 %	43 %
Nombre moyen d'années d'ancienneté	3,2	20,2	9,7
% d'enseignants ayant suivi plusieurs sessions de formation continue au cours des 5 dernières années	29,1 %	38,8 %	32,8 %
% d'enseignants qui choisiraient à nouveau le métier d'enseignant	67,1 %	53,1 %	61,7 %
% d'enseignants qui échangent régulièrement avec leurs collègues sur les problèmes professionnels	89,9 %	87,8 %	89,1 %
% d'enseignants ayant une classe à simple flux	51,9 %	49 %	50,8 %
% d'enseignants ayant une classe à double flux	21,5 %	26,5 %	23,4 %
Taille moyenne de la classe	67	72	69
Caractéristiques des écoles			
% d'enseignants en zone rurale	50,6 %	30,6 %	43 %
% de contractuels dans l'école	54,4 %	36,7 %	47,6 %
Index d'équipement de l'école	3,1	3,1	3,1

Pour ce qui est des enseignants de deuxième année, des différences importantes sont à considérer. C'est avant tout au niveau de l'ancienneté et de la formation professionnelle que les enseignants contractuels et fonctionnaires se différencient. Les premiers n'ont qu'une ancienneté moyenne d'un peu plus de trois ans quand leurs collègues affichent plus de 20 ans d'expérience dans le métier en moyenne. Ensuite, la proportion de contractuels ayant suivi une formation professionnelle d'un an ou plus n'atteint pas 8 % quand la totalité des fonctionnaires a suivi une formation de ce type. Ainsi, la formation professionnelle et l'ancienneté sont clairement liées au type d'enseignant. En revanche, en matière de niveau académique, il n'y a pas de différence sensible, la très large majorité des enseignants (80,5 %) possède un niveau second cycle secondaire, mais sans avoir obtenu le baccalauréat. L'écart observé pour la formation continue est probablement en grande partie imputable à la faible ancienneté des contractuels dans le système puisque la question portait sur les cinq années précédentes et, qu'en moyenne, ils n'excèdent guère trois années de présence dans le système. Il est intéressant de remarquer des réponses assez éloignées sur la question du choix du métier qui pourrait traduire des différences de motivation entre enseignants. Alors que plus de 67 % des contractuels affirment que, si c'était à refaire, ils choisiraient le même métier, ce n'est le cas que pour 53 % des titulaires. Ceci pourrait avoir des conséquences sur l'efficacité des deux catégories d'enseignants qu'il conviendra de prendre en compte dans les analyses ultérieures. Enfin, on observe aussi des écarts très nets quant aux caractéristiques des écoles où exercent les deux types de maître. En effet, les contractuels sont plus souvent affectés dans des écoles situées en zone rurale⁶⁴ (50,6 % contre 30,6 %) et où les enseignants contractuels tendent à être majoritaires (54,4 % contre 36,7 %). Ainsi, malgré le protocole d'enquête spécifique suivi, on constate que les deux catégories de maître n'exercent pas nécessairement leur métier dans les mêmes conditions au sein des écoles de l'échantillon. Cela montre que l'affectation des enseignants est effectivement dictée par certains critères pas nécessairement explicites, ce qui peut constituer une source de biais à ne pas négliger dans les analyses.

Pour ce qui est de la 5^{ème} année (cf. tableau 3.1.2), on constate d'emblée une différence notable avec la 2^{ème} année puisque les enseignants fonctionnaires sont cette fois majoritaires (76 contre 58). Bourdon, Fröelich et Michaelowa (2006) relèvent le même phénomène au Niger dans une étude du PASEC similaire. Cela tend à renforcer notre observation précédente sur l'affectation plus fréquente des contractuels en 2^{ème} année et donc probablement de façon

⁶⁴ D'après la déclaration du directeur d'école.

plus générale en début de cycle primaire. On note que les déperditions sont équivalentes à celles observées en 2^{ème} année, même si cette fois ce sont les élèves des maîtres titulaires qui sont un peu plus concernés (13,1 % contre 10 %). Les résultats aux tests présentent également une différence majeure avec la situation observée en 2^{ème} année. Les écarts sont moins marqués même s'ils restent significatifs en début et en fin d'année mais, surtout, ils sont cette fois en faveur des élèves des enseignants contractuels dès le début d'année. Contrairement à la 2^{ème} année, les enseignants contractuels se voient confier en 5^{ème} année des élèves d'un niveau légèrement supérieur à ceux de leurs collègues fonctionnaires. Les coefficients de corrélation intraclasse sont à nouveau assez élevés, particulièrement en fin d'année, et pas très éloignés pour les deux catégories, si bien qu'on ne peut pas conclure à une différence d'homogénéité des élèves selon le type d'enseignants. On relève également des situations contrastées pour les variables relatives au niveau de vie des familles des élèves⁶⁵. Les élèves des enseignants fonctionnaires appartiennent plus souvent au quintile au niveau de vie le plus élevé (24,6 % contre 16,1 %) et moins souvent au quintile le plus pauvre (16,1 % contre 29,4 %) que les élèves des maîtres contractuels. Aucune différence notable n'avait été observée en 2^{ème} année sur ce plan. Il n'est alors pas très surprenant de voir les élèves des enseignants titulaires posséder plus fréquemment les manuels de français et mathématiques que leurs camarades (20,7 % contre 15,9 %). La proportion d'élèves redoublant la 5^{ème} année est un autre point de divergence, elle est en effet nettement plus élevée dans les classes des fonctionnaires (32,9 % contre 24,4 %). Il faut donc relever des différences entre les élèves des deux catégories d'enseignants beaucoup plus nombreuses qu'en 2^{ème} année.

La tendance est quelque peu différente au niveau des caractéristiques des enseignants. Les écarts observés en 2^{ème} année sont encore présents, mais ils tendent à s'atténuer pour certaines caractéristiques. En effet, la proportion d'enseignants contractuels ayant suivi une formation professionnelle longue est nettement plus importante qu'en 2^{ème} année (20,7 % contre 7,6 %). Sur le plan de la formation académique, on remarque que les contractuels sont plus souvent bacheliers que leurs collègues titulaires (20,7 % contre 6,6 %) mais aussi que les contractuels enseignant en 2^{ème} année (7,6 %). Ils ont également un an d'ancienneté en plus que ces derniers (4,2 ans en moyenne) mais restent loin derrière leurs collègues fonctionnaires qui ont en moyenne 20 ans d'ancienneté. Enfin, l'écart en matière de formation continue s'est aussi

⁶⁵ Cf. annexe 3.1.4 pour la construction de ces variables.

réduit entre contractuels et fonctionnaires, il n'est plus statistiquement significatif au seuil de 10 % ce qui est à rapprocher de l'ancienneté plus grande qu'en 2^{ème} année des contractuels.

Tableau 3.1.2 : Caractéristiques des élèves, des enseignants et des écoles de l'échantillon de 5^{ème} année

	Contractuels	Fonctionnaires	Ensemble
Caractéristiques des élèves			
Nombre d'élèves présents en début et fin d'année	765	979	1744
% d'élèves n'ayant pas passé le test en fin d'année	10 %	13,1 %	11,8 %
Score moyen global au pré-test (sur 100)	32,2	30,7	30,8
Score moyen global au post-test (sur 100)	33,4	30,4	31,7
Coefficient de corrélation intraclasse en début d'année (score global)	0,51	0,45	0,48
Coefficient de corrélation intraclasse en fin d'année (score global)	0,68	0,59	0,64
% de filles	41,3	44 %	42,8 %
Age moyen des élèves	12,1	12,1	12,1
% d'élèves vivant avec leurs deux parents	74,8 %	71,6 %	73 %
Moyenne de l'indice de richesse	2,7	3,3	3
% d'élèves appartenant au quintile le plus pauvre	29,4 %	16,1 %	22 %
% d'élèves appartenant au quintile le plus riche	16,1 %	24,6 %	20,9 %
% d'élèves ayant redoublé leur 4 ^{ème} année	22 %	19 %	20,3 %
% d'élèves ayant redoublé leur 5 ^{ème} année	24,4 %	32,9 %	29,2 %
% d'élèves possédant les manuels de français et de mathématiques	15,9 %	20,7 %	18,6 %
Caractéristiques des enseignants			
Nombre d'enseignants	58	76	134
% d'enseignants ayant le niveau 2 nd cycle secondaire	67,2 %	84,2 %	76,9 %
% d'enseignants ayant le Bac et plus	20,7 %	6,6 %	12,7 %
% d'enseignants ayant suivi une formation professionnelle d'un an ou plus	20,7 %	100 %	65,7 %
Nombre moyen d'années d'ancienneté	4,2	20	13,2
% d'enseignants ayant suivi plusieurs sessions de formation continue au cours des 5 dernières années	32,8 %	36,8 %	35,1 %
% d'enseignants qui choisiraient à nouveau le métier d'enseignant	70,7 %	72,4 %	71,6 %
% d'enseignants qui échangent régulièrement avec leurs collègues sur les problèmes professionnels	89,7 %	84,2 %	86,6 %
% d'enseignants ayant une classe à simple flux	53,4 %	50 %	51,5 %
% d'enseignants ayant une classe à double flux	27,6 %	32,9 %	30,6 %
Taille moyenne de la classe	60	67	64
Caractéristiques des écoles			
% d'enseignants en zone rurale	48,3 %	35,5 %	41 %
% de contractuels dans l'école	57,7 %	39,3 %	47,3 %
Index d'équipement de l'école	3,2	3	3,1

Ce constat vaut également pour la motivation puisque les fonctionnaires se montrent tout aussi disposés à choisir à nouveau le même métier que leurs jeunes collègues contractuels (respectivement 70,7 % et 72,4 %). La différence est frappante avec la 2^{ème} année où les fonctionnaires n'étaient que 53,1 % à déclarer qu'ils choisiraient à nouveau le même métier. On constate ainsi que le profil des enseignants contractuels de 5^{ème} année est différent de celui de leurs collègues de 2^{ème} année ; ils apparaissent plus formés et un peu plus expérimentés. On note aussi pour les enseignants fonctionnaires une différence sensible par rapport à leurs collègues de 2^{ème} année sur le plan de la motivation. Il semble donc que, de façon générale, les enseignants de 5^{ème} année se distinguent de leurs collègues de 2^{ème} année. Il reste néanmoins des différences notables entre contractuels et fonctionnaires.

A l'instar de la 2^{ème} année, même si l'écart s'est un petit peu réduit, les enseignants contractuels sont nettement plus souvent en poste dans des écoles situées en zone rurale. Cela est sans doute à mettre en relation avec le fait que les enseignants fonctionnaires ont un peu plus souvent des classes à double flux. Enfin, les enseignants contractuels exercent le plus souvent dans des écoles où ils sont majoritaires, ce qui les distingue à nouveau très nettement de leurs collègues titulaires.

Les différentes statistiques présentées permettent de souligner les différences entre les deux catégories d'enseignants qui sont étudiées, mais aussi entre les contextes dans lesquels ils évoluent. Autant d'éléments qui devront être pris en compte pour les analyses ultérieures. Enfin, nous avons aussi pu constater des situations relativement différentes entre la 2^{ème} et la 5^{ème} année qui montrent qu'il est utile de considérer des classes de début et de fin de cycle dans les analyses.

3.2 Statut et acquisitions scolaires : quelles évidences ?

L'analyse descriptive comparée entre enseignants fonctionnaires et contractuels a mis en évidence un certain nombre de différences entre ces deux catégories, ceci malgré le protocole d'enquête retenu. L'objectif de la modélisation est de pouvoir comparer toutes choses égales par ailleurs ces deux catégories d'enseignants afin de déterminer si l'une ou l'autre s'avère plus efficace pour faire progresser les élèves au cours d'une année scolaire. Compte tenu des différences observées, il était nécessaire, dans un premier temps, d'introduire des variables de

contrôle dans les modèles estimés. Dans la suite des discussions du chapitre 1, nous avons estimé plusieurs types de modèle. Le premier est basé sur les MCO (M1 et M2) mais il prend en compte la possibilité de non-indépendance des résidus liée aux effets de groupement et permet l'estimation d'intervalles de confiance robustes (Cf. annexe 1.2.2). Le second type de modèle correspond à un modèle multiniveau classique (M3) tel que présenté au chapitre 1 et reprenant les mêmes variables de contrôle que le modèle précédent (M2). Enfin, la dernière catégorie de modèle (M4) fait intervenir des variables croisées entre le statut et la localisation de l'école dans le cadre d'un modèle multiniveau (Cf. annexe 3.2.1).

Bien que les éléments dont nous disposons ne laissent pas supposer la présence d'endogénéité dans le modèle⁶⁶, on ne peut complètement exclure que les enseignants contractuels exercent leur métier dans des contextes spécifiques dont certaines caractéristiques ne sont pas prises en compte correctement par les variables de contrôle des modèles estimés. On peut notamment faire l'hypothèse que les contractuels, avant tout affectés dans des écoles afin de répondre à une demande scolaire, se retrouvent *in fine* dans des établissements avec des dynamiques spécifiques. Pour contourner cette difficulté, une tentative d'instrumentation a été réalisée (Cf. annexe 3.2.3). La variable contractuelle a été instrumentée avec les variables d'ancienneté et de formation professionnelle, toutes deux très corrélées au statut. Les résultats confortent ceux des modèles classiques et tendent à confirmer l'absence d'endogénéité. Toutefois, cette instrumentation est discutable dans la mesure où l'ancienneté et la formation professionnelle sont elles-mêmes censées influencer les acquis scolaires⁶⁷. Il reste donc utile de discuter les biais qui pourraient intervenir dans l'estimation et notamment le sens dans lequel ils peuvent jouer.

La même structure d'analyse a été utilisée pour les deux niveaux étudiés, mais naturellement les variables utilisées diffèrent pour chacun d'entre eux.

⁶⁶ Le test réalisé en annexe 3.2.3 suggère que la variable contractuelle n'est pas endogène, toutefois la fragilité du test qui se fonde sur une estimation d'une variable dichotomique avec des MCO invite à estimer un modèle avec variables instrumentales pour confirmer les résultats.

⁶⁷ Cette question est discutée en annexe 3.2.3.

3.2.1 Les résultats en 2^{ème} année : Les enseignants contractuels à leur avantage

Un premier constat porte sur la faible part de variance expliquée qu'implique l'introduction des différentes variables de contrôle puisque le R^2 ne progresse que de 10 points passant de 29 % (M1) à 39 % (M2). Le R^2 du modèle avec le seul score initial est de 26 %, ce qui implique que l'introduction de la variable du statut de l'enseignant permet un gain d'environ 3 %. En outre, la modélisation multiniveau (Cf. annexe 3.2.1, M3) permet de montrer que la variabilité au niveau élève est légèrement plus importante que la variabilité au niveau de la classe.

Il ressort des différents modèles estimés (Cf. tableau 3.2.1) une stabilité tout à fait remarquable du coefficient attaché au statut de contractuel, celui-ci oscille entre 0,31 et 0,36 selon les modèles, ainsi que le risque de sa non-significativité toujours en deçà du seuil de 1 %. La modélisation confirme l'observation réalisée à partir des scores moyens des élèves, à savoir que les élèves scolarisés avec des enseignants contractuels tendent à progresser de façon plus marquée que leurs camarades. Si on se réfère au tableau 3.1.1 de la section précédente, on peut constater que le score moyen des élèves des maîtres contractuels en début d'année est significativement inférieur au score moyen des maîtres titulaires (Cf. annexe 3.1.3). Il semble donc que si un biais existe, l'écart n'est pas si important pour l'affirmer, il soit en défaveur des enseignants contractuels puisque ceux-ci héritent en début d'année d'élèves de niveau plus faible, ce qui reviendrait ici à une sous-estimation du coefficient. Ceci tend donc à conforter le résultat obtenu.

Les variables croisées (M4) amènent des informations complémentaires puisqu'elles font ressortir que c'est essentiellement en zone rurale que les enseignants contractuels se révèlent plus performants alors que les fonctionnaires tendent à avoir des résultats comparables aux contractuels en zone urbaine.

Tableau 3.2.1 : Modèles explicatifs des scores de fin de 2^{ème} année au Mali

	M1	M2	M3	M4
Score standardisé de français et mathématiques de début d'année	0.53*** (0.04)	0.62*** (0.03)	0.61*** (0.02)	0.61*** (0.02)
Enseignant contractuel	0.36*** (0.11)	0.31*** (0.10)	0.32*** (0.10)	/
L'élève est une fille		-0.07 (0.04)	-0.06** (0.03)	-0.06* (0.03)
L'élève est plus âgé que l'âge normal pour la classe		0.12** (0.05)	0.07* (0.04)	0.07* (0.04)
L'élève appartient au quintile le plus pauvre		-0.18*** (0.06)	-0.10** (0.04)	-0.10** (0.04)
L'élève a redoublé sa 1 ^{ère} année		-0.15** (0.06)	-0.10** (0.04)	-0.10** (0.04)
L'élève a redoublé sa 2 ^{ème} année		-0.21*** (0.06)	-0.18*** (0.05)	-0.18*** (0.05)
L'enseignant est titulaire du BAC voire d'un diplôme du supérieur		-0.17 (0.13)	-0.19 (0.18)	-0.16 (0.18)
Si c'était à refaire l'enseignant choisirait à nouveau le même métier		0.20** (0.10)	0.20** (0.10)	0.20** (0.10)
L'enseignant discute régulièrement avec ses collègues des difficultés rencontrées en classe		0.35*** (0.12)	0.36** (0.16)	0.36** (0.15)
La salle de classe est construite en dur		-0.22** (0.11)	-0.21* (0.11)	-0.21* (0.11)
La classe fonctionne en double flux		-0.37*** (0.12)	-0.39*** (0.12)	-0.41*** (0.12)
Coefficient de variation par classe (écart-type /score moyen)		0.63*** (0.21)	0.65*** (0.17)	0.70*** (0.18)
Ancienneté du directeur d'école dans la fonction de directeur (en années)		0.02*** (0.01)	0.02** (0.01)	0.02*** (0.01)
Association de parents d'élèves jugée active ou très active par le directeur		-0.17 (0.10)	-0.18* (0.10)	-0.17* (0.10)
Proportion d'enseignants contractuels dans l'école		-0.05 (0.24)	-0.11 (0.25)	/
L'école est en zone rurale		0.04 (0.11)	0.04 (0.11)	/
Enseignant contractuel travaillant en zone urbaine				0.18 (0.13)
Enseignant contractuel travaillant en zone rurale				0.30** (0.13)
Enseignant fonctionnaire travaillant en zone rurale				-0.20 (0.18)
Observations	1662	1662	1662	1662
R ²	0.29	0.39	/	/
Nombre d'écoles	128	128	128	128

Erreurs type robustes entre parenthèses

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

D'ailleurs, on observe un coefficient élevé de signe négatif (-0,20), quoique non significatif, associé aux enseignants fonctionnaires exerçant en milieu rural. La non-significativité est à rapprocher du faible nombre d'enseignants puisqu'ils ne sont que 15 dans cette situation.

Quand on met en référence, dans le modèle, les contractuels exerçant en zone urbaine et qu'on insère la variable des fonctionnaires en zone urbaine (Cf. annexe 3.2.1, modèle 5), on remarque que les fonctionnaires en zone rurale sont nettement moins performants que les enseignants contractuels, ce qui n'est pas vérifié pour leurs collègues des zones urbaines. Ainsi, assez paradoxalement, les enseignants contractuels s'avèrent plus performants en zone rurale tandis qu'à l'inverse leurs collègues fonctionnaires y apparaissent moins efficaces. On voit ainsi derrière la tendance générale indiquant une meilleure performance des enseignants contractuels des différences sensibles selon la localisation de l'école. Il semble clair que le statut ne constitue pas une explication satisfaisante de ce type de résultats. Outre le statut, les contractuels se différencient des fonctionnaires par une ancienneté nettement moins importante. Il est possible de s'interroger sur le lien direct ou indirect que pourrait entretenir cette caractéristique avec le résultat observé. Comme dans tout métier, l'expérience est précieuse dans le métier d'enseignant, mais on peut se demander si l'ancienneté, à partir d'un certain seuil, ne pourrait pas être associée à un découragement relatif notamment pour les enseignants évoluant dans les conditions les plus difficiles ou les moins gratifiantes. Notre hypothèse est donc que le résultat observé pourrait s'expliquer avant tout par des questions de motivation. D'un côté, de nouveaux enseignants présentant l'enthousiasme des débutants, de l'autre, des maîtres qui auraient subi l'usure d'un métier exigeant et exercé dans un contexte difficile. Il a bien été tenté de contrôler un effet éventuel de la motivation de l'enseignant dans les modèles en intégrant une variable qui prend en compte le souhait du maître de choisir à nouveau le même métier et dont l'effet se révèle positif et significatif. Cependant, comme le relève Michaelowa (2002), cette variable est plus un indicateur de la satisfaction professionnelle qu'une mesure de la motivation des enseignants. Il n'est donc pas certain que cette variable permette de contrôler correctement l'incidence de la motivation des enseignants. On ne peut ici que se limiter à des hypothèses ; toutefois, il est clair que l'effet de la variable de statut incorpore d'autres effets et que son interprétation doit être prudente et que, compte tenu de ce qui vient d'être évoqué, cette enquête présente une situation diachronique alors que les phénomènes de la motivation et de l'expérience devraient être étudiés longitudinalement sur le moyen terme. Il n'en reste pas moins que l'extension de la scolarisation au moyen de cette catégorie d'enseignants s'est accompagnée à court terme d'une incidence positive sur la qualité des apprentissages en 2^{ème} année de l'enseignement primaire qui reste malgré tout globalement extrêmement modeste dans notre échantillon (PASEC, 2004b). Cependant, compte tenu de ce qu'il vient d'être dit, il serait tout à fait

hasardeux de considérer que ce résultat représente une tendance stable de l'impact des enseignants contractuels sur les acquisitions scolaires.

Les modèles estimés nous apportent d'autres informations sur différents facteurs intervenant dans le processus d'acquisition. Toutefois, en se référant au travail de Leamer (1983), seuls les résultats qui sont cohérents entre les différents modèles estimés sont considérés comme fiables. En effet, on ne peut accorder trop de confiance au résultat d'une variable dont la significativité et/ou le coefficient change selon les modèles estimés ici.

En ce qui concerne les caractéristiques des élèves, quatre variables ont un effet relativement stable selon les modèles. Tout d'abord, le fait d'être plus âgé que l'âge théorique pour la classe est associé à de meilleures progressions, même si l'incidence reste modérée autour de 10 points d'écart-type. Ce résultat est souvent mis en rapport avec l'avantage que confèrerait une plus grande maturité dans les premières années de scolarité. Il ne faudrait toutefois pas conclure trop rapidement que retarder l'âge d'entrée à l'école primaire serait un levier d'amélioration de la qualité de l'enseignement. En effet, ce constat pourrait ne pas être valide pour les classes de fin de cycle et, de plus, le coût d'opportunité d'envoyer un enfant à l'école augmentant avec l'âge de celui-ci, cela pourrait avoir des conséquences sur les abandons scolaires. En revanche, appartenir au quintile le plus pauvre se traduit par de moins bonnes progressions pour les élèves concernés. Là aussi, l'amplitude demeure assez modérée avec des coefficients négatifs compris entre 10 % et 18 % d'écart-type selon les modèles. Le redoublement apparaît également être lié négativement avec les progressions des élèves. L'effet des variables de redoublement apparaît un peu plus marqué que pour les variables précédentes surtout pour le redoublement de la 2^{ème} année où le coefficient oscille autour de 20 % d'écart-type. Ce résultat rejoint ceux trouvés par Bernard, Simon et Vianou (2005) dans d'autres pays d'Afrique francophone et s'inscrit dans la problématique complexe du redoublement dans les systèmes éducatifs francophones. Toutefois, les données disponibles ici ne sont pas de nature à permettre les analyses nécessaires à un traitement satisfaisant de la problématique du redoublement.

Pour les caractéristiques des enseignants, en dehors de leur statut, deux variables ressortent clairement avec des coefficients et des significativités très stables. Ainsi, le fait que l'enseignant déclare qu'il choisirait à nouveau ce métier si le choix se présentait à nouveau est associé à des progrès plus importants des élèves. Les coefficients sont stables entre 0,20 et

0,21 selon les modèles et significatifs au seuil de 5 %. Il a déjà été souligné que cette variable pouvait être considérée comme une proxy de la motivation. Cependant, elle mesure surtout la satisfaction de la position professionnelle plutôt que la motivation de l'enseignant à exercer au mieux son métier. Une possibilité est que la question ne revête pas exactement la même signification selon l'ancienneté et le contexte de travail de l'enseignant. La relation de cette variable avec les acquisitions des élèves demeure un aspect intéressant à considérer, de même que le fait que les contractuels sont plus souvent enclins à répondre qu'ils choisiraient le même métier que leurs collègues fonctionnaires. Elle tend à montrer que des aspects psychologiques sont à l'œuvre dans les résultats qui sont observés. La seconde variable influente a trait à la propension de l'enseignant à consulter régulièrement ses collègues pour discuter des difficultés rencontrées dans la classe. Cette variable, l'une des plus influentes, se traduit par de meilleures progressions (entre 35 et 37 % d'écart-type) qui peuvent traduire différentes choses. Outre, la détermination de l'enseignant à chercher des solutions à ses difficultés professionnelles, elle peut aussi refléter le niveau d'échanges entre enseignants au sein d'un établissement et ainsi prendre en compte l'incidence de la dynamique d'équipe et de l'ambiance de travail.

Si l'on excepte le score initial, les caractéristiques de la classe sont les variables les plus influentes des modèles qui ont été estimés. Le coefficient de variation, qui rapporte l'écart-type au score moyen de la classe au test de début d'année, apparaît le plus influent avec un coefficient qui varie entre 0,63 et 0,70 toujours significatif au seuil de 1 %. Ce résultat indique que les apprentissages en 2^{ème} année sont meilleurs quand il y a une relative hétérogénéité des niveaux des élèves⁶⁸. Mingat (1987) trouve des résultats similaires en 1^{ère} année de l'école primaire française. Le fonctionnement de la classe en double flux⁶⁹ est en revanche associé à de nettement moins bons niveaux d'acquisitions pour les élèves concernés. Ce résultat n'est pas surprenant dans la mesure où il a déjà été observé dans d'autres pays africains (Behaghel, Coustère et Lepla, 1999). En outre, même si les modalités de mise en œuvre peuvent fluctuer selon les pays, le double flux implique généralement pour des raisons pratiques d'organisation une diminution du temps d'enseignement effectif dont bénéficient les élèves. Il n'y a pas de débat autour de la question de l'efficacité de ce mode d'organisation scolaire, il s'agit d'un pis-aller pour faire face à l'insuffisance de locaux. Certains auteurs préconisent de n'avoir

⁶⁸ Des analyses complémentaires avec l'introduction de la variable au carré n'ont pas permis de mettre en évidence un effet de seuil.

⁶⁹ Le même enseignant travaille avec un groupe d'élèves le matin et un autre groupe l'après-midi.

recours à ce mode d'organisation qu'au-delà d'un effectif d'élèves très élevé par classe ; Michaelowa (2001a) établit le seuil à 82 élèves. Cela suppose cependant que la taille des salles de classes permette d'accueillir un si grand nombre d'élèves dans des conditions acceptables. Enfin, un dernier résultat plus ambigu touche le type de salle de classe. Les progressions des élèves sont moins bonnes dans les salles de classe en béton. Il est vrai que le coefficient est peu élevé et associé avec une significativité modérée mais stable à travers les modèles. Le résultat est contre-intuitif et les travaux antérieurs ont plutôt signalé l'absence de lien entre le type de construction et les acquisitions des élèves (Mingat et Suchaut, 2000). Toutefois, aucune explication satisfaisante n'a pu être mise en évidence par les différents traitements entrepris. Il est vraisemblable que ce résultat soit le produit d'effets de contexte qui nous échappent ici.

Enfin, la dernière catégorie de variables prises en compte caractérise l'école. Cependant, une seule des quatre variables retenues s'avère posséder un effet stable. Il s'agit de l'ancienneté du directeur dans la fonction qui est corrélée positivement avec les acquis des élèves. Le coefficient est assez modeste puisque pour 10 ans d'ancienneté, l'écart se situe entre 10 % et 20 % d'écart-type. Toutefois, cela tend à montrer que la gestion de l'établissement peut avoir une incidence sur les résultats des élèves et que le directeur joue un rôle dans le processus d'acquisitions. Il est aussi possible que l'ancienneté soit corrélée avec la possibilité de choix de l'école par les directeurs et que ceux-ci choisissent les établissements les plus dynamiques. L'effet ainsi mesuré serait alors plus un effet de l'établissement qu'un effet de l'ancienneté du directeur de l'école.

3.2.2 Les résultats en 5^{ème} année : un écart moins net en faveur des enseignants contractuels

Les analyses descriptives ont montré des différences très sensibles entre la 2^{ème} année et la 5^{ème} année de l'enseignement fondamental malien. Il s'agit maintenant de voir si ces différences ont des répercussions sur les résultats des modèles explicatifs des acquisitions scolaires des élèves de 5^{ème} année.

Pour commencer, le poids du score initial apparaît nettement plus marqué puisqu'il explique 35 % de la variance des scores de fin d'année (Cf. annexe 3.2.2, modèle 1) contre seulement 26 % en 2^{ème} année. En 5^{ème} année, le niveau initial semble donc plus déterminant dans la réussite scolaire. Toutefois, ce résultat pourrait aussi s'expliquer par une mesure moins précise des acquis en début de scolarité liée notamment à l'utilisation de tests écrits en début de 2^{ème} année. L'introduction de la variable CONTRACTUEL, quoique non significative, entraîne un gain de 0,7 % du R² soit quatre fois moins qu'en 2^{ème} année. L'ajout de l'ensemble des variables de contrôle fait progresser le R² de 0,36 à 0,44, ce qui est encore plus modeste que ce qui a été observé en 2^{ème} année puisque le gain était de 10 % de variance expliquée. Il s'agit là d'un constat important et commun aux deux niveaux : en dehors du score initial, les variables utilisées dans les modèles explicatifs n'expliquent qu'une faible part de la variance des scores de fin d'année. Par ailleurs, contrairement à la 2^{ème} année, on constate que la variabilité entre classes est nettement plus marquée que la variabilité individuelle (annexe 3.2.2, modèle 4). Ce sont donc plus les caractéristiques du contexte de la classe et de l'école que celles des élèves qui influent sur les acquisitions en cours d'année alors que paradoxalement les variables contextuelles expliquent une part modérée de la variance. On reviendra un peu plus loin sur ce point qui mérite attention.

On observe (tableau 3.2.2) à nouveau une stabilité assez remarquable du coefficient de la variable relative au statut de contractuel, même si l'on constate une ampleur plus modérée et une significativité moindre qu'en 2^{ème} année. Le coefficient n'est pas significatif au seuil de 10 % dans le modèle sans variables de contrôle. L'influence associée au statut de contractuel ne se dément pas, mais apparaît beaucoup plus modeste. En outre, cette fois, ce sont les enseignants contractuels exerçant en zone urbaine qui semblent les plus efficaces. Les faibles significativités sont cependant à rapprocher des petits effectifs de chacune des catégories,

respectivement 30 et 28 enseignants contractuels en zone rurale et en zone urbaine. De fait, les coefficients sont proches et l'écart n'est pas significatif entre les deux catégories d'enseignants contractuels.

Tableau 3.2.2 : Modèles explicatifs des scores de fin de 5^{ème} année au Mali

	M1	M2	M3	M4
Score standardisé de français et mathématiques de début d'année	0.58*** (0.04)	0.55*** (0.04)	0.45*** (0.02)	0.45*** (0.02)
Enseignant contractuel	0.17 (0.11)	0.22** (0.10)	0.20* (0.11)	
L'élève est une fille		-0.06 (0.04)	-0.10*** (0.02)	-0.10*** (0.02)
L'élève est plus âgé que l'âge normal pour la classe		-0.09 (0.06)	-0.06** (0.03)	-0.06** (0.03)
L'élève appartient au quintile le plus riche		0.00 (0.06)	0.04 (0.04)	0.04 (0.04)
L'élève a redoublé sa 4 ^{ème} année		-0.14** (0.07)	-0.04 (0.03)	-0.04 (0.03)
L'enseignant habite à proximité de l'école		0.34*** (0.09)	0.34** (0.16)	0.34** (0.16)
L'enseignant est titulaire du BAC voire d'un diplôme du supérieur		-0.22* (0.12)	-0.21 (0.17)	-0.21 (0.17)
La classe fonctionne en double flux		-0.33*** (0.10)	-0.37*** (0.12)	-0.37*** (0.12)
L'ensemble des enseignants et le directeur se réunissent une fois par mois		0.21** (0.10)	0.23** (0.11)	0.23** (0.11)
Ancienneté du directeur d'école dans la fonction de directeur (en années)		0.01** (0.01)	0.01** (0.01)	0.01** (0.01)
Association de parents d'élèves jugée active ou très active par le directeur		-0.18* (0.09)	-0.16 (0.10)	-0.16 (0.10)
L'école est en zone rurale		-0.10 (0.10)	-0.09 (0.11)	
Enseignant fonctionnaire travaillant en zone urbaine				0.09 (0.15)
Enseignant contractuel travaillant en zone urbaine				0.29* (0.17)
Enseignant contractuel travaillant en zone rurale				0.20 (0.17)
Observations	1730	1730	1730	1730
R ²	0.36	0.44	/	/
Nombre d'écoles	134	134	134	134

Erreurs type robustes entre parenthèses

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

Toutefois, on constate cette fois que le score moyen en début d'année des élèves des enseignants contractuels est légèrement, mais significativement supérieur, au score moyen des élèves des enseignants fonctionnaires (32,2 contre 30,7 sur 100). La faiblesse de l'écart ne semble cependant pas plaider pour l'existence d'un biais marqué en faveur des enseignants

contractuels, mais elle confirme que leur affectation en 5^{ème} année répond à d'autres critères que ceux de 2^{ème} année. Ceci est à rapprocher de la différence de profils observée entre ceux exerçant en 2^{ème} année et ceux affectés en 5^{ème} année, en moyenne plus expérimentés et plus diplômés.

A l'instar de la 2^{ème} année, la variable relative au statut contractuel regroupe de fait d'autres caractéristiques (salaire, ancienneté et formation) qui distinguent ces enseignants de leurs collègues fonctionnaires. Ces variables sont colinéaires avec le statut, si bien qu'il est impossible de connaître leur poids respectif dans le résultat observé. On peut cependant faire remarquer que les enseignants fonctionnaires étant en moyenne mieux formés, mieux payés et beaucoup plus expérimentés, il est pour le moins délicat d'attribuer leur moins bonne performance relative à leurs caractéristiques professionnelles. L'incidence du statut sur les acquisitions scolaires transite donc manifestement par d'autres vecteurs, dont la motivation pourrait être l'un des plus importants. En effet, l'enthousiasme lié à la jeunesse et aux débuts dans le métier associé à un découragement des plus expérimentés semble l'une des explications les plus plausibles du résultat observé. Il est cependant fort probable que la motivation se combine avec d'autres aspects comme par exemple un moindre absentéisme des contractuels du fait du type de contrat qui les régit et d'un contrôle plus contraignant.

Pour ce qui est des autres facteurs pris en compte dans les modèles, il faut signaler des différences par rapport à la 2^{ème} année. Il a déjà été souligné que leur pouvoir explicatif était moindre mais il faut ajouter que ce ne sont pas tous les mêmes. Il y a tout d'abord moins de variables dans les modèles car peu sont significatives en 5^{ème} année et de nouvelles variables font leur apparition.

Pour les caractéristiques de l'élève, on retrouve le symétrique de la 2^{ème} année pour ce qui concerne les variables utilisées. Toutefois, le fait d'être plus âgé que l'âge normal de la classe n'apparaît plus comme un avantage ; le coefficient est même négatif mais pas significatif dans tous les modèles. Cela montre bien qu'il ne faut pas conclure trop rapidement sur l'influence de l'âge comme aurait pu y inciter le résultat de 2^{ème} année. La variable relative au niveau de vie de l'élève (ici le fait d'appartenir au quintile le plus élevé) n'apparaît plus du tout significative. La mesure imprécise basée sur la déclaration des élèves peut être avancée, mais cela n'explique pas pourquoi la variable était influente en 2^{ème} année. Une hypothèse est que la variabilité des situations socio-économiques est moindre en 5^{ème} année du fait d'une

sélection liée à l'abandon scolaire. Dès lors, la mesure imprécise dont on dispose ne permettrait pas d'identifier correctement les différents niveaux de vie, ce qui expliquerait l'absence d'effet de cette variable. Une autre hypothèse plausible, qui n'exclut pas nécessairement la précédente, est que l'effet du niveau de vie serait plus marqué en début de cycle alors qu'il serait essentiellement capté par le score initial en 5^{ème} année. L'effet du redoublement est également plus modéré et plus instable qu'en 2^{ème} année. Le redoublement de la 4^{ème} année est associé à un coefficient négatif dans les différents modèles mais celui-ci n'est pas toujours significatif.

Les caractéristiques des enseignants, quant à elles, se révèlent assez peu influentes en 5^{ème} année. En dehors du statut, seul le fait que l'enseignant habite à proximité de l'école (dans le village ou le quartier) s'avère avoir une incidence sur les apprentissages. Cette caractéristique est associée à de meilleures acquisitions avec un coefficient très stable (0,34), et toujours significatif au moins au seuil de 5 %. Si ce résultat semble logique, puisqu'il pourrait être justifié par des retards et des absences plus nombreuses pour ceux qui sont loin de l'école, il faut toutefois s'interroger sur le fait que cette variable n'est pas ressortie des analyses de 2^{ème} année. La prudence dans l'interprétation est donc de mise d'autant que cela ne correspond qu'à un petit nombre d'enseignants (16).

Au niveau de la classe, on retrouve l'incidence négative du fonctionnement en double flux. Les coefficients attachés à cette variable restent très élevés, entre -0,33 et -0,37, et sont toujours significatifs au seuil de 1 %. La répétition de ce résultat en 2^{ème} et 5^{ème} années montre que ce mode de fonctionnement, tel qu'il était pratiqué au moment de l'évaluation, constituait une problématique à ne pas négliger pour la politique éducative, d'autant qu'il était relativement répandu. Toutefois, le double flux visant à faire face au problème d'insuffisance de locaux, il est difficile d'envisager de le supprimer à court terme. Les résultats plaident plutôt pour la mise en œuvre d'une politique de constructions scolaires ciblée.

En ce qui concerne l'école, on retrouve à peu près le même résultat qu'en 2^{ème} année pour l'ancienneté du directeur. Ceci tend à indiquer la relation entre la gestion de l'école et les acquisitions scolaires même si cette variable n'en rend compte que de façon très imparfaite. Il faudrait recourir à d'autres variables pour être en mesure de prendre en compte le type de management et d'autres dimensions du fonctionnement de l'école. D'ailleurs, on observe que les réunions mensuelles entre le directeur et l'ensemble des enseignants sont associées à de

meilleures acquisitions scolaires. Cela conforte l'idée que le management de l'école joue un rôle important dans le processus d'apprentissage et qu'il faudrait tenter de mieux le prendre en compte dans ce type d'analyses.

3.2.3 Les enseignants contractuels et la qualité des apprentissages

Les différents traitements économétriques réalisés ont amené des résultats assez stables qui tendent à renforcer la confiance qu'on peut leur accorder. Les différentes analyses réalisées laissent penser que le risque d'endogénéité est limité. La question de l'endogénéité, même si elle n'est jamais totalement écartée, ne semble donc pas de nature à remettre en cause les résultats obtenus. En effet, si des biais avaient échappé aux analyses, ils auraient plutôt tendance à sous-estimer l'effet de la variable « contractuel ».

Deux grands enseignements peuvent être tirés des résultats présentés dans cette section. Tout d'abord, le statut d'enseignant contractuel est plutôt associé à de meilleures acquisitions scolaires. Le résultat est plus net en 2^{ème} année qu'en 5^{ème} année. Le constat est donc en complète contradiction avec les discours courants sur la dégradation de la qualité de l'éducation qui serait essentiellement imputable à ces nouveaux enseignants. Toutefois, il convient d'être prudent et de souligner que ces résultats ne permettent pas de tirer de conclusions sur le long terme tant les aspects de motivation semblent jouer un rôle important et sont peu contrôlés dans les analyses menées. De plus, on se situe dans un contexte de faible qualité des apprentissages où il serait plus approprié de dire que les élèves obtiennent de moins mauvais résultats avec les enseignants contractuels.

Par ailleurs, un autre résultat marquant est que les différentes variables introduites dans les modèles estimés, en dehors du score initial de l'élève, n'expliquent qu'une part très modeste de la variance (13 % en 2^{ème} année et 9 % en 5^{ème} année). Ainsi, les principales caractéristiques des enseignants, dont leur statut, n'expliquent qu'une part modeste de la variance des scores. Ce qui confirme le faible lien entre le statut des enseignants et la qualité des apprentissages. Cela semble donc indiquer qu'il faille chercher ailleurs les principaux leviers d'amélioration de la qualité des apprentissages dans le système éducatif malien.

3.3 Où se situent les marges de manœuvre pour l'amélioration de la qualité des apprentissages dans l'enseignement fondamental malien ?

Si l'objectif de l'étude réalisée au Mali par le PASEC était bien d'analyser l'impact des enseignants contractuels sur la qualité des acquisitions scolaires à l'école fondamentale, les principales conclusions obtenues sont quelque peu paradoxales. En effet, il ressort clairement que les caractéristiques contextuelles prises en compte dans l'étude, dont les caractéristiques des enseignants, n'expliquent finalement que très peu de la variance des scores des élèves. Cela signifie que leur rôle potentiel dans l'amélioration de la qualité des apprentissages, sans être négligeable, reste limité. Or le contexte malien est caractérisé par des résultats en matière d'acquisitions scolaires particulièrement bas. L'amélioration du niveau des acquisitions scolaires est donc une priorité pour les autorités et appelle à rechercher d'autres marges de manœuvre pour la politique éducative. Les analyses menées dans ce chapitre n'ayant pas permis de les identifier, l'objectif se limite ici à tenter d'identifier où se situent potentiellement ces marges de manœuvre.

Pourtant, si le niveau des acquisitions est globalement bas dans l'échantillon étudié, on observe tout de même une grande variabilité des résultats. Il est intéressant de noter que cette variabilité ne se situe pas qu'au niveau des élèves mais aussi des classes, comme le mettent en évidence les modèles hiérarchiques (Cf. annexe 3.3.1). Dans le tableau 3.3.1, deux modèles sont estimés pour chaque niveau. Le premier modèle est un modèle vide encore appelé modèle hiérarchique d'analyse de la variance, car il permet d'estimer les différentes composantes de la variance (Cf. annexe 3.3.1). Le second modèle est le modèle multiniveau qui a été estimé dans la section précédente pour chacune des deux classes considérées. On observe qu'en 2^{ème} année la variabilité est un peu plus importante au niveau des élèves (la variance s'élève à 0,54 contre 0,44 dans le modèle vide). Toutefois, la variabilité au niveau des classes représente 45 à 46 % de la variance totale, ce qui n'est pas négligeable. La part augmente très légèrement dans le modèle avec l'ensemble des variables explicatives. En 5^{ème} année, la variabilité au niveau des classes est cette fois plus élevée puisqu'elle s'étend de 65 % à 62 % de la variance totale. Il semble donc qu'avec la progression dans les niveaux scolaires, la variabilité des résultats soit de plus en plus imputable à la variabilité entre classes et/ou établissements.

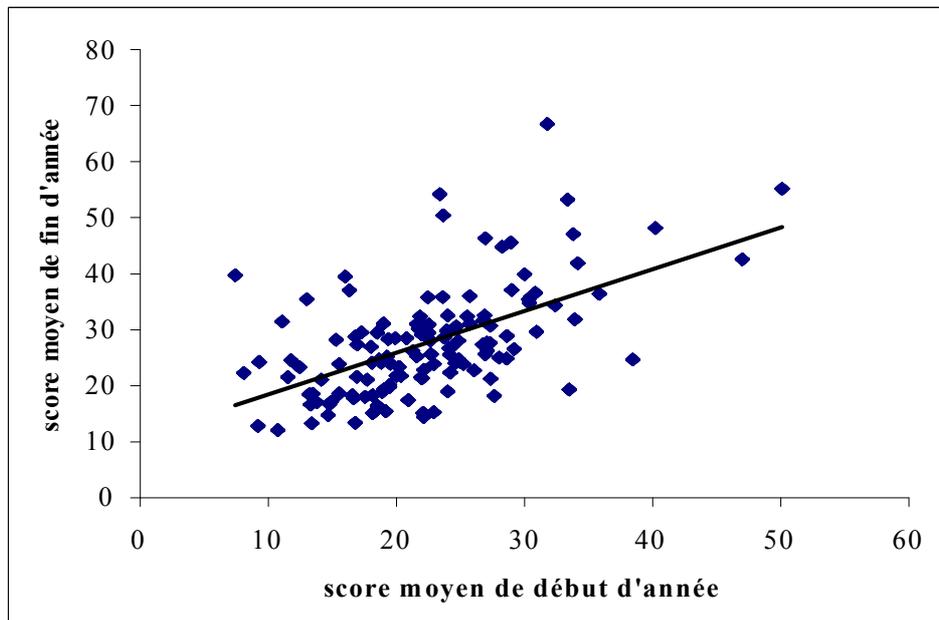
Tableau 3.3.1 : Les variances inter-classes et inter-élèves

	2ème année		5ème année	
	M1 (vide)	M2	M3 (vide)	M4
Variance inter-classes	0,44 (0,06)	0,29 (0,04)	0,64 (0,08)	0,37 (0,05)
Variance inter-élèves	0,54 (0,02)	0,34 (0,01)	0,34 (0,01)	0,23 (0,01)
Corrélation intra-classe	0,45	0,46	0,65	0,62

Erreurs types entre parenthèses

Le graphique 3.3.1 nous fournit une illustration de ces résultats sur la variabilité des résultats au niveau de la classe. S'il y a bien une relation positive entre le score moyen de la classe en début d'année et le score moyen en fin d'année, cette relation est loin d'épuiser toute la variabilité observée puisque le R^2 est seulement de 0,31. On observe ainsi des classes avec des scores moyens semblables en début d'année et qui affichent en fin d'année des résultats très éloignés, on est même tenté de dire trop éloignés. Cela montre que de grandes différences se créent entre les classes sur le plan des acquis scolaires entre le début et la fin d'une année scolaire, ce qui pose de sérieuses questions sur les inégalités entre classes et établissements au sein du système éducatif malien.

Graphique 3.3.1 : Relation entre score moyen initial et score moyen final au niveau des classes en 5^{ème} année de l'école fondamentale au Mali



Naturellement, il ne s'agit ici que d'une analyse bivariée et les différences observées peuvent s'expliquer par de multiples facteurs dont ceux pris en compte dans les analyses précédentes. Pour approfondir et voir quelle est la part de la variance totale expliquée par des différences inter-classes⁷⁰, des effets fixes « classe » ont été introduits en reprenant la spécification des modèles économétriques estimés dans la section précédente mais cette fois sous la forme de MCO classiques. En outre, les valeurs extrêmes qui avaient été écartées car potentiellement sources de biais ont ici été prises en compte car sinon, la variance de l'échantillon est réduite artificiellement.

⁷⁰ La notion de R^2 ou de variance expliquée est plus complexe dans les modèles multiniveau (Cf. Snijders et Bosker, 1999), aussi nous avons préféré ici utiliser les MCO classiques.

Tableau 3.3.2 : Effets fixes « classe » en 2^{ème} et 5^{ème} années au Mali

	2 ^{ème} année	5 ^{ème} année
R ² ajusté ⁷¹ avec le score initial	18,9 %	29,1 %
R ² ajusté avec l'ensemble des variables du modèle (M2)	31 %	38,3 %
Gain de R ² avec l'introduction des variables explicatives hors score initial	12,1 %	9,2 %
R ² ajusté avec les variables muettes « classes »	62,8 %	75,4 %
Gain de R ² avec l'introduction des variables muettes classe	31,8 %	37,1 %

On constate, en 2^{ème} année comme en 5^{ème} année, que la part de variance expliquée par l'introduction des variables muettes « classes » est très nettement supérieure à celle de l'ensemble des variables explicatives des modèles estimés dans la section précédente à l'exclusion du score initial (Cf. tableau 3.3.2). Ainsi, cette part est deux fois et demie plus importante en 2^{ème} année et quatre fois plus importante en 5^{ème} année. Concrètement, une partie conséquente des différences d'acquisitions qui se créent au cours d'une année donnée est imputable à la classe et probablement aussi à l'établissement⁷² où est scolarisé l'élève et ce indépendamment des multiples variables prises en compte dans l'analyse comme les caractéristiques des élèves, des enseignants, des classes et des écoles. Ainsi, les facteurs traditionnels comme la formation, le statut des enseignants, les caractéristiques de la classe, etc. laissent inexplicée une grande partie des différences qui se créent dans les acquis scolaires. C'est d'autant plus troublant que la plupart de ces facteurs ont des implications en matière de coût et que finalement leur rôle dans le domaine de la qualité des apprentissages apparaît beaucoup moins déterminant qu'on ne le suppose généralement. Le point le plus important qu'on peut entrevoir à travers ces résultats, c'est que des leviers majeurs d'amélioration de la qualité des apprentissages échappent aux analyses. En effet, les effets

⁷¹ Dans les modèles estimés dans la section précédente, seul le R² traditionnel est disponible du fait de l'estimation robuste des intervalles de confiance, ici l'utilisation des MCO classiques permet de disposer du R² ajusté ce qui est important du fait de l'introduction d'un grand nombre de variables dans le modèle.

⁷² Le PASEC n'évaluant qu'une seule classe par établissement, il est donc probable que l'effet de la classe recoupe en partie l'effet de l'établissement.

fixes nous montrent qu'il existe de grandes disparités entre classes, mais ils ne nous en donnent pas les causes⁷³. Or on peut penser que les différences entre classes et établissements, ou tout au moins une grande partie d'entre elles⁷⁴, peuvent être directement l'objet de mesures relevant de la politique éducative. Il est donc nécessaire d'essayer de mieux comprendre les causes de ces disparités entre classes et entre établissements, ce qui implique d'envisager de nouvelles pistes de recherche pour les travaux à venir en accordant une attention prioritaire à cet aspect. En outre, on doit aussi s'interroger sur l'aspect spécifique ou non du cas malien. Les autres pays africains sont-ils confrontés à des situations comparables ? Au-delà du continent africain, cette problématique est-elle prégnante ?

3.4 Statut des enseignants et qualité des apprentissages : une question essentiellement politique

La question enseignante est au cœur de la problématique de la scolarisation primaire universelle comme cela a été souligné dans le chapitre 2. La nécessité de recruter massivement de nouveaux enseignants à des coûts qui sont supportables par les Etats s'est peu à peu imposée aux différents acteurs des systèmes éducatifs. Il ne faut cependant pas sous-estimer la sensibilité et la complexité de cette question sur le plan politique. Le recrutement massif d'enseignants contractuels, s'il est entré dans les faits dans nombre de pays africains francophones, continue à provoquer des remous sociopolitiques qui peuvent fragiliser les systèmes éducatifs. Le Mali n'échappe pas à la règle et si l'étude du PASEC avait été commanditée par les autorités maliennes, c'était en partie parce que celles-ci étaient convaincues qu'elle allait montrer l'incidence négative de ces nouveaux enseignants sur la qualité des apprentissages et ainsi nuancer les conclusions « quantitativistes » qui leur étaient opposées par certains bailleurs de fonds. Cette question de la relation entre les « nouveaux enseignants » et la qualité des apprentissages est en effet un argument très souvent avancé dans le débat passionné sur les enseignants non fonctionnaires. L'étude réalisée au Mali permet de fournir des éléments factuels à ce débat. Les résultats obtenus prennent à contre-pied la plupart des argumentaires qu'on entend ici ou là sur la question. Les élèves scolarisés avec des enseignants contractuels ont tendance à mieux apprendre que leurs camarades

⁷³ Pour reprendre une formule d'Alain Mingat prononcée lors d'un colloque en 2005 à Dijon : « fixed effects are very nice words to say ignorance ».

⁷⁴ Il est également possible que cette mesure incorpore des effets de contexte.

scolarisés avec des enseignants fonctionnaires. Ce résultat est, il est vrai, plus marqué en 2^{ème} année qu'en 5^{ème} année. Cela dit, il faut resituer ce résultat dans un contexte de faible performance du système éducatif puisque globalement les élèves maliens apprennent peu, les scores moyens obtenus étant très nettement en dessous de la moyenne et relativement modestes en comparaison à d'autres pays africains (Cf. PASEC, 2004b). De plus, il faut aussi relativiser puisque ce résultat est le fruit d'une étude réalisée quelques années seulement après le début du recrutement massif des enseignants contractuels. Il ne permet donc pas de rendre compte d'éventuelles dynamiques. On peut notamment s'interroger sur ce qu'il adviendra quand ces nouveaux enseignants auront une ancienneté plus importante si, comme on en a fait l'hypothèse, celle-ci est liée avec la motivation de l'enseignant.

Toutefois, il ressort des analyses que le statut n'explique qu'une part modeste de la variance des scores et ne peut à lui seul apparaître comme un facteur majeur d'amélioration ou de dégradation de la qualité des apprentissages. Ce constat peut être étendu dans une moindre mesure à l'ensemble des variables explicatives (à l'exception du score de début d'année) qui ont été prises en compte dans les modèles estimés. Elles expliquent entre 8,2 % et 12,9 % de la variance totale. Cela laisse une partie considérable des différences d'acquisitions sans explication et tend à indiquer qu'il faille chercher ailleurs que dans les facteurs traditionnels de nouveaux leviers d'amélioration de la qualité des apprentissages. Des analyses complémentaires ont montré que 31,8 % de la variance des scores en 2^{ème} année et 37,1 % en 5^{ème} année étaient directement imputables aux classes et aux établissements fréquentés. Ces chiffres sont très nettement supérieurs à ceux trouvés pour l'ensemble des autres variables explicatives, si l'on excepte le score initial, ce qui laisse supposer qu'il existe des marges importantes d'amélioration de la qualité des apprentissages. On sait cependant relativement peu de choses sur les causes de ces disparités entre classes et établissements, si ce n'est qu'elles ne sont pas imputables aux différents facteurs utilisés dans les modèles. Il s'agit donc d'une voie à explorer qui soulève de nombreuses questions, à commencer par savoir si ce constat s'étend à d'autres pays que le Mali. Autant d'éléments qui seront abordés dans le chapitre suivant.

Chapitre 4 : l'amélioration de la qualité des apprentissages à l'école primaire en Afrique : les enjeux de la gestion des systèmes éducatifs

Les chapitres précédents ont amené à relativiser le rôle des caractéristiques observables des enseignants dans le processus d'apprentissage. La formation professionnelle, le niveau académique, l'ancienneté ou encore le statut sont apparus comme des facteurs moins déterminants que généralement supposés pour la qualité des apprentissages. Sans pour autant les négliger, il apparaît utile de rechercher de nouvelles pistes d'amélioration de la qualité des apprentissages dans les pays africains. Le chapitre précédent a mis en lumière l'importance des disparités entre classes dans la variance des scores de fin d'année à l'école fondamentale malienne. La part de variance expliquée par ces disparités interclasses est apparue beaucoup plus marquée que la part expliquée par l'ensemble des autres variables explicatives relevant directement de la politique éducative. Ainsi, les différences entre classes observées, qu'on ne sait pas expliquer, pourraient constituer l'une des principales sources d'amélioration de la qualité des apprentissages à l'école primaire ou fondamentale.

Il apparaît cependant opportun de voir dans quelle mesure le résultat obtenu au Mali peut s'étendre à d'autres pays, africains ou non (4.1). Des analyses complémentaires ont donc été menées sur d'autres pays africains francophones et anglophones sur la base des données du PASEC et du SACMEQ. Il s'agit dans un premier temps de voir la part de la variabilité du niveau classe dans la variabilité totale du score des élèves. Pour une meilleure comparabilité entre ces différentes enquêtes, c'est le score de mathématiques qui a été retenu. Le cas du Mali n'apparaissant pas du tout spécifique dans le contexte africain, le même type d'analyse est mené à partir des données TIMSS pour les pays développés, ainsi que pour deux pays africains, et à partir des données LLECE pour les pays d'Amérique latine. Partout les disparités entre classes ou écoles se révèlent relativement importantes mais, sur ce plan, on constate une différence substantielle entre pays développés et pays en développement. Pour approfondir ce résultat, on se concentre sur les pays du PASEC où on dispose d'un score initial, ce qui autorise, compte tenu de l'approche en valeur ajoutée, une analyse plus fine de l'effet classe et de son impact. Le résultat conforte les résultats antérieurs sur l'importance des disparités entre classes. Les variables de contrôle introduites amènent à discuter de la distinction entre « effet-classe » et « effet-maître ». Des analyses complémentaires sur le cas

maliens montrent qu'il est également important de prendre en compte des effets-établissements (4.2).

Sur le plan théorique, la fonction de production n'est pas d'une très grande aide pour rendre compte des effets-classe et effet-école observés. On retrouve certaines conclusions du chapitre 1 où l'accent était mis sur la nécessité d'expliquer les différences d'efficacité. C'est notamment l'apport de la « X-efficacité » de Leibenstein, adaptée à l'éducation par Levin (1997). Toutefois, cette théorisation de la boîte noire, si elle écarte la notion de fonction de production et le concept d'optimalité qui lui est attaché, nous dit peu de choses sur les rouages du processus éducatif. Elle laisse poindre la question du comportement des acteurs comme essentielle, mais sans fournir de clés d'analyse générales.

Ces constats amènent également à s'interroger sur la gestion de la qualité des apprentissages au sein des systèmes éducatifs (4.3). En effet, même si la recherche fournit encore peu d'éléments pour expliquer ces grandes disparités entre établissements, leur existence même pose la question de la myopie des systèmes éducatifs devant ces disparités. Un point essentiel est d'identifier des indicateurs permettant de fournir, à des fins de gestion, l'information nécessaire sur la qualité des apprentissages pour l'ensemble des établissements. Il faut donc une information qui soit exhaustive, c'est-à-dire portant sur l'ensemble des établissements scolaires, mais aussi produite de façon régulière par le système éducatif afin de permettre un véritable pilotage de la qualité des apprentissages. Des indicateurs sont proposés pour répondre aux besoins de la gestion. Toutefois, il apparaît clairement que leur utilisation est conditionnée à une évolution notable des pratiques de gestion des systèmes éducatifs.

4.1 Le poids des disparités entre écoles et entre classes dans le processus d'apprentissage

Les résultats relatifs au Mali ont mis en évidence l'importance des disparités entre classes dans l'explication de la variance des scores à l'école fondamentale. Il s'agit ici de voir si ce résultat peut s'étendre à d'autres pays africains, mais aussi à des pays d'autres régions du monde. Il n'existe pas d'enquête sur laquelle on pourrait se baser pour comparer tous les pays, il faut donc avoir recours à plusieurs enquêtes. Les données des enquêtes LLECE, PASEC, SACMEQ et TIMSS ont été mobilisées. Les mathématiques ont été retenues comme discipline sur laquelle se basent les comparaisons. Enfin, il faut signaler que les enquêtes ne ciblent pas toutes le même niveau éducatif ; ainsi pour LLECE et TIMSS c'est la 4^{ème} année de l'enseignement primaire qui est ciblée tandis qu'il s'agit de la 5^{ème} année pour le PASEC et la 6^{ème} année pour le SACMEQ. Ceci est à prendre en compte dans la mesure où on peut émettre l'hypothèse que les disparités entre établissements peuvent s'accroître en avançant dans le cycle. Dans un premier temps, il s'agit de connaître la part de variance expliquée par l'introduction de variables muettes pour chaque classe et/ou pour chaque école dans un modèle explicatif du score de mathématiques (Cf. annexe 4.1.1). Dans un second temps, à partir des données PASEC, on cherche à voir quelle est la proportion de cet effet classe ou école qui est expliquée par les variables explicatives usuelles.

4.1.1 La variabilité des résultats imputables au niveau classe ou école

L'analyse menée ici se cantonne à distinguer ce qui relève du niveau de la classe ou de l'école de ce qui relève du niveau individuel de l'élève dans l'explication de la variance des scores de mathématiques. L'idée sous-jacente est que, si toutes les classes et les écoles étaient identiquement efficaces, il n'y aurait pas de différences entre classes et entre écoles. Si des différences apparaissent et si elles sont importantes, cela suppose que des classes et des écoles sont beaucoup plus efficaces que d'autres, ce qui pose un double problème d'efficacité globale du système éducatif mais aussi d'équité. Toutefois, il faut souligner qu'en l'absence de variables de contrôle et d'une approche expérimentale, il est fort possible qu'une partie des écarts observés soient imputables à la composition du public d'élèves plutôt qu'à l'école elle-même. Cela n'enlève rien à l'existence d'inégalités entre classes et écoles mais il faut nuancer l'interprétation de leur origine.

Le tableau 4.1.1 montre pour les pays francophones du PASEC qu'en moyenne un peu plus de 45 % de la variance est imputable au niveau classe, ce qui est assez marqué. On remarque cependant des différences importantes selon les pays. Ainsi, le Mali, traité dans le chapitre précédent, présente la proportion la plus élevée (63,4 %) mais on constate que d'autres pays comme le Tchad et la Mauritanie dépassent les 50 %. Le cas du Mali n'est donc pas atypique. Le Sénégal présente quant à lui la part la plus faible avec seulement 26 %, ce qui est nettement en deçà de ce qu'on observe dans les autres pays.

Tableau 4.1.1 : Part de variance du score de mathématiques de 5^{ème} année du PASEC expliquée par les variables muettes « classe »

Pays	% de variance expliquée
Burkina Faso (1996)	33,3 %
Cameroun (1996)	47,1 %
Côte d'Ivoire (1996)	40,0 %
Madagascar (1998)	48,3 %
Mali (2002)	63,4 %
Niger (2002)	33,5 %
Sénégal (1996)	26,0 %
Tchad (2004)	55,2 %
Mauritanie (2004)	59,6 %
Moyenne	45,2 %

La même analyse a été faite pour les pays africains ayant participé aux enquêtes du SACMEQ. Cette fois, à l'exception de l'Ouganda, on dispose d'élèves appartenant à plusieurs classes d'une même école, ce qui permet de distinguer le niveau-classe et le niveau-école. Au niveau de la classe, la variabilité des résultats est un peu plus marquée : la part de variance expliquée varie de 15 % au Malawi à 64,8 % en Afrique du Sud. De plus, trois pays sont au-delà de 60 %, ce qui conforte l'appréciation donnée pour le Mali. Au niveau de l'école, la part de variance expliquée est systématiquement inférieure à celle du niveau-classe. Ceci est tout à fait normal puisque les différentes classes de l'école sont regroupées dans la mesure du niveau-école, ce qui tend à diminuer la variabilité. Plus l'écart est petit entre les deux mesures, niveau-classe et niveau-école, et plus les classes au sein d'une même école sont d'un niveau homogène. Seuls deux pays présentent des écarts très marqués, il s'agit de Maurice et

des Seychelles. Ces deux pays sont connus pour pratiquer des classes de niveau, c'est-à-dire que les classes d'une école sont composées en fonction du niveau scolaire des élèves. Les élèves faibles vont être regroupés ensemble dans une même classe, tandis que les élèves forts seront dans une autre classe. Hormis les cas cités, on observe en général un écart modéré entre la part de variance expliquée par le niveau-classe et celle expliquée par le niveau-école, ce qui laisse entrevoir l'importance des dynamiques propres aux établissements. Ce constat est intéressant dans une perspective de gestion des systèmes éducatifs.

Tableau 4.1.2 : Part de variance du score de mathématiques de 6^{ème} année du SACMEQ expliquée par les variables muettes « classes » et « écoles »

Pays	% de variance expliquée niveau-classe	% de variance expliquée niveau-école
Afrique du sud	64,8 %	63,7 %
Botswana	25,1 %	21,9 %
Kenya	38,1 %	35,6 %
Lesotho	30,9 %	29,9 %
Malawi	15,0 %	15,1 %
Maurice	37,4 %	24,8 %
Mozambique	24,9 %	20,4 %
Namibie	55,7 %	54,8 %
Seychelles	61,9 %	7,5 %
Swaziland	25,8 %	25,2 %
Tanzanie	26,4 %	24,9 %
Ouganda	62,4 %	/
Zambie	23,0 %	20,3 %
Zanzibar	34,4 %	33,0 %
moyenne	37,6 %	29,0 %

Les données du PASEC et du SACMEQ, bien que portant sur des niveaux et des pays différents, nous donnent une vision similaire : une part non négligeable de la variance des résultats scolaires est imputable à des différences entre classes et/ou établissements. Il apparaît intéressant à ce stade de vérifier s'il s'agit là d'une spécificité africaine ou si c'est une problématique qu'on retrouve dans d'autres pays du monde. Pour cela, les données du LLECE 1997, de l'enquête TIMSS 2003 pour la 4^{ème} année de l'enseignement primaire ont été mobilisées. En ce qui concerne les pays d'Amérique Latine et des Caraïbes, on observe des résultats relativement proches de ceux observés dans les pays africains. Seule la variabilité au niveau des écoles a pu être exploitée, et on constate que son niveau moyen est supérieur à

celui des pays SACMEQ. Il y a donc une problématique commune entre ces pays et les pays africains. Les fortes disparités entre établissements ne sont pas *a priori* une spécificité africaine. On constate également des différences entre pays puisque l'Argentine ou le Pérou présentent un chiffre supérieur à 46 % tandis que le Chili dépasse à peine les 24 %. Là encore, c'est un point commun avec les pays africains.

Tableau 4.1.3 : Part de variance du score de mathématiques de 4^{ème} année du LLECE 1997 expliquée par les variables muettes « écoles »

Pays	% de variance expliquée niveau-école
Argentine	46,4 %
Bolivie	39,2 %
Brésil	35,4 %
Chili	24,1 %
Colombie	28,4 %
Cuba	37,7 %
Honduras	39,1 %
Mexique	28,0 %
Paraguay	38,1 %
Pérou	46,4 %
République dominicaine	31,4 %
Vénézuela	26,6 %
Moyenne	35,1 %

Les données TIMSS présentent l'intérêt de concerner très majoritairement des pays développés et c'est précisément ces pays qui sont ciblés ici. Il s'agit de vérifier s'il existe des différences notables avec les pays en développement que nous venons d'évoquer. Un autre point intéressant est que deux pays africains, non subsahariens toutefois, le Maroc et la Tunisie, ont participé à l'enquête TIMSS 2003 pour la 4^{ème} année de l'école primaire, ce qui permet une comparaison directe. Si l'on regarde dans un premier temps les 11 pays développés considérés, on constate que les parts de variances expliquées au niveau de la classe et de l'école sont sensiblement inférieures à ce qui a été observé dans les pays africains et d'Amérique latine. En moyenne, cette part s'élève à 22,7 % au niveau de la classe quand elle varie entre 37,6 % et 45,2 % pour les pays africains. Si on se place au niveau de l'école, l'écart est moindre mais reste assez net. Ce qu'il est intéressant de relever, c'est la situation relative du Maroc et de la Tunisie. En effet, les deux pays africains, qui sont réputés avoir les enseignements primaires parmi les plus performants d'Afrique (UNESCO-BREDA, 2007),

présentent des résultats plus élevés que la plupart des autres pays et assez proches des pays africains pour le niveau classe, avec des chiffres de l'ordre de 35 %. Il faut toutefois souligner qu'il existe aussi une assez grande variété des résultats parmi les pays développés. Ainsi, le Japon, qui n'atteint pas 6 % (ce qui illustre une homogénéité remarquable entre classes et entre établissements), côtoie des pays comme les Etats-Unis, la Nouvelle Zélande ou l'Italie qui dépassent tous les 30 %. Il faut donc être relativement prudent dans les enseignements qui peuvent être tirés de ces comparaisons entre les différents pays.

Tableau 4.1.4 : Part de variance du score de mathématiques de 4^{ème} année de TIMSS 2003 expliquée par les variables muettes « classes » et « écoles »

	% de variance expliquée niveau-classe	% de variance expliquée niveau-école
Australie	26,4 %	25,1 %
Belgique	15,2 %	14,1 %
Angleterre	25,0 %	23,0 %
Hong Kong	27,8 %	24,5 %
Italie	34,7 %	29,8 %
Japon	5,8 %	/
Hollande	15,3 %	14,5 %
Nouvelle Zélande	36,0 %	31,2 %
Norvège	11,3 %	9,0 %
Etats-Unis	33,4 %	28,8 %
Ecosse	19,2 %	17,8 %
Moyenne	22,7 %	21,8 %
Maroc	34,6 %	/
Tunisie	35,2 %	/

L'idée générale selon laquelle la part de la variance des résultats en matière d'acquis scolaires imputable à la classe et à l'école est plus marquée dans les pays africains, mais aussi dans d'autres pays en développement, que dans les pays développés, ressort plutôt confortée de ces dernières analyses menées. Toutefois, un autre constat important tient à la grande variété des situations entre les pays qu'ils soient africains ou non. On peut bien sûr faire valoir un ensemble d'arguments pour expliquer cette situation. Tout d'abord, les modes d'organisation des systèmes éducatifs peuvent varier grandement entre pays, du plus centralisé au plus décentralisé. On pourra cependant observer avec malice que les systèmes éducatifs d'Afrique francophone, pour la plupart encore très centralisés, n'en présentent pas moins des variabilités

interclasses extrêmement élevées. On peut également souligner les limites de la mesure dont nous disposons, en effet les variabilités observées ne tiennent pas compte d'un ensemble de caractéristiques des élèves, du contexte scolaire et du contexte extrascolaire⁷⁵. En effet, rien de surprenant à voir une plus grande variabilité dans des pays pauvres où des écoles se situent dans des zones rurales isolées caractérisées par des populations très pauvres et majoritairement analphabètes. Ceci incite donc à affiner les analyses notamment pour essayer de différencier ce qui est attribuable à des facteurs scolaires et ce qui est attribuable à des facteurs extrascolaires.

4.1.2 Les disparités entre classes : « Effet-maître », « effet-classe » ou ?

L'un des points laissés en suspens par les analyses précédentes tient à l'origine de la variabilité interclasses : est-elle surtout imputable à des facteurs relevant directement de la politique éducative ou bien est-elle due avant tout à de fortes différenciations entre les publics scolaires ? Pour affiner l'appréciation, il convient de prendre en compte tout un ensemble de facteurs. Un aspect particulièrement important est de pouvoir contrôler le niveau initial des élèves, car il est possible qu'une part des disparités observées soit liée au public d'élèves plutôt qu'à l'établissement scolaire. Les données du PASEC présentent un intérêt réel pour ce type de question du fait de l'administration d'un test en début d'année aux élèves, et l'utilisation d'un même protocole dans un grand nombre de pays. Différents travaux ont déjà été réalisés sur cette question et sur les données du PASEC par Bernard (1999a) et Bernard, Tiyab et Vianou (2004). Dans le prolongement de ces derniers, de nouvelles analyses ont été réalisées pour 11 pays⁷⁶ ayant participé aux enquêtes du PASEC. L'objectif est tout d'abord de voir quelle est l'importance de la variabilité inter-classes une fois prises en compte les principales variables de contrôle relatives aux caractéristiques des élèves, des enseignants, des classes et des écoles. Pour ce faire, on observe l'évolution de la variance dans les modèles explicatifs des acquis scolaires, modèles à valeur ajoutée, avec l'introduction des variables muettes des classes (Cf. Annexe 4.1.1). Techniquement, on estime un effet-classe qui est

⁷⁵ Par ailleurs, l'hypothèse d'un biais mécanique lié à la différence des effectifs d'élèves testés selon les classes est réfutée en annexe 4.1.3.

⁷⁶ Il s'agit du Burkina Faso, du Cameroun, de la Côte d'Ivoire, de la Guinée, de Madagascar, du Mali, de la Mauritanie, du Niger, du Sénégal, du Tchad et du Togo.

souvent qualifié d'effet-maître dans la littérature (Bressoux, 2000). Nous reviendrons sur ce point plus loin.

Le tableau 4.1.5 présente les résultats obtenus dans les 11 pays, en 2^{ème} et en 5^{ème} années. La variabilité des résultats des modèles en fonction de la spécification ne permet pas de prétendre à une très grande précision dans ces mesures. Aussi faut-il considérer ces chiffres comme des ordres de grandeur⁷⁷. On observe des effets-classe relativement marqués, ils atteignent en moyenne 23,2 % et ils apparaissent un peu plus élevés en 5^{ème} année qu'en 2^{ème} année (24,5 % contre 21,8 %). Toutefois, Il n'existe pas de règle stricte : il y a des pays comme le Sénégal, la Côte d'Ivoire, le Niger et le Togo où l'effet-classe est plus important en 2^{ème} année. En outre, on observe que les écarts entre la 2^{ème} année et la 5^{ème} année sont en général peu marqués, ce qui signifie que la variabilité n'est pas très différente entre les deux niveaux, ce qui s'explique probablement en partie par le fait que les classes des deux niveaux se situent dans un même établissement. En effet le protocole d'enquête du PASEC prévoit que, dans un même établissement, on enquête la classe de 2^{ème} année et celle de 5^{ème} année quand elles existent⁷⁸.

⁷⁷ La structure des modèles est similaire dans les différents pays avec un corps de variables qui revient pour tous les pays. Toutefois, les spécifications sont ajustées selon les pays et il y a aussi des variables qui ne sont utilisées que dans certains pays.

⁷⁸ Dans certains pays, il existe des écoles incomplètes parfois assez nombreuses.

Tableau 4.1.5 : Part de variance expliquée par les variables muettes « classes » dans les modèles à valeur ajoutée sur les données du PASEC

Pays	2 ^{ème} année	5 ^{ème} année	Ensemble
Burkina Faso	13,4 %	15,0 %	14,2 %
Cameroun	19,6 %	24,9 %	22,3 %
Côte d'Ivoire	15,9 %	11,5 %	13,7 %
Guinée	25,3 %	40,7 %	33,0 %
Madagascar	27,7 %	36,7 %	32,2 %
Mali	31,8 %	37,1 %	34,5 %
Mauritanie	12,4 %	31,4 %	21,9 %
Niger	25,4 %	22,3 %	23,9 %
Sénégal	13,2 %	6,1 %	9,7 %
Tchad	29,8 %	29,7 %	29,8 %
Togo	16,6 %	11,4 %	14,0 %
Moyenne	21,8 %	24,5 %	23,2 %

On observe cependant deux exceptions avec la Guinée et la Mauritanie, avec dans les deux cas un effet-classe beaucoup plus marqué en 5^{ème} année qu'en 2^{ème} année. En Mauritanie, il existe un très grand nombre d'écoles à cycle incomplet (plus de 80 % des écoles), si bien que dans 32 écoles, seule la 2^{ème} année a été enquêtée et dans 13 écoles seule la 5^{ème} année a été enquêtée. L'écart observé pourrait venir en partie de ces écoles différentes. Une explication complémentaire tient au fait que ce ne sont pas les mêmes tests PASEC que dans les autres pays qui ont été utilisés en 2^{ème} année et, qui plus est, ces tests étaient administrés en langue arabe, la langue d'enseignement en 2^{ème} année. En revanche, les tests de 5^{ème} année sont bien les tests PASEC standards, le français devenant langue d'enseignement pour les mathématiques à partir de la 3^{ème} année. La différence de tests et l'aspect linguistique sont probablement à l'origine de l'écart observé. En Guinée, à nouveau, ce ne sont pas les tests PASEC standards qui ont été utilisés⁷⁹. Il est fort possible que les tests soient à l'origine du constat réalisé. On peut imaginer, par exemple, que les tests de 2^{ème} année laissaient moins de part à la variance des résultats de par leur degré de difficulté relatif que le test de 5^{ème} année. Dans tous les cas, le point commun à ces deux pays est la différence des tests utilisés, ce qui montre l'intérêt de disposer de tests comparables pour ce type d'analyses.

⁷⁹ En fait, certains items des tests standards ont été inclus mais il y avait des différences importantes.

On constate également une grande variabilité entre pays pour un même niveau. Ainsi, en 2^{ème} année, l'effet-classe oscille entre 12,4 % en Mauritanie et 31,8 % au Mali. En 5^{ème} année, il varie de 6,1 % au Sénégal à 40,7 % en Guinée. Il faut donc considérer des situations très différentes selon les pays, mais on doit tout de même relever que les chiffres moyens de l'effet-classe sont dans l'ensemble relativement importants. Ils traduisent ainsi de fortes différenciations entre classes qui correspondent à une réalité très concrète que rappellent Bernard, Nkengne Nkengne et Robert (2007) avec des classes aux niveaux moyens extrêmement différents au sein d'un même système éducatif (Cf. tableau 4.1.5). Ce qui est essentiel et découle des analyses réalisées, c'est qu'une très grande partie de ces différences entre classes n'est pas imputable aux facteurs traditionnels comme les caractéristiques observables des élèves, des enseignants, des classes ou des écoles.

Tableau 4.1.6 : Comparaison des scores moyens minimaux et maximaux de français aux tests PASEC de 5^{ème} année de l'école primaire

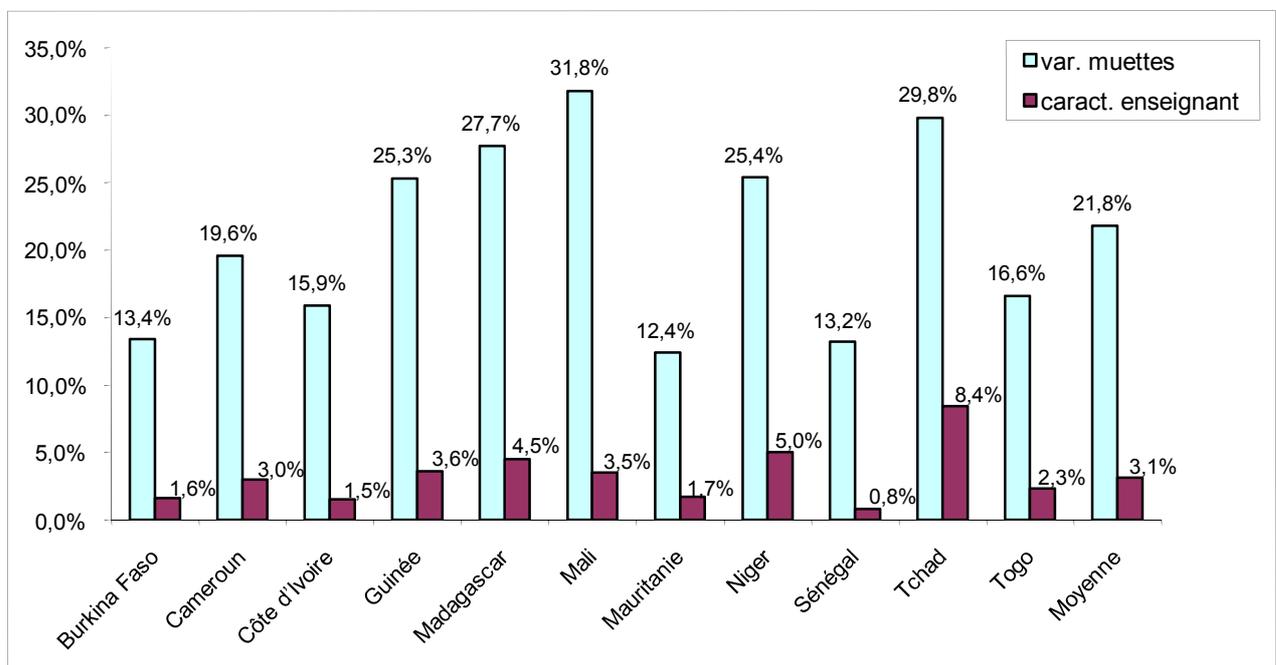
	Score moyen de la classe (sur 100)	
	minimum	maximum
Burkina Faso 1996	19,2	84,7
Cameroun 1996	31,4	87,3
Côte d'Ivoire 1996	31,6	80,2
Madagascar 1998	22,3	89,4
Mauritanie 2004	3,2	61,2
Sénégal 1996	17,5	76,3
Tchad 2004	11,1	79,0

Source : Bernard, Nkengne Nkengne, et Robert (2007)

Généralement, on impute ces différences entre classes non expliquées par les variables des modèles à l'enseignant ; la terminologie d'« effet-maître » ou « teacher effects » est employée dans la littérature (Bressoux, 2000 ; Konstantopoulos, 2007). Cela peut se justifier par le fait qu'il n'y a qu'un enseignant par classe à l'école primaire et donc, une fois contrôlées les caractéristiques des élèves, de la classe et de l'école, on peut avoir une mesure en creux de l'effet global de l'enseignant. Nous reviendrons plus loin sur les limites de cette approche. Il est intéressant de relever que des variables caractérisant le maître sont incluses dans les modèles, on peut donc comparer la part de variance expliquée par les variables muettes « classe », qu'on préfère qualifier d'effet-classe, avec la part de variance expliquée par les

variables caractérisant l'enseignant. Pour tous les pays, on retrouve, parmi ces variables, la formation académique et professionnelle de l'enseignant, son ancienneté et son statut. Le graphique 4.1.1 présente les résultats obtenus pour la 2^{ème} année : ils apparaissent très tranchés. La part de variance expliquée par les variables caractérisant l'enseignant est très modeste, elle oscille entre 0,8 % au Sénégal et 8,4 % au Tchad, et reste surtout très inférieure à l'effet-classe.

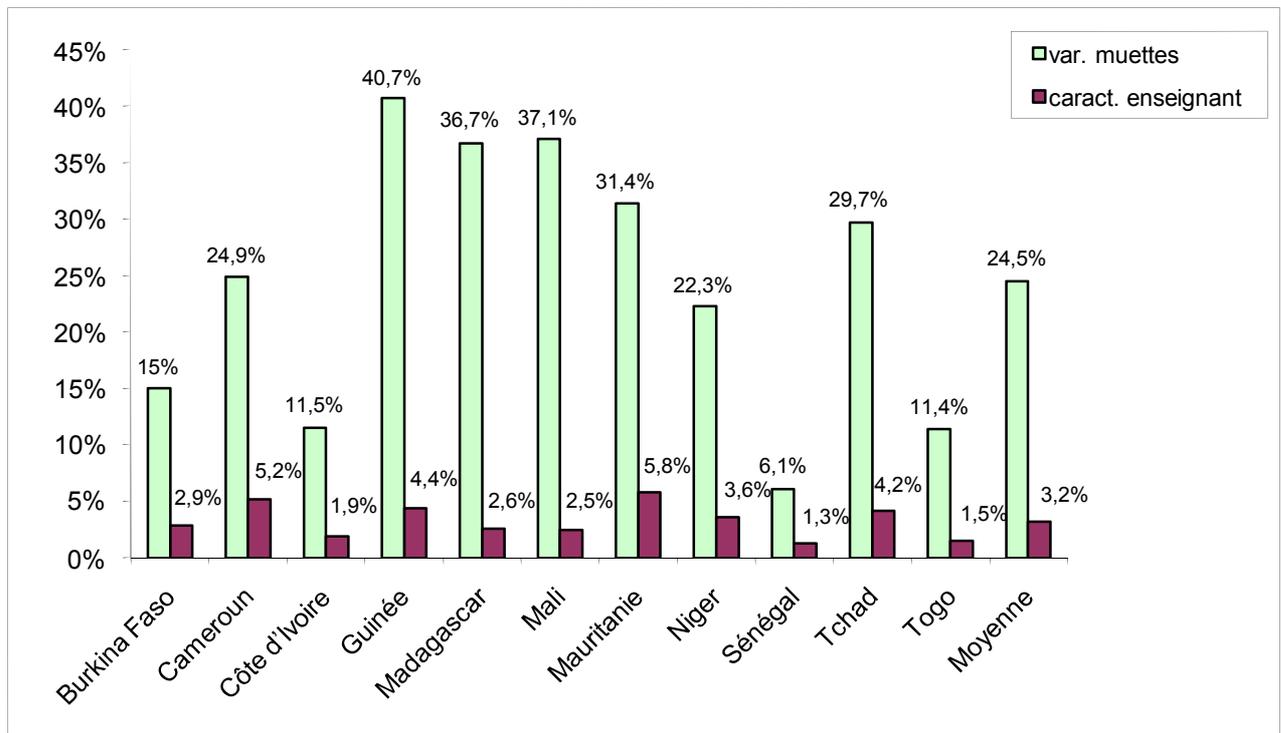
Graphique 4.1.1 : Comparaison des parts de variance expliquée par les variables caractérisant les enseignants et les variables muettes « classe » en 2^{ème} année



Le graphique 4.1.2 montre que la situation est similaire en 5^{ème} année avec des écarts même un peu plus marqués qu'en 2^{ème} année. En ce qui concerne la part de variance expliquée par les caractéristiques des enseignants, c'est le Mali qui, cette fois, affiche le chiffre le plus bas avec 1 %, devant de peu le Sénégal, tandis que la Mauritanie possède la part la plus importante avec 5,8 %. Pour ce qui est de l'effet-classe, c'est la Guinée et le Mali qui possèdent les chiffres les plus élevés avec respectivement 40,7 % et 40,2 %. Le Sénégal avec 6,1 % présente un résultat tout à fait atypique par rapport à l'ensemble des autres pays.

Ces résultats rejoignent les conclusions d'autres recherches, dont Rivkin et al. (2005) ou encore Murnane et al. (2005), évoquées dans le chapitre 2. Il s'agit d'un point particulièrement important sous l'angle de la politique éducative, non pas qu'il faille négliger les caractéristiques des enseignants mais plutôt qu'il convient d'envisager de rechercher de nouvelles pistes sur la base de ces résultats.

Graphique 4.1.2 : Comparaison des parts de variance expliquée par les variables caractérisant les enseignants et les variables muettes « classe » en 5^{ème} année



Le tableau 4.1.7 présente la part de variance expliquée par l'ensemble des variables explicatives, à l'exception du score initial, des modèles à valeur ajoutée qui ont été estimés. On constate de façon un peu surprenante que le pouvoir explicatif de l'ensemble des variables explicatives est aussi sensiblement inférieur à celui des variables muettes « classe ». On obtient en moyenne 7 % en 2^{ème} année, avec des variations assez prononcées entre 2,2 % au Togo et 18,6 % au Tchad. En 5^{ème} année, la moyenne est un peu plus élevée (8,4 %) mais avec aussi des variations importantes entre pays. Le Sénégal avec 3,9 % a le résultat le plus modeste tandis que la Mauritanie avec 18,1 % présente le chiffre le plus élevé.

Tableau 4.1.7 : Part de variance expliquée par l'ensemble des variables explicatives à l'exception du score initial dans les modèles à valeur ajoutée sur les données du PASEC

Pays	2^{ème} année	5^{ème} année	Ensemble
Burkina Faso	3,6 %	6,2 %	4,9 %
Cameroun	8,1 %	6,9 %	7,5 %
Côte d'Ivoire	5,1 %	5,1 %	5,1 %
Guinée	7,8 %	8,0 %	7,9 %
Madagascar	8,1 %	9,2 %	8,7 %
Mali	12,1 %	9,2 %	10,7 %
Mauritanie	6,4 %	18,1 %	12,3 %
Niger	9,5 %	4,4 %	7,0 %
Sénégal	2,8 %	3,9 %	3,4 %
Tchad	18,6 %	15,5 %	17,1 %
Togo	2,2 %	8,8 %	5,5 %
Moyenne	7,0 %	8,4 %	7,7 %

Enfin, ce dernier résultat pourrait être à double tranchant ou tout au moins amener une double interprétation. D'un côté, l'importance de l'effet-classe peut être vue comme l'incapacité des modèles à prendre en compte certaines variables importantes comme l'environnement familial de l'élève et notamment la pauvreté des familles. De l'autre côté, cela montre clairement que les variables traditionnelles laissent une grande part de la variance des résultats inexpliquée et que les différences entre classes permettent de mieux en rendre compte. Cependant, la question centrale est de savoir ce qui se cache derrière cette mesure en creux qu'est l'effet-classe. Plusieurs hypothèses peuvent être envisagées. Celle la plus communément admise dans les pays développés est qu'il s'agit d'un effet-maître (Bressoux, 2000), la mesure rendrait compte de caractéristiques inobservables de l'enseignant comme son charisme, sa motivation ou encore son talent pédagogique. Il est vrai que l'enseignant étant associé à la classe, il est tentant d'assimiler l'effet-classe à un effet-maître. Toutefois, plusieurs aspects sont à prendre en considération. Le premier est d'ordre conceptuel et est mis en avant par Bressoux (2000, p. 143) : « (...) il n'y a pas d'efficacité en soi des enseignants. (...) Cette efficacité n'est jamais que le produit d'une interaction entre un enseignant et des élèves ». Cet auteur nous rappelle que la production de connaissances n'est pas l'unique fait de l'enseignant, mais repose aussi sur les élèves. On est proche ici de la notion de

coproduction développée par Mc Meekin (2003) où l'enseignant et les élèves sont coproducteurs du produit éducatif⁸⁰. Bressoux (2000, p. 144) précise : « *Concevoir l'effet-maître comme le produit d'une interaction permet d'envisager que l'art de faire d'un enseignant ne rencontre pas toujours les conditions de son plein exercice* ». Cela revient à considérer que l'effet-maître n'est pas seulement attribuable à l'enseignant et n'est donc pas à proprement parler un effet du maître.

Au-delà, on peut aussi revenir sur cette assimilation entre la classe et l'enseignant et la questionner dans le contexte africain. Des exemples concrets nous permettent de montrer de nouvelles limites. Par exemple, il est assez fréquent que les enseignants soient informés tardivement de leur nomination et soient donc dans l'incapacité de rejoindre à temps pour la rentrée scolaire les écoles où ils sont affectés. Cela peut atteindre parfois plusieurs semaines pour les zones reculées et ce temps d'enseignement qui ne sera pas récupéré pénalisera nécessairement les acquisitions scolaires. Mais doit-on considérer cela comme un effet-maître ? Autre exemple, lui aussi très fréquent, les enseignants s'absentent parfois plusieurs jours par mois, dans certaines zones, pour aller percevoir leurs salaires car leur école est située dans une zone reculée et qu'ils doivent se rendre à la capitale régionale. Là encore, le temps d'enseignement perdu pénalise les élèves, mais est-ce un effet-maître ou un effet-administration ? Un autre exemple concerne les élèves et le fait qu'en zone rurale, lors de la période des récoltes, les élèves sont dans les champs plutôt que dans les classes. Là aussi le temps d'apprentissage perdu se ressentira sur les acquisitions scolaires et intégrera notre effet-classe. Pourtant, son origine est totalement en dehors de l'école et il serait plus approprié de parler alors de l'effet-récolte... Ce dernier point nous renvoie à la distinction entre ce qui relève de l'influence de l'environnement et ce qui dépend de l'école elle-même. Sur ce plan, il faut également aborder la question du contexte familial de l'élève et notamment de la mesure du niveau de vie. Nous avons déjà signalé, dans le chapitre 1, les problèmes que cela posait notamment du fait que l'information était collectée auprès de l'élève et donc relativement imprécise. Dès lors, l'effet-classe pourrait incorporer des effets liés à la composition du public d'élèves et en fait traduire des différences entre communautés. Il faut toutefois nuancer cette conclusion. Les modèles à valeur ajoutée comprennent le score initial de l'élève, qui traduit son passé scolaire et extrascolaire. Aussi peut-on penser que cette variable absorbe une partie

⁸⁰ Définition cohérente d'ailleurs avec celle utilisée par l'économie pour définir un service : une relation unique et non exactement reproductible entre une offre et une demande individuelle.

de l'effet lié à l'environnement de l'élève. On ne peut cependant pas complètement exclure que l'effet-classe incorpore une partie de l'influence de la communauté. De fait, on voit que cet effet-classe semble constitué de beaucoup de composantes différentes spécifiques au contexte africain. Ceci est probablement à rapprocher du fait que les mesures obtenues avec les données PASEC sont beaucoup plus élevées que ce qu'on observe dans les pays développés. Le tableau 4.1.8 présente les résultats d'études comparables menées aux Etats-Unis et en France sous le label « effet-maître » ou « teachers effect ». Il faut bien sûr être très prudent dans les comparaisons car il ne s'agit pas souvent d'échantillons représentatifs. Quatre de ces études concernent les pauvres ou des minorités (Armour, 1976 ; Hanushek, 1971 ; Hanushek, 1992 ; Murnane et Phillips, 1981). Deux seulement (Goldhaber et Brewer, 1997 ; Rowan, Correnti, & Miller, 2002) sont basées sur des données représentatives au niveau national. En outre, les variables de contrôle utilisées varient selon les études. Certaines ne contrôlent pas avec les caractéristiques de la classe et de l'école quand d'autres le font. Deux études utilisent les modèles multiniveaux pour leurs estimations (Rowan, Correnti, et Miller, 2002 ; Bressoux, 1996). Enfin, les études concernent des niveaux différents et utilisent des tests distincts. Il est donc préférable de prendre en considération uniquement les ordres de grandeur.

Tableau 4.1.8 : Les résultats des études sur les effets-classe aux Etats-Unis et en France

Etudes	Pays	Discipline testée	Niveau	Effet-classe (ΔR^2)
Armour (1976)*	Etats-Unis	Anglais (reading)	6	0,07-0,14
Goldhaber et Brewer (1997)*	Etats-Unis	Mathématiques	10	0,12
Hanushek (1971)	Etats-Unis	SAT ⁸¹	2-3	0,09-0,13
Hanushek (1992)*	Etats-Unis	Vocabulaire	2-6	0,16
Hanushek (1992)*	Etats-Unis	Lecture (reading)	2-6	0,1
Murname et Phillips (1981)*	Etats-Unis	Vocabulaire	3-6	0,1-0,21
Rivkin, Hanushek et Kain (2005)	Etats-Unis	Lecture (reading)	3-7	0,08
Rivkin, Hanushek et Kain (2005)	Etats-Unis	Mathématiques	3-7	0,14
Rowan, Correnti, et Miller (2002)*	Etats-Unis	Anglais (reading)	3-6	0,03-0,13
Rowan, Correnti, et Miller (2002)*	Etats-Unis	Mathématiques	3-6	0,06-0,13
Nye, Konstantopoulos, Hedge (2004)	Etats-Unis	Lecture (reading)	1-3	>0,07-0,07
Nye, Konstantopoulos, Hedge (2004)	Etats-Unis	Mathématiques	1-3	0,12-0,14
Mingat (1984)	France	Lecture	1	0,16
Mingat (1984)	France	Mathématiques	1	0,12
Mingat (1991)	France	Français et Mathématiques	1	0,14
Bressoux (1995)	France	Lecture	3-5	0,11-0,13
Bressoux (1996)	France	français	3	0,1-0,11
Bressoux (1996)	France	Mathématiques	3	0,14-0,19

* Cité par Nye, Konstantopoulos et Hedge (2004)

On constate qu'on se situe largement en deçà des chiffres africains, puisque les moyennes en 2^{ème} et 5^{ème} années sont au-delà de 20 %, alors qu'aucune étude citée dans ce tableau n'atteint ce chiffre. On se situe plutôt autour de 10 % avec des résultats un peu plus élevés dans le cas français.

Cela tend à confirmer que l'effet-classe ne se limite pas à un effet-maître, puisqu'*a priori* il n'y a pas d'explication à des écarts majeurs entre les effets-maîtres des différents pays⁸². On

⁸¹ Scholastic Aptitude Test, ce test d'admission dans les universités présente l'inconvénient d'être passé sur une base volontaire, il correspond donc à un sous ensemble sélectionné de la population.

⁸² Les biais de mesure propres à chaque étude peuvent causer mécaniquement des écarts limités comme un bruit de second ordre mais ils ne peuvent pas justifier des écarts importants.

montre, par la même occasion, la limite de cette mesure qui vaut pour tous les contextes. Cet effet-classe ne peut donc pas être *stricto sensu* un effet-maître, il peut même en être une mesure très médiocre dans certains contextes. En revanche, il s'agit d'une mesure qui nous indique les différences d'efficacité entre classes, et ces différences apparaissent très marquées dans un grand nombre de pays africains. La question reste posée sur la composition de cet effet-classe et, incontestablement, il faudra que les recherches futures y accordent une attention soutenue car il apparaît clairement que des marges de manœuvre majeures pour l'amélioration de la qualité des apprentissages restent à identifier. Une piste souvent évoquée (Bressoux, 2000 ; Bernard et al., 2004 ; UNESCO-BREDA, 2007) est le temps scolaire, mais peu d'études empiriques sont disponibles pour quantifier son impact sur les apprentissages et notamment dans le contexte africain. Duflo et Hanna (2005) ont pu montrer la relation positive entre le temps de présence de l'enseignant et les acquis scolaires dans une étude réalisée en Inde. Il reste cependant beaucoup de chemin à parcourir pour mieux appréhender cette variable dans les études empiriques. Il ne s'agira pas bien sûr de la seule piste à explorer. D'ailleurs, la question du temps scolaire est souvent liée à la gestion locale qui doit elle aussi être considérée comme un objet d'étude privilégié. Un point en suspens tient à la capacité d'observer ce qui relève de l'enseignant et de ce qui ne lui est pas imputable en matière de temps scolaire.

4.1.3 Les questions soulevées par l'effet-classe

Les différentes analyses réalisées dans cette section et les travaux mentionnés signalent tous l'importance de la variabilité entre classes, que ce soit dans les pays développés ou dans les pays pauvres. Toutefois, les amplitudes apparaissent particulièrement importantes dans ces derniers et notamment sur le continent africain. Un point marquant relevé dans de nombreuses études est que la part de variance expliquée par les caractéristiques de l'enseignant est très nettement inférieure à la part de variance expliquée par l'introduction des variables muettes identifiant les classes. Dans la littérature, ce type de mesure est généralement assimilé à un effet-maître, puisque le principal facteur qui change entre deux classes, c'est bien l'enseignant. Toutefois, cette assimilation est discutable (Bressoux, 2000) et notamment dans le contexte africain où l'ampleur des effets tend à confirmer que d'autres facteurs interviennent dans cette mesure. Il reste à identifier ces facteurs et surtout à quantifier leur impact dans les recherches futures, ce qui ne sera pas nécessairement très simple. Dans le

contexte africain, il apparaît important de mieux prendre en compte l'environnement familial de l'élève et notamment son niveau de vie ; mais aussi de parvenir à rendre compte de la grande variabilité du temps d'apprentissage, selon les classes et les écoles, qui pourrait être une clé déterminante de la qualité des apprentissages mais dont la mesure s'avère délicate.

Ces différences entre classes qui ne s'expliquent pas par les principaux inputs qui constituent le coût de l'éducation (comme la formation, l'ancienneté, le statut de l'enseignant, la taille de classe, le redoublement, etc.) illustrent des différences d'efficacité dans l'utilisation des inputs de la classe. On est d'ailleurs assez proche ici de la « x-*efficiency* » de Leibenstein (1966). Cependant, l'idée que ces différences soient imputables au seul enseignant semble par trop réductrice. La recherche doit encore amener des réponses sur ces aspects notamment dans une perspective de pilotage des systèmes éducatifs, ce que n'ont pas souvent permis les travaux sur l'efficacité des écoles qui sont très descriptifs, mais se transcrivent assez mal en mesures de politique éducative (Levin, 1997). Il n'en demeure pas moins que ces résultats interpellent la gestion des systèmes éducatifs. L'existence de fortes différenciations entre lieux de scolarisation n'est pas compatible avec l'idée d'un système cohérent et équitable propre à tout système d'éducation.

4.2 De l'effet-classe à l'effet-école

L'effet-classe traité précédemment traduit de fortes disparités entre classes, et vraisemblablement aussi entre écoles, qui ne s'expliquent pas par les facteurs traditionnels qui ont été pris en compte dans les analyses. Quelle que soit l'origine de ces disparités, elles posent un sérieux problème sur le plan de l'équité du système éducatif puisque, selon le lieu de scolarisation, on doit s'attendre à des résultats différents. De plus, des simulations montrent que cette situation pénalise aussi assez sensiblement le niveau moyen de performance d'un pays. Ainsi, dans le rapport Dakar +7 (UNESCO-BREDA, 2007), pour étudier les conséquences de ces disparités entre classes, différentes simulations ont été réalisées à partir des données PASEC⁸³. L'objectif était d'identifier des classes présentant des résultats très en deçà de la moyenne⁸⁴ et d'estimer l'impact de ces classes sur le niveau moyen

⁸³ Cette analyse a été menée par nos soins.

⁸⁴ Le critère retenu était un score moyen de la classe inférieur de 0,75 écart-type ou plus au score moyen de l'ensemble des élèves. Cela représente entre 10 % et 20 % des écoles de l'échantillon selon les pays.

des élèves dans un pays. Pour cela, des scores ont été estimés pour les élèves de ces classes en tenant compte de leurs caractéristiques personnelles (dont leur niveau en début d'année) mais en leur affectant un effet-classe moyen. En fait, il a été procédé à une simulation économétrique des scores que ces élèves auraient obtenus s'ils avaient été scolarisés dans une classe avec des résultats moyens plutôt que dans une classe sous-performante. On peut ainsi estimer l'impact de ces classes sous-performantes sur la performance globale du système éducatif. Pour cela, on regarde à la fois l'évolution du score moyen et la proportion d'élèves qui a obtenu au moins 40 % de bonnes réponses aux tests (cf. tableau 4.2.1).

Tableau 4.2.1 : Incidence des classes sous-performantes sur la qualité des apprentissages au sein des systèmes éducatifs de quatre pays francophones

		Score moyen de 5 ^{ème} année (sur 100)		Proportion d'élèves avec au moins 40 % de réponses correctes	
		réel	simulé	Réelle	simulée
Mauritanie	Français	20,9	23,3	10 %	10 %
	Mathématiques	22,9	26,1	16 %	17 %
Madagascar	Français	31,4	32,8	22 %	22 %
	Mathématiques	51,3	53,4	77 %	84 %
Cameroun	Français	45,1	48,8	60 %	72 %
	Mathématiques	46,4	50,9	64 %	77 %
Tchad	Français	28,8	32,8	22 %	24 %
	Mathématiques	32,6	37,7	31 %	46 %
Gain moyen	Français	2,9		3,8 %	
	Mathématiques	3,7		9,1 %	

Source : UNESCO-BREDA (2007)

Les simulations réalisées aboutissent à un gain moyen pour les scores aux tests sur les quatre pays de près de trois points en français et près de quatre en mathématiques. Pour ce qui est de la proportion d'élèves qui atteignent le seuil de 40 % de bonnes réponses aux tests, seuil jugé souhaitable pour l'ensemble des élèves, le gain est proche de 4 points de pourcentage en français et légèrement supérieur à 9 points de pourcentage en mathématiques. Il faut noter sans surprise qu'il existe un peu de variété selon les pays et selon l'indicateur retenu⁸⁵, mais,

⁸⁵ Ainsi, pour la Mauritanie, on observe un effet assez net sur le score moyen mais très peu d'effet sur la proportion d'élèves qui atteignent le seuil de 40 % de bonnes réponses, ce qui s'explique par la faiblesse générale du niveau des élèves.

dans l'ensemble, les ordres de grandeur obtenus montrent l'incidence de ces écoles déviantes sur la qualité des apprentissages au plan national.

On parle ici d'écoles déviantes et non de classes déviantes, dans la mesure où on peut penser que les acquis des élèves en fin de 5^{ème} année sont le produit de l'ensemble de la scolarité réalisée dans l'établissement. On fait l'hypothèse assez réaliste que la mobilité des élèves n'est pas suffisante pour remettre en cause le fait que les acquis en fin de scolarité soient imputables à l'établissement. Cependant, sur le plan empirique les estimations que nous avons pu réaliser, à partir des données, PASEC n'ont concerné que des effets-classe. En effet, le PASEC n'enquêtant qu'une classe par niveau testé, il n'est pas possible de distinguer un effet-classe d'un effet-école. Toutefois, l'existence d'un tel effet peut être testée de manière indirecte. Si un effet-école suffisamment important existe, les résultats des élèves de 2^{ème} année et de 5^{ème} année d'une même école devraient être corrélés. Ainsi, il est possible d'introduire le score moyen des élèves de 5^{ème} année de la même école que les élèves de 2^{ème} année, qui ont suivi presque une scolarité complète dans l'établissement, dans le modèle explicatif des acquis de 2^{ème} année. Naturellement, l'existence d'élèves ayant changé d'école tend à diminuer la significativité de la variable. Un autre problème, qui se pose de façon plus ou moins importante selon les pays, touche aux écoles à cycle incomplet où seul l'un des deux niveaux a pu être testé, l'autre n'existant pas. De ce fait, l'analyse proposée ne se fait pas sur le même échantillon de départ, les écoles n'ayant pas de classe de 2^{ème} année ou de 5^{ème} année étant exclues, ce qui peut faire redouter des biais importants et notamment que le résultat soit surtout lié à la distinction entre écoles à continuité éducative et écoles à cycle incomplet. Le cas du Mali traité dans le chapitre précédent permet de limiter ce type de problème puisque seules quatre écoles sont perdues dans l'exercice, ce qui représente seulement 53 élèves en moins sur 1662 dans l'échantillon de départ. La comparaison entre les modèles M1 et M2, qui ne se distinguent que par l'échantillon sur lequel ils portent, montre que les résultats sont très stables. La perte des quatre écoles n'introduit donc pas de biais dans l'estimation. L'introduction du score moyen de la classe de 5^{ème} année comme variable explicative des acquis de 2^{ème} année se révèle pertinente. La variable est très significative et le coefficient indique une relation positive assez marquée (Cf. tableau 4.2.2). Pour un écart-type en plus au score moyen de 5^{ème} année, on associe, en moyenne, environ un quart d'écart-type supplémentaire au score de fin de 2^{ème} année. Le résultat est stable selon qu'on utilise un

modèle avec estimation robuste des intervalles de confiance (M3) ou un modèle multiniveau (M4)⁸⁶. Enfin, on observe que l'introduction de la variable du score moyen de la classe de 5^{ème} année se traduit par un accroissement de 5 % du R² alors que l'ensemble des autres variables explicatives, hors score initial, augmente le R² de 10 % (Cf. chapitre 3, tableau 3.2.1). Compte tenu des limites de la mesure évoquées précédemment, on peut considérer que ce résultat est assez marqué et confirme l'existence d'un effet-école qui ne saurait être négligé.

⁸⁶ Cf. annexe 4.1.4 pour les sorties STATA des modèles.

Tableau 4.2.2 : La relation entre le score moyen de 5^{ème} année et les acquis en 2^{ème} année

	M1	M2	M3	M5
Score standardisé de français et mathématiques de début d'année	0.62*** (0.03)	0.61*** (0.03)	0.56*** (0.04)	0.60*** (0.02)
Enseignant contractuel	0.31*** (0.10)	0.29*** (0.10)	0.29*** (0.09)	0.31*** (0.10)
L'élève est une fille	-0.07 (0.04)	-0.07 (0.04)	-0.07* (0.04)	-0.05* (0.03)
L'élève est plus âgé que l'âge normal pour la classe	0.12** (0.05)	0.12** (0.05)	0.10* (0.05)	0.06* (0.04)
L'élève appartient au quintile le plus pauvre	-0.18*** (0.06)	-0.19*** (0.07)	-0.22*** (0.07)	-0.11** (0.05)
L'élève a redoublé sa 1 ^{ère} année	-0.15** (0.06)	-0.16*** (0.06)	-0.15** (0.06)	-0.10** (0.04)
L'élève a redoublé sa 2 ^{ème} année	-0.21*** (0.06)	-0.22*** (0.06)	-0.22*** (0.06)	-0.18*** (0.04)
Le maître est titulaire du BAC voire d'un diplôme du supérieur	-0.17 (0.13)	-0.19 (0.14)	-0.21* (0.12)	-0.21 (0.18)
Si c'était à refaire l'enseignant choisirait à nouveau le même métier	0.20** (0.10)	0.18* (0.10)	0.15* (0.08)	0.16 (0.08)
Le maître discute régulièrement avec ses collègues des difficultés rencontrées en classe	0.35*** (0.12)	0.35*** (0.14)	0.30** (0.13)	0.32* (0.16)
La salle de classe est construite en dur	-0.22** (0.11)	-0.21* (0.11)	-0.08 (0.10)	-0.09 (0.11)
La classe fonctionne en double flux	-0.37*** (0.12)	-0.37*** (0.12)	-0.31*** (0.11)	-0.34*** (0.12)
Coefficient de variation par classe (écart-type /score moyen)	0.63*** (0.21)	0.61*** (0.22)	0.52*** (0.19)	0.60*** (0.18)
Ancienneté du directeur d'école dans la fonction de directeur (en années)	0.02*** (0.01)	0.01*** (0.01)	0.01 (0.01)	0.01 (0.01)
Association de parents d'élèves jugée active ou très active par le directeur	-0.17 (0.10)	-0.17 (0.10)	-0.16 (0.10)	-0.17* (0.10)
Proportion d'enseignants contractuels dans l'école	-0.05 (0.24)	-0.04 (0.26)	-0.16 (0.24)	-0.20 (0.26)
L'école est en zone rurale	0.04 (0.11)	0.04 (0.11)	0.04 (0.11)	0.04 (0.11)
Score moyen standardisé de français et mathématiques des élèves de 5 ^{ème} année de l'école	/	/	0.24*** (0.01)	0.23*** (0.01)
Observations	1662	1609	1609	1609
R ²	0.39	0.38	0.43	/
Nombre d'écoles	128	124	124	124

Erreurs-type robustes entre parenthèses

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

On constate, en outre, que certaines variables ont vu leur coefficient et/ou leur significativité changer avec la prise en compte de cette nouvelle variable, sans pour autant qu'il y ait de problèmes de colinéarités dans le modèle (Cf. annexe 4.1.4). C'est notamment le cas de la variable relative à la construction de salles de classe en dur qui présente un coefficient négatif assez élevé (-0,21) et significatif dans le modèle initial et qui voit son coefficient diminuer

fortement (-0,08) et perdre sa significativité. Il semble donc que cette variable captait des effets de contextes pas nécessairement liés au type de construction scolaire. Une autre variable perd sa significativité, il s'agit de l'ancienneté du directeur. Ce qui tendrait à montrer que les directeurs les plus anciens ont tendance à se retrouver dans des écoles performantes mais ne sont pas nécessairement les principaux artisans de cette performance. On ne peut toutefois pas aller trop loin dans l'interprétation de ces résultats. Le fait que le niveau des élèves en fin de 5^{ème} année soit fortement lié aux acquis des élèves en 2^{ème} année démontre qu'il existe dans le cas malien un effet-école assez marqué. Toutefois, à l'instar de l'effet-classe, nos résultats ne nous permettent pas de dire de quoi se compose cet effet-école. On peut bien sûr avancer qu'il se compose d'une somme d'effets-classe (Bressoux, 2000) mais cela nous renvoie aux points de suspension de la discussion précédente concernant l'effet-classe.

Pourtant, pour les systèmes éducatifs, il y a un enjeu important à réduire les disparités entre écoles, notamment avec celles obtenant les résultats les plus faibles, pour améliorer la qualité des apprentissages et aussi l'équité. Il est vrai que, pour l'heure, les résultats des recherches fournissent assez peu d'éléments sur les causes de ces importants écarts de performance, mais différents éléments permettent de penser qu'une gestion locale appropriée permettrait une réduction de ces écarts. Cela suppose également l'existence d'indicateurs pertinents relatifs à la qualité des apprentissages mais aussi leur utilisation dans la gestion courante du système.

4.3 Disparités entre écoles et gestion de la qualité des apprentissages

L'existence de très fortes disparités, entre classes et entre établissements, concourt à la faible performance des systèmes éducatifs africains et à leur manque d'équité (UNESCO-BREDA, 2007). Si l'on dispose encore d'assez peu d'éléments factuels sur les sources de ces disparités, qu'on soupçonne d'être multiples et très variables selon les cas, certaines pistes sont à privilégier.

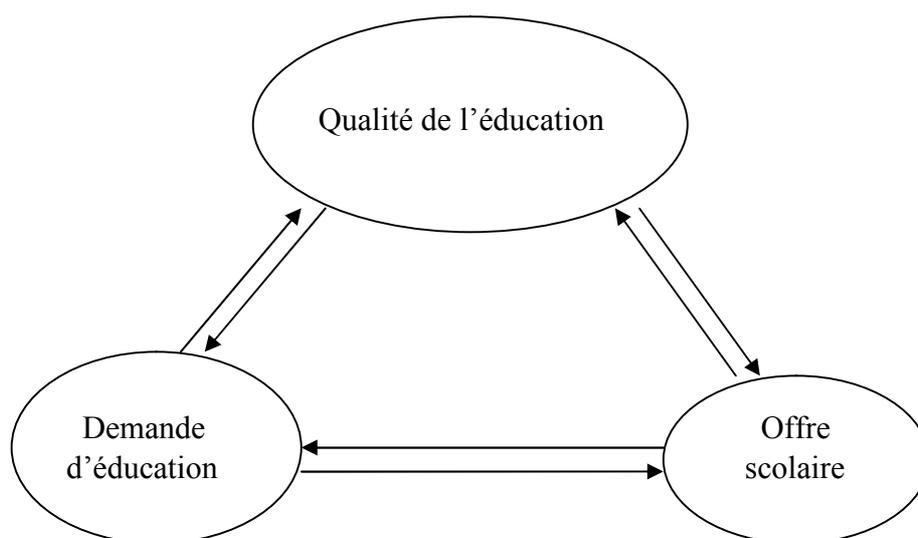
4.3.1 Les sources potentielles des disparités entre classes ou écoles

Sur le plan théorique, il est habituel de distinguer les facteurs d'offre et de demande scolaires. L'offre renvoie à l'institution scolaire et à la qualité des services dispensés aux populations. Il est donc assez facile de lier les problèmes de qualité des apprentissages à une défaillance de l'offre scolaire. C'est d'ailleurs ce que soulignent généralement les accusations fréquentes de déliquescence des services publics en général et éducatifs en particulier qu'on peut entendre ici ou là. Ainsi, l'insuffisance de formation des enseignants, leur manque de motivation, leur absentéisme marqué sont des arguments souvent avancés. Il faut toutefois faire preuve de circonspection devant ces arguments car, nous l'avons vu, ils ne permettent pas d'expliquer de façon satisfaisante les observations réalisées précédemment qui soulignent des cas de réussite spectaculaires comme des cas d'échec non moins spectaculaires et une variété de situations assez marquée. Si les problèmes d'offre scolaire existent bien, ils ne sont pas seuls en cause et il est nécessaire de prendre aussi en compte les facteurs de demande scolaire et probablement leurs interactions. Les attentes envers l'éducation peuvent être extrêmement variables au sein d'un même pays, entre zones rurales et zones urbaines mais aussi selon les milieux sociaux et culturels. Il ne faut pas perdre de vue que certaines populations s'opposent encore à la scolarisation de leurs enfants pour des raisons culturelles, religieuses ou tout simplement pratiques. Au-delà de ces cas extrêmes, on doit considérer que dans certaines zones, l'école est confrontée à des situations complexes où se combinent parfois son absence de légitimité auprès des familles et des conditions socio-économiques extrêmement difficiles. Ainsi, les enfants peuvent être des éléments clés de la survie des familles les plus pauvres à travers la main-d'œuvre qu'ils fournissent à certains moments de la vie des communautés comme les récoltes. On ne peut pas attendre de ces familles qu'elles envoient leurs enfants à l'école à ces périodes critiques. Là où la demande d'école des populations est fragile, il faudrait une offre

scolaire adaptée qui sache prendre en compte le contexte local. Dans ces contextes difficiles, c'est l'école qui doit s'adapter à son milieu et non l'inverse. Or, on peut légitimement s'interroger sur la capacité des systèmes éducatifs africains à s'adapter de la sorte. Toutefois, l'offre et la demande sont souvent interdépendantes. Ainsi, une demande inexistante constitue un environnement extrêmement difficile pour les acteurs du système éducatif et principalement les enseignants. Ceux-ci peuvent se retrouver dans des contextes à la fois très pauvres, différents sur le plan culturel et où leur statut social est remis en cause du fait de l'absence de considération des populations pour l'institution scolaire. Il est alors possible d'anticiper que ce type d'environnement ait des conséquences sur l'enseignement dispensé. On entre alors dans un cercle vicieux bien difficile à rompre. On ne peut bien sûr pas généraliser un tel cas de figure mais il serait tout aussi erroné de penser qu'il s'agit de cas exceptionnels.

On voit bien à travers ces différents arguments que les facteurs d'offre et de demande interagissent dans le processus scolaire et qu'une attention particulière doit leur être portée. Le schéma 4.1 illustre très simplement les relations entre offre scolaire, demande d'éducation et qualité de l'éducation.

Schéma 4.1 : Les interactions entre offre scolaire, demande d'éducation et qualité de l'enseignement



Une faible demande d'éducation peut constituer un environnement difficile pour l'école et influencer directement et indirectement, via l'offre, la qualité de l'éducation. Ainsi, quand la légitimité de l'école et de ses enseignants est remise en cause ou simplement ignorée, on peut aboutir à des phénomènes de découragement des équipes éducatives qui se traduiront dans la qualité du service offert. En outre, l'absence de contrôle social que cela implique contribue à inciter des phénomènes comme l'absentéisme ou le « tir-au-flanc » conformément à la théorie de l'agence (Mc Meekin, 2003). De plus, on peut supposer que les parents ne donneront que peu d'incitation à leurs enfants pour qu'ils fournissent le travail scolaire nécessaire ; il est même probable que les absences des élèves seront fréquentes, ce qui peut également contribuer à une faible qualité des apprentissages. On voit qu'on peut vite arriver à un cercle vicieux. A l'inverse, une offre scolaire déficiente (rentrée tardive, fort absentéisme des enseignants, manque de motivation...) va se traduire par de faibles acquisitions des élèves et une mauvaise image auprès des parents. Cela risque de remettre en cause l'intérêt de l'école pour les parents⁸⁷ et susciter différentes réactions. La plus problématique est le retrait des enfants de l'école par les familles dont l'intérêt pour l'éducation est limité, ce qui est évidemment assez dramatique quand cela intervient avant la fin de l'école primaire. Un autre phénomène observé ailleurs, mais à d'autres échelles est le « school choice ». Ceci nous renvoie au constat de C. Tiebout (1956) et de son « vote avec les pieds » où est introduite une liaison entre la mobilité géographique et les différences d'assortiments biens collectifs locaux/impôts offerts par les différentes collectivités, les agents choisissant leur localisation en fonction de l'assortiment le plus avantageux. Dans ce modèle, le vote avec les pieds révèle les préférences des contribuables/consommateurs comme la demande de biens privés révèle sur le marché leur disposition à payer. Il est évident que cette approche est liée à des hypothèses restrictives difficilement assimilables au cas de l'Afrique subsaharienne (mobilité parfaite, information parfaite, absence d'externalités géographiques). Toutefois, l'envoi des enfants dans les écoles privées, qui tend à se développer en Afrique, répond à cette logique même s'il n'implique pas nécessairement un déménagement de la famille. Par ailleurs, le phénomène très répandu de « confiage » des enfants peut, dans certains cas, être rapproché de cette démarche. Ainsi, quand les parents confient leurs enfants à des membres de la famille (parfois très éloignés) qui habitent dans des zones urbaines pour que leurs enfants soient scolarisés dans de supposées meilleures écoles, on se rapproche du concept de Tiebout, à cette

⁸⁷ Pour Prichett (2001), l'école africaine est dans une impasse car une partie de la population est consciente qu'elle ne remplit que peu son rôle,

nuance que seul l'enfant sera concerné. Toutefois, cela suppose à la fois une information sur la qualité de l'enseignement, un réseau social développé et une forte demande scolaire, trois conditions que ne peuvent pas remplir un grand nombre de ménages africains. En tout état de cause, le départ des écoles des enfants des familles qui portent le plus d'intérêt à l'école constitue plutôt un facteur aggravant de la situation des écoles qui perdent les élèves potentiellement les plus motivés.

Sur le plan empirique, différents résultats confortent certains arguments présentés. Ainsi, dans des études menées en Mauritanie et au Tchad, l'implication des parents dans l'école était associée à de meilleures acquisitions scolaires des élèves (PASEC, 2006a et 2006b). Ces résultats tendent à confirmer le lien entre demande d'éducation, mesurée ici par l'implication des parents, et la qualité de l'éducation. Cela va dans le sens d'une plus grande implication des communautés dans l'école que prônent bon nombre d'acteurs oeuvrant sur le terrain de l'éducation. Une conséquence de ce type de constat est qu'une sensibilisation des populations aux enjeux éducatifs à grande échelle pourrait avoir des retombées importantes sur l'amélioration du fonctionnement des systèmes éducatifs.

Cependant, l'analyse qui est proposée suppose que différents cas de figure peuvent être à l'origine des sous-performances marquées de certaines écoles. On peut distinguer les cas où les problèmes d'offre sont dominants, ceux où les problèmes de demande sont les plus marqués, et enfin les cas où les problèmes relèvent des deux dimensions. Les origines de ces problèmes, tant sur le plan de l'offre que de la demande, peuvent être multiples et donc nécessiter des réponses différentes en termes de mesures de remédiation. Cela suppose une grande capacité des systèmes éducatifs à établir des diagnostics locaux, à l'échelon de l'école, et à mettre en œuvre des mesures adaptées. Cela suppose également que les gestionnaires de l'éducation disposent d'indicateurs pertinents pour assurer le suivi.

4.3.2 Des indicateurs pour la gestion locale de la qualité des apprentissages

Une première difficulté, quand on recherche des indicateurs sur la qualité des apprentissages, tient à la disponibilité d'informations au sein des systèmes éducatifs. On sait par exemple que la notation des enseignants présente des biais importants qui interdisent toute comparaison entre classes (Cf. Bernard, Simon et Vianou, 2005). D'un autre côté, il n'existe pas dans les

pays africains, comme dans certains pays développés, de tests standardisés qui seraient administrés à toutes les écoles. Les enquêtes sur les acquis des élèves, quand elles existent, portent sur un échantillon limité d'écoles et ne sont pas réalisées de façon régulière. Finalement, la seule mesure disponible de la qualité des apprentissages, comparable entre établissements, reste les résultats aux examens nationaux. Ces résultats sont d'ailleurs disponibles, au mieux, sous forme de taux de réussite à l'examen. Il s'agit d'une mesure imprécise puisque les examens n'ont pas vocation à faire le bilan des acquis des élèves mais elle correspond toutefois à une norme que s'est fixé le système qui rend son utilisation légitime. Toutefois, un problème surgit rapidement, celui du risque d'un biais de sélection important du fait des fortes déperditions qu'enregistrent certaines écoles. On risque alors d'être confronté à des écoles ayant d'excellents résultats aux examens du fait qu'elles ont sélectionné les élèves. Il est donc utile de prendre en compte cette dimension. De plus, en utilisant simultanément des indicateurs de qualité et de quantité, on se rapproche d'une mesure plus satisfaisante de l'effectivité du droit à l'éducation (Bernard, 2003).

La démarche de Michaelowa (2001a), à partir des tests PASEC, ouvre des perspectives intéressantes. Elle propose un indicateur combinant les deux dimensions : le taux de connaissance de base (TCB) et le taux d'accès en 5^{ème} année. Le principe est de considérer un objectif d'apprentissage minimal pour le niveau retenu et de voir la proportion d'élèves qui atteignent cet objectif. A ce stade, l'indicateur ne concerne que les enfants qui sont toujours à l'école, il faut donc le combiner avec un indicateur permettant de nous ramener à l'ensemble des enfants.

Michaelowa définit ainsi le taux de connaissance de base :

$$\begin{aligned}
 \text{TCB} &= \frac{\text{nombre d'élèves avec connaissances de bases minimales à la fin du primaire}^{88}}{\text{nombre d'enfants à l'âge correspondant}} \\
 &= \left[\frac{\text{nombre d'élèves scolarisés à la fin du primaire}}{\text{nombre d'enfants à l'âge correspondant}} \right] \\
 &\quad \times \\
 &\quad \left[\frac{\text{nombre d'élèves avec connaissances de base minimales à la fin du primaire}}{\text{nombre d'élèves scolarisés à la fin du primaire}} \right]
 \end{aligned}$$

Cet indicateur nous permet de connaître la proportion d'enfants d'âge scolaire qui possède les connaissances minimales à la fin de l'école primaire.

Derrière ce principe séduisant, il est nécessaire de préciser les indicateurs retenus. Les connaissances minimales à la fin du primaire ne sont pas forcément appréciées de la même façon, et si elles n'ont pas été définies *a priori* par les personnes qui ont élaboré les tests, comme c'est le cas pour les tests PASEC, il faut trouver des critères *a posteriori*. Michaelowa a choisi *a posteriori* de retenir un taux de 40 % de réponses correctes aux tests comme seuil minimal. Bien sûr, ce seuil est discutable et il s'agit là d'un choix de l'auteur ; l'idéal reste la définition *a priori* sur une base pédagogique du seuil minimal. En ce qui concerne l'indicateur quantitatif, si l'auteur a retenu le taux de scolarisation en 5^{ème} année, il apparaît plus pertinent de prendre en compte le taux d'accès en 5^{ème} année. Finalement, pour notre cas particulier, on obtiendra le taux de connaissances de base en multipliant le taux d'accès en 5^{ème} année avec le taux d'élèves ayant un minimum de 40 % de réponses correctes dans les évaluations du PASEC.

⁸⁸ Dans la classe retenue.

Tableau 4.3.1 : Les taux de connaissances de base en 5^{ème} année de l'enseignement primaire dans cinq pays d'Afrique francophone

	Burkina Faso (1996)	Cameroun (2005)		Côte d'Ivoire (1996)	Madagascar		Sénégal (1996)
		1996	2005		1998	2005	
% d'élèves avec au moins 40 % de bonnes réponses*	60 %	64,4 %	60,5 %	59,5 %	55,8 %	69,6 %	33,7 %
Taux d'accès en 5 ^{ème} année ^{89**}	29 %	50 %	60 %	52 %	25 %	56,5 %	56 %
Taux de connaissance de base	17,4 %	32,2 %	36,3 %	30,9 %	14 %	39,3 %	21.2 %

* Source : PASEC

Cet indicateur combiné nous permet d'avoir une meilleure appréciation de la réalité éducative. Ainsi, on voit qu'en 1996, un pays comme le Burkina, qui est classé second sur la base des résultats aux tests, se retrouve en 4^{ème} position du fait de son très faible taux d'accès en 5^{ème} année. Inversement, le Sénégal qui avait les plus faibles résultats aux tests se hisse en troisième position grâce à une meilleure rétention. Toutefois, on constate que globalement, une faible proportion d'enfants accède à un seuil de compétences minimal avec des écarts importants entre pays. On observe un TCB d'un peu moins de 40 % au Cameroun contre moins de 18 % à Madagascar. Certes, on peut toujours objecter que le seuil minimal de compétences retenu est discutable, mais l'ampleur des chiffres ne laisse aucun doute sur la faible efficacité des systèmes éducatifs en question. D'énormes efforts restent à faire, tant sur le plan quantitatif que qualitatif, si l'on veut qu'une majorité des enfants sachent lire, écrire et compter à l'issue de l'enseignement primaire. Les évolutions observées au Cameroun et surtout à Madagascar, qui progresse sur les deux dimensions, entre 1996 et 2005 sont encourageantes mais encore insuffisantes par rapport aux objectifs EPT de 2015.

Cet indicateur, s'il peut encore être amélioré, a le mérite de fournir une image plus précise de la réalité éducative dans les pays. Sa principale limite est qu'il n'y a pas de données communes à l'ensemble des pays africains permettant un calcul homogène pour tous les pays. Une autre limite est qu'il se calcule sur la base d'évaluations ponctuelles basées sur des échantillons représentatifs. On ne peut donc pas le calculer régulièrement, tant que des évaluations régulières ne sont pas menées dans les pays. Enfin, le TCB ne peut pas être utilisé pour la

⁸⁹ Année 1998 ou proche.

gestion courante au niveau des établissements scolaires du système éducatif qui est l'objectif poursuivi ici⁹⁰. La démarche peut cependant être adaptée pour créer des indicateurs de gestion courante.

Il nous faut donc revenir aux examens nationaux qui sont la seule mesure régulière des acquisitions scolaires permettant de comparer l'ensemble des établissements au niveau national dans la majorité des pays africains. Ils sont clairement incontournables dans la réflexion sur les indicateurs de la qualité des apprentissages. De plus, il s'agit d'une information produite par le système dans son fonctionnement courant, elle n'implique donc pas, *a priori*, de coût supplémentaire. Cependant, outre les limites des examens en termes de mesure d'acquis scolaires déjà évoquées, d'autres problèmes pratiques se posent pour leur utilisation. Tout d'abord, dans beaucoup de pays africains, on ne peut pas rattacher les notes des élèves aux établissements scolaires. Ils sont enregistrés par centre d'examen, et des élèves de plusieurs écoles se retrouvent dans un même centre d'examen. On est donc réduit à utiliser l'information, moins précise en termes de mesure des acquis, relative aux taux de réussite. En outre, on sait aussi que les taux de réussite des établissements scolaires peuvent être sujets à caution. En effet, certains directeurs et chefs d'établissements ont tendance à développer des stratégies par rapport aux examens pour améliorer l'image de leur établissement. Ainsi, on observe parfois une sélection des élèves qui seront appelés à présenter l'examen, les plus faibles étant découragés de le faire ou incités à le faire en tant que candidats libres. On peut donc trouver des écoles avec 100 % de réussite mais avec très peu de candidats présentés à l'examen. Le taux de réussite usuel qui rapporte le nombre de reçus à l'examen au nombre de candidats présents est donc insuffisant. Afin d'améliorer cette mesure, il est utile de ramener le nombre de reçus d'une école aux effectifs de la classe de l'examen pour approcher un taux de réussite plus proche de la réalité que nous appellerons taux de réussite corrigé. Voilà donc un premier indicateur assez simple qui fournit une information sur la capacité d'un établissement à amener les élèves de dernière année au niveau de la norme qualitative fixée par le système éducatif qu'est l'examen. Toutefois, cet indicateur masque un aspect essentiel à savoir la rétention scolaire. On peut très bien avoir de bons résultats sur la base de cet indicateur en pratiquant une sélection importante tout au long du cycle, avec par exemple des taux de redoublement et d'abandons élevés, et finalement ne conserver dans l'établissement

⁹⁰ Il est aussi évident que cette mesure ne s'aurait s'étudier en longitudinal indépendamment de la dynamique du développement et particulièrement du bien-être des familles.

que les bons élèves⁹¹. Or, un bon établissement c'est avant tout un établissement qui va amener un maximum d'élèves en fin de cycle et non pas un établissement qui va sélectionner les meilleurs. Pour prendre en compte cette dimension, il faut connaître la proportion d'élèves qui accèdent en dernière année du cycle, c'est ce que mesure le taux de survie. Ces deux indicateurs, taux de réussite et taux de survie, donnent finalement une image assez bonne de la capacité d'une école à amener le maximum d'élèves au niveau requis par le système. Ils ont également l'avantage d'être basés sur une information existante, donc sans coût supplémentaire, et relativement familière aux acteurs du système éducatif.

On peut encore être plus synthétique en combinant les deux indicateurs, c'est-à-dire en multipliant le taux de survie avec le taux de réussite corrigé. Le résultat nous indiquera la probabilité pour un enfant qui entre dans un établissement de réussir l'examen du cycle, autrement dit de valider le cycle. Cet indicateur, directement inspiré du TCB de Michaelowa, est appelé taux de validation (Cf. Bernard, 2003). Il se construit comme suit :

Taux de réussite = reçus / présents

Taux de réussite corrigé = reçus / effectifs dans la classe de l'examen

Taux de survie = proportion d'élèves qui accèdent en dernière année du cycle.

Taux de validation = taux de réussite corrigé x Taux de survie en dernière année du cycle

Illustrons les propos précédents par un exemple chiffré :

	taux de réussite	taux de réussite "corrigé"	Taux de survie
Ecole A	100 %	90 %	40 %
Ecole B	90 %	50 %	80 %
Ecole C	60 %	60 %	80 %

On a trois types d'écoles, une école A où le taux de réussite apparent est excellent (100 %) mais où le taux de réussite "corrigé" est un peu plus faible (90 %) alors que la rétention est mauvaise (40 %). Dans l'école B, le taux de réussite apparent est assez élevé (90 %) mais le taux de réussite "corrigé" est nettement plus faible (50 %). On voit ici que l'établissement sélectionne soigneusement ses candidats à l'examen. Par contre, la survie dans l'établissement

⁹¹ Stratégie bien connue dans le cas français et mise en évidence par les indicateurs annuels de la DEPP sur la performance des Lycées.

s'avère nettement meilleure (80 %). Enfin, dans l'école C le taux de réussite est plus faible (60 %) mais tous les élèves de la classe ont été présentés (taux de réussite= taux de réussite corrigé). Par ailleurs, le taux de survie est identique à l'école B (80 %). Voyons maintenant comment le taux de validation va rendre compte de ces trois situations :

	taux de réussite "corrigé"	Taux de survie	Taux de validation
Ecole A	90 %	40 %	36 %
Ecole B	50 %	80 %	40 %
Ecole C	60 %	80 %	48 %

Le taux de validation fait ressortir l'école C comme la plus efficace et l'école A comme la moins efficace. Dans la première, la probabilité pour un élève de valider le cycle est de 48 % contre 36 pour la seconde. On constate que le classement par rapport à l'indicateur traditionnel de taux de réussite est singulièrement différent puisque l'école C qui était en dernière position se trouve maintenant à la première place.

Toutefois, dans certains contextes, le taux de réussite « corrigé » peut présenter certaines limites qu'il faudrait prendre en compte pour être plus précis. En effet, par exemple, dans certains pays l'accès au collège est conditionné par un concours différent de l'examen de fin de cycle primaire. Une pratique fréquente est que les élèves qui ont eu l'examen, mais ont échoué pour l'accès au collège « redoublent » pour préparer à nouveau l'entrée en collège. Même si en règle générale cette pratique n'est pas autorisée, puisqu'il s'agit de faux redoublants ayant en fait validé le cycle, elle reste relativement fréquente. Aussi, les écoles avec une proportion plus importante de faux redoublants vont être avantagées par rapport aux autres. Une façon de mieux apprécier les résultats du taux de validation est alors de considérer la proportion de redoublants dans la classe de l'examen.

Par ailleurs, la survie et la réussite à l'examen ont le même poids dans le calcul de l'indicateur. Ainsi, une école qui a un taux de survie de 80 % et un taux de réussite de 40 % aura un taux de validation de 32 % tout comme une école qui a un taux de survie de 40 % et de réussite de 80 %. Or, cela implique des choses différentes en matière d'appui à ces deux écoles. C'est pourquoi la décomposition de cet indicateur (taux d'achèvement et taux de réussite corrigé) est importante pour analyser la situation des écoles et identifier le type de problème (qualité *versus* rétention). Il est donc préférable d'utiliser ces trois indicateurs

simultanément. De façon générale, il faut plaider pour l'utilisation de différents indicateurs pour apprécier le mieux possible la situation des différents établissements. L'intérêt du taux de validation est qu'il résume en un seul chiffre la situation de l'établissement, ce qui permet d'identifier les établissements en difficulté. Pour une meilleure perception des causes des difficultés, il apparaît nécessaire de compléter avec d'autres indicateurs comme ceux que nous avons mentionnés précédemment. Cela appelle ensuite des actions de remédiation pour aider les établissements connaissant des difficultés.

Le taux de validation, comme tout indicateur, présente des limites qu'il convient de prendre en compte pour une utilisation appropriée. Une première limite de cet indicateur tient à la prise en compte de la qualité des apprentissages à travers un taux de réussite plutôt qu'à travers les acquis des élèves. De ce fait, on ne mesure pas la proportion d'enfants qui ont développé les compétences requises pour le cycle, mais seulement la proportion de ceux qui ont atteint la norme qualitative fixée par le système (réussite à l'examen). Si la sélection à l'examen est très marquée, on aura une image un peu éloignée des compétences réelles des élèves et inversement s'il y a peu de sélection. Ainsi, les enquêtes ponctuelles sur les apprentissages demeurent importantes pour cerner le niveau de performance du système éducatif sur le plan des acquisitions. Une dernière insuffisance réside dans la non-prise en considération par cet indicateur de la durée moyenne nécessaire pour obtenir un diplômé. C'est un aspect de l'efficacité qui n'est pas traité et il faut donc compléter avec d'autres indicateurs (nombre d'années élèves pour un diplômé, coefficient d'efficacité interne). Enfin, comme nous l'avons déjà souligné, cet indicateur est avant tout conçu pour piloter au niveau le plus fin (circonscription scolaire), ce qui demande une couverture statistique exhaustive. En revanche, il ne peut être utilisé dans le cadre de comparaisons internationales pour situer le niveau de performance moyen d'un pays par rapport à ses pairs.

La discussion précédente a montré qu'il était possible d'utiliser des indicateurs permettant d'identifier les écarts de performance marqués entre les établissements scolaires. Il est d'ailleurs intéressant de remarquer qu'il est nécessaire de combiner des informations sur la qualité des apprentissages avec des informations plus quantitatives pour avoir une bonne image de la situation. Toutefois, l'existence d'indicateurs relativement simples d'usage ne garantit pas leur utilisation.

4.4 La gestion locale comme moyen de réduire les disparités entre écoles

Dans ce chapitre, il a été mis en évidence que les disparités entre classes et entre établissements scolaires étaient une problématique commune à de nombreux systèmes éducatifs et pas seulement dans les pays les plus pauvres. Il faut cependant reconnaître que le phénomène apparaît beaucoup plus marqué dans les pays pauvres et singulièrement dans les pays africains. La comparaison peut toutefois être hasardeuse dans la mesure où cela pourrait simplement retranscrire les plus fortes inégalités sociales qui caractérisent les pays les plus pauvres. La variété observée, au sein même des pays africains, permet de montrer que cette explication ne suffit pas à rendre compte de ces disparités. Les analyses plus poussées basées sur des effets fixes classes prenant un certain nombre de facteurs en compte amènent à s'interroger sur les causes de ces disparités entre classes et établissements. L'association entre effet-classe et effet-maître, courante dans la littérature, est apparue très contestable dans le contexte considéré. On dispose, à l'heure actuelle, de bien peu d'éléments pour expliquer cet effet-classe mais il existe des arguments importants pour avancer qu'il ne se limite pas à un effet-maître et inclut d'autres dimensions. Le temps effectif d'enseignement est un bon exemple de piste de recherche à explorer. L'importance de l'effet-classe est la mesure de notre ignorance, elle appelle de nouveaux travaux de recherche.

Si l'on se place du point de vue du responsable et du gestionnaire du système éducatif, il est évident que la situation actuelle qui traduit de fortes inégalités ne peut pas attendre les résultats des recherches en cours ou à venir. Les sources des inégalités, qu'elles relèvent de l'offre scolaire et/ou de la demande d'éducation, sont nécessairement multiples et il est assez improbable qu'une seule et même recette puisse être appliquée à tous les établissements scolaires. Sous l'angle de la gestion, il est d'abord essentiel d'être en mesure d'identifier les différences de performances entre établissements ; il faut ensuite être en mesure, au cas par cas, d'apporter des solutions aux problèmes que rencontrent certaines écoles. L'information existe et des indicateurs relativement simples peuvent permettre aux gestionnaires d'identifier les établissements connaissant de grandes difficultés ou s'écartant nettement de la moyenne. Une telle démarche, qui n'apparaît pas totalement inaccessible aux systèmes éducatifs, permettrait probablement de réduire significativement les écarts considérables observés.

Il reste à se demander pourquoi, jusqu'à présent, les systèmes éducatifs ont échoué à prendre en compte ces aspects dans leur gestion courante. Un premier élément de réponse tient au fait

que la gestion pédagogique, qui s'attache à ce que les moyens mis à disposition des écoles se transforment en apprentissages scolaires effectifs, est avant tout normative. L'inspection reste le mode de régulation le plus fréquent dans les pays d'Afrique francophone. Outre les problèmes divers et variés qui font que l'inspecteur inspecte finalement très peu, il n'y a guère de place pour la prise en compte des résultats dans les procédures administratives suivies. De ce point de vue, il apparaît nécessaire qu'il y ait une évolution notable des pratiques de gestion pour que celles-ci soient plus axées sur les résultats. Cela dit, sans sous-estimer la complexité de mise en œuvre d'une telle réforme, on peut aussi penser qu'une plus grande implication des parents dans le suivi de l'école serait de nature à améliorer la situation. Au-delà du contrôle social au plus proche de l'école que cela induit, cette implication peut aussi avoir un aspect motivant dans le sens où les enseignants peuvent voir, à travers l'intérêt porté à leur travail, une forme de reconnaissance sociale. A vrai dire, tout cela n'est pas très nouveau, mais on sait encore mal comment l'opérationnaliser. Dans ce domaine, l'approche quasi expérimentale pourrait être d'une grande utilité pour évaluer de nouvelles procédures de gestion et des mesures incitatives en direction des familles.

Conclusion générale

Les vertus heuristiques de la fonction de production expliquent probablement son succès auprès des économistes de l'éducation dans l'analyse des systèmes éducatifs. Toutefois, il n'est pas évident que son statut épistémologique soit le même pour tous. En effet, il a été souligné, dans le premier chapitre, que considérer la fonction de production éducative comme un cadre théorique du processus d'acquisitions scolaires soulevait de nombreuses difficultés. Si ce cadre est rarement évoqué explicitement dans les recherches, il sous-tend assez fréquemment les interprétations des travaux empiriques. Ainsi, la présentation par divers auteurs de l'absence, ou plutôt de l'insuffisance de convergence, des résultats des études empiriques comme une limite, semble se rattacher à cette vision. En effet, l'idée implicite d'une nécessaire convergence des résultats incorpore, entre autres, l'idée d'optimalité de la fonction de production. D'un point de vue purement empirique, il n'est pas évident *a priori* que les mêmes inputs⁹² doivent nécessairement avoir, dans les différentes estimations réalisées, le même effet dans des contextes différents, ni même en moyenne sur un grand nombre d'études⁹³. Utiliser cet argument pour rejeter les travaux réalisés est pour le moins excessif, mais surtout résulte d'une vision biaisée, pour ne pas dire erronée, de la démarche suivie. Bien sûr, il ne s'agit pas ici de nier le problème de qualité inégale des études réalisées et les sérieux problèmes méthodologiques que posent certaines d'entre elles, sans évoquer les questions posées par la qualité des mesures sur les entrants, d'une part, et le produit éducatif par l'évaluation des élèves, d'autre part. Il est toutefois possible de se donner des critères scientifiques d'appréciation de la qualité des études comme on l'a vu dans les deux premiers chapitres et aussi d'adopter une certaine prudence dans l'interprétation des résultats. Par ailleurs, le constat de l'insuffisance de convergence dans les résultats correspond à un certain angle de lecture des travaux empiriques. L'examen de la relation entre les caractéristiques des enseignants et les acquisitions des élèves, effectué dans le deuxième chapitre, a montré qu'il y avait bien un certain nombre de convergences dans les études considérées. Ainsi, dans ces études, au Nord comme au Sud, au travers de ces démarches, explicites ou implicites, de la

⁹² A supposer qu'ils soient réellement identiques ce dont on peut douter.

⁹³ Il n'est d'ailleurs pas sûr que les travaux répondant à de nouvelles approches, quasi expérimentale et/ou expérience naturelle, présentent un degré de convergence beaucoup plus marqué.

fonction de production de l'école, le rôle des caractéristiques des enseignants apparaît modéré et rend peu compte de la variété des résultats des élèves.

Le risque, en tournant le dos aux résultats de ces travaux, est qu'on se prive de pistes de recherche tout à fait pertinentes. Les limites des analyses entreprises dans le cadre de la fonction de production éducative devraient au contraire être soigneusement considérées. Si l'on part du constat que la relation entre les inputs et les acquis des élèves est pour le moins instable, cela doit induire un questionnement. Celui-ci peut être ponctuel et porter sur l'un ou l'autre des facteurs d'intérêt. Par exemple, l'absence d'effet de la formation professionnelle des enseignants sur les acquis des élèves, analysée au chapitre deux, doit amener à se poser certaines questions : la mesure de la formation professionnelle dont on dispose pour les analyses est-elle adéquate ? La formation est-elle appropriée à la situation que rencontrent les maîtres dans la classe où ils sont affectés ? Les enseignants mettent-ils en œuvre ce qu'ils ont appris en formation dans les salles de classe ? Autant de questions qui peuvent déboucher sur de nouveaux travaux. Si l'on prend l'analyse réalisée au chapitre 3 sur les enseignants contractuels au Mali, certaines questions restent en suspens. Certes, les résultats sont relativement robustes et indiquent que les enseignants contractuels obtiennent des résultats assez flatteurs comparés à leurs collègues fonctionnaires, notamment en 2^{ème} année. Il demeure néanmoins une incertitude quant à la durabilité de ce résultat tant il semble dépendant de questions de motivation, qui peuvent être intimement liées à l'ancienneté dans le métier et/ou à la représentation que l'enseignant se fait de son métier. Par ailleurs, à l'évidence si nous nous projetons en avant, rien ne dit que, d'ici quinze ans, le contractuel ne sera pas tout autant démotivé que l'est le fonctionnaire qui possède aujourd'hui cette ancienneté. Difficile donc de conclure sur ce terrain mouvant, et de nouvelles études devront essayer de prendre en compte les dynamiques à l'œuvre dans le système éducatif malien. Une perspective temporelle est désormais nécessaire pour aller plus loin sur cette question. Ceci est certainement un objectif stratégique d'études dans les travaux d'analyse du suivi du processus EPT.

De façon plus générale, la question de la relation instable entre inputs et acquis induit directement celle de l'efficacité des écoles. Un constat, marquant et récurrent, réside en effet dans les disparités entre classes et établissements. Là encore, c'est un constat commun aux pays développés et en développement, mais avec une intensité plus importante dans ces derniers. Cette différence conforte d'ailleurs les réserves qu'on peut avoir sur l'assimilation

de l'effet-classe à l'effet-maître, puisqu'elle laisse entrevoir l'incorporation d'effets de contexte dans la mesure utilisée. Les effets fixes sont une mesure résiduelle dont l'interprétation doit être des plus prudentes, il n'est pas tout à fait exagéré de dire qu'ils mesurent surtout notre ignorance. De fait, les recherches ont jusqu'ici amené peu de réponses, tout au plus ont-elles soulevé des pistes de recherche, comme le temps d'enseignement en Afrique, dont les observations de terrain indiquent qu'il peut être très variable d'une école à l'autre. Il s'agit d'une voie qui devra être explorée par les recherches futures. Toutefois, d'autres questionnements accompagnent cette problématique des disparités entre classes et établissements. Il s'agit essentiellement de l'impuissance de la gestion actuelle des systèmes éducatifs à réduire ces disparités. Celles-ci, compte tenu des problèmes d'équité qu'elles soulèvent, devraient constituer une préoccupation majeure pour les gestionnaires de l'éducation. Il semble, au contraire, que les systèmes éducatifs soient complètement myopes devant ce problème. L'incapacité chronique dans la gestion courante des écoles à prendre en compte des résultats est probablement un facteur important à considérer. Pourtant, les résultats du chapitre 4 ont permis de montrer que les informations disponibles dans les systèmes éducatifs permettaient le calcul d'indicateurs de la performance des établissements. L'utilisation de tels indicateurs permettrait d'identifier les établissements rencontrant des difficultés et, dans un second temps, de leur apporter un appui. Malheureusement, on tarde à voir ce schéma se mettre en place. De fait, la gestion courante ignore le plus souvent les performances des établissements sur le continent africain. Le processus de pilotage y est à sens unique du centre vers la périphérie ; sans qu'en retour le centre ne s'intéresse aux difficultés et spécificités d'application d'un levier de la politique éducative et aux nécessaires adaptations de terrain, ceci surtout vis-à-vis des populations les plus fragiles pour leur accès à l'école⁹⁴.

L'absence d'un cadre théorique permettant de rendre compte du comportement des acteurs dans le processus d'apprentissage constitue une limite à ce stade. En effet, les différences d'efficacité des établissements tendent à mettre en avant le rôle prépondérant des acteurs.

⁹⁴ Une illustration très actuelle de ceci dans le contexte de l'Afrique subsaharienne réside dans les impacts de la gratuité scolaire. Cette mesure, à priori généreuse et visant à accompagner l'EPT, prise par les autorités centrales, entraîne des difficultés d'application sur le terrain. Dans certains pays francophones (Cameroun, Bénin), mais encore plus pour l'Afrique anglophone, cette mesure a gonflé la demande d'éducation des plus défavorisés, des zones reculées, sans que l'offre scolaire ne s'adapte à ces nouvelles populations. Pire, les ressources que représentaient les frais de scolarité au niveau des écoles sont difficilement remplacées par des subventions de l'Etat, laissant parfois les établissements sans ressources pour faire face à l'afflux de nouveaux élèves. Les conditions se dégradent et les frustrations s'accumulent.

D'ailleurs, les quelques réflexions théoriques menées récemment sur les processus scolaires pointent toutes l'importance d'une approche comportementale des individus et des groupes. S'il n'y a pas de cadre théorique global et cohérent qui se dégage pour l'heure, certaines approches ouvrent des pistes intéressantes. C'est notamment le cas de l'approche principal-agent utilisée par Pritchett et Filmer (1999) et McMeekin (2003). Ce dernier propose une analyse du processus d'apprentissage où s'imbriquent plusieurs relations entre principal (P) et agent (A) : entre l'enseignant (P) et l'élève (A), entre les parents (P) et l'élève (A), entre les parents (P) et l'enseignant (A) et encore entre l'administration scolaire (P) et l'enseignant (A). L'enseignant, au cœur du processus, est celui qui détient le plus d'informations, sur l'effort des élèves, mais aussi sur son propre travail. Sous cet angle, l'asymétrie de l'information sur ce qui se passe dans la classe devient un enjeu central. On retrouve là des questions familières à la théorie microéconomique avec cette différence près que le produit dont il est question ici est incorporé à l'individu et donc pas directement observable, ce qui rend la question de l'information encore plus complexe et peut-être plus sensible. Si ces pistes de recherche, encore récentes, nécessitent d'être approfondies, elles ont le mérite de mettre en évidence l'importance de l'information sur la production scolaire dans la gestion des systèmes éducatifs. Le fait qu'à l'heure actuelle, les systèmes éducatifs africains produisent peu d'informations sur les performances des établissements apparaît déjà comme une difficulté. L'attention des chercheurs s'est aussi beaucoup focalisée sur les incitations à donner aux différents acteurs pour qu'ils fournissent les efforts nécessaires. L'idée dominante, pas très originale, est qu'une plus grande concurrence entre établissements permettrait d'améliorer l'efficacité de ceux-ci. Maurin (2007), passant en revue les différents travaux empiriques portant sur les réformes concurrentielles menées au Chili (voucher school), aux Etats-Unis (charter schools) et en Grande-Bretagne, souligne que « séduisants sur le papier, les principes concurrentiels s'appliquent ainsi assez difficilement à l'univers scolaire » (p. 226). Il rappelle que l'un des problèmes qui se posent pour que la concurrence soit effective est l'existence d'une information fiable sur la capacité des établissements à faire progresser, de manière endogène, les écoles. Cette information étant très difficile à connaître, même pour de bons économètres, les parents sont surtout sensibles à la composition sociale des établissements, ce qui aboutit à des phénomènes de ségrégation scolaire. Il faut ajouter à l'argumentaire de Maurin sur les limites de la concurrence entre établissements la dimension spatiale, prégnante en Afrique subsaharienne. Les écoles ne sont pas distribuées identiquement sur le territoire et changer d'école peut impliquer, faute d'écoles proches, des déplacements importants tout simplement impossibles pour de nombreuses familles, c'est particulièrement frappant pour les

zones rurales en Afrique. On ne voit donc pas très bien comment cette concurrence pourrait être effective dans ce contexte. Au-delà, les mesures d'incitation ciblant les enseignants, les parents (primes) ou les établissements (subventions) se heurtent toujours à la même difficulté, identifier le critère pertinent sur lequel baser l'incitation au risque d'effets pervers induits par les comportements opportunistes. A l'heure actuelle, nos connaissances sur le sujet sont fragiles et la plus grande prudence est de mise avec ce genre de politiques. Si la piste concurrentielle semble risquée, il n'en reste pas moins que la question des disparités entre classes et établissements demande des réponses en matière de gestion et de pilotage. La réflexion théorique doit se poursuivre sur ces questions mais il semble que de nouveaux travaux empiriques sur ces aspects pourraient aussi être utiles. L'approche quasi-expérimentale pourrait s'avérer très utile pour tester la mise en œuvre de nouvelles pratiques de gestion notamment sur le continent africain. C'est en partie la démarche suivie par Glewwe, Ilias et Kremer (2003), au Kenya, où ils évaluent des incitations ciblées sur les enseignants avec des résultats pour le moins mitigés. Toutefois, il serait bon de faire des analyses un peu plus originales sur la base des résultats des travaux déjà menés. Ainsi, la question du temps d'enseignement devrait faire l'objet d'une attention particulière. On sait, par exemple, que les calendriers scolaires sont inadaptés aux contextes de certaines zones, il serait intéressant d'expérimenter des mesures d'adaptation de ces calendriers au niveau des établissements et de voir s'il y a des conséquences sur la fréquentation et les acquis des élèves. Des initiatives intéressantes ont été menées en Asie du sud, en partie autour de l'initiative BRAC, mais en l'état, elles restent une référence lointaine compte tenu des différences de contexte. On pourrait aussi évaluer de façon plus rigoureuse des expériences d'implication des familles dans la gestion de l'école. Il serait également intéressant d'évaluer un éventuel impact d'une diffusion publique de l'information sur la performance des écoles aussi bien en zone urbaine qu'en zone rurale. On peut multiplier les exemples, car très peu a été fait jusqu'ici pour mieux appréhender les enjeux autour de la gestion et c'est probablement une priorité pour les recherches futures.

Au regard des enjeux pour l'Afrique, les connaissances apportées par la recherche peuvent sembler dérisoires ou tout au moins insuffisantes. Si la tâche reste immense, il faut cependant rappeler le développement des travaux et les progrès réalisés au cours des deux dernières décennies qui ont permis d'accumuler un certain nombre de résultats utiles aujourd'hui. Les recherches, effet induit de l'objectif de l'EPT, sont désormais de plus en plus nombreuses en Afrique et il faut souhaiter qu'elles sachent accorder la priorité aux principaux enjeux de

politique éducative. Assurer une scolarisation primaire complète de qualité à tous les enfants du continent africain reste un défi majeur auquel les chercheurs, au même titre que les autres acteurs de la communauté internationale, doivent s'attacher pour tenter de le relever.

ANNEXES

ANNEXE 1.1.1 : Construction des scores ajustés

L'objectif de la variable de score ajusté est d'estimer le score qu'aurait l'élève moyen dans chacune des classes enquêtées. Nous décrivons ci-dessous les étapes de la construction

(1) Création des variables muettes classes

```
tabulate NUMECOLE, generate(C)
```

(2) Modèle

```
reg SFIN5A100 SINI5A100 AGEPLUS_ES FILLE NBRED PAUVRE DOMHASSAN_ES /*  
*/ C1-C67 C69-C121
```

(3) Calcul des valeurs moyennes pour les principales caractéristiques élèves

```
egen M5A1=mean(SINI5A100)  
sum M5A1  
tab AGEPLUS  
tab FILLE  
egen NBMOYRED=mean(NBRED)  
tab PAUVRE  
tab DOMHASSAN
```

(4) Conservation d'une observation par classe

```
sort NUMECOLE  
by NUMECOLE: drop if _n!=_N  
count  
sort NUMECOLE  
gen C=_n
```

(5) Calcul du score ajusté : on récupère les coefficients (significatifs) du modèle précédent (2) qu'on met en relation avec les valeurs moyennes des différentes variables

```
gen SFIN5A_ajust=0  
  
local i=1  
while `i'<67 {  
    replace  
    SFIN5A_ajust=_cons+(_b[SINI5A100]*M5A1)+(_b[AGEPLUS_ES])+(_b[FILLE])+(_b[NBRED]*  
NBMOYRED)+(_b[PAUVRE])+(_b[DOMHASSAN_ES])+(_b[C`i']) if C==`i'  
        local i=`i'+1  
    }  
  
local i=67  
while `i'<68 {  
    replace  
    SFIN5A_ajust=_cons+(_b[SINI5A100]*M5A1)+(_b[AGEPLUS_ES])+(_b[FILLE])+(_b[NBRED]*  
NBMOYRED)+(_b[PAUVRE])+(_b[DOMHASSAN_ES]) if C==`i'  
        local i=`i'+1  
    }  
  
local i=69  
while `i'<122 {  
    replace  
    SFIN5A_ajust=_cons+(_b[SINI5A100]*M5A1)+(_b[AGEPLUS_ES])+(_b[FILLE])+(_b[NBRED]*NBM  
OYRED) +(_b[PAUVRE])+(_b[DOMHASSAN_ES])+(_b[C`i']) if C==`i'  
        local i=`i'+1  
    }  
}
```

ANNEXE 1.2.1 : Méthodes et instruments du Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC)

Les éléments repris ici sont présentés dans la plupart des rapports PASEC ; pour plus de détails, le lecteur peut se référer à l'un de ces rapports (Cf. bibliographie).

La méthodologie du PASEC

Le principe de base de la méthodologie du PASEC repose sur la comparaison. Il consiste à tirer parti de la variété des situations scolaires pour identifier les modèles de scolarisation les plus performants. L'objectif étant avant tout d'identifier les facteurs qui agissent sur la progression des élèves, il faut pouvoir mettre en relation les niveaux moyens d'acquisition des élèves en arabe, en français et en mathématiques avec les conditions matérielles et organisationnelles dans lesquelles ils sont scolarisés.

Cependant, on ne saurait se limiter aux scores moyens des élèves à la fin de l'année scolaire. En effet, tout le passé des élèves interviendrait dans cette mesure et l'on ne saurait pas réellement ce qui est imputable à l'année évaluée par rapport aux années antérieures. Cela implique d'avoir une mesure du niveau des acquis des élèves en début et en fin d'année scolaire. A ce stade, on ne sait rien des conditions de scolarisation respectives des différents élèves. Certains sont scolarisés dans des classes à double vacation, d'autres dans des classes à grands effectifs ou multigrades, d'autres sont issus de milieux favorisés, etc. Il y a de nombreux paramètres qui interviennent dans le processus d'acquisition des élèves, s'ils n'étaient pas pris en compte, ils pourraient nous amener à tirer des conclusions erronées. Il faut donc avoir recours à des techniques capables de prendre en compte simultanément les principaux facteurs qui interviennent dans le processus d'apprentissage et de dégager leur effet propre, indépendamment de l'effet des autres facteurs.

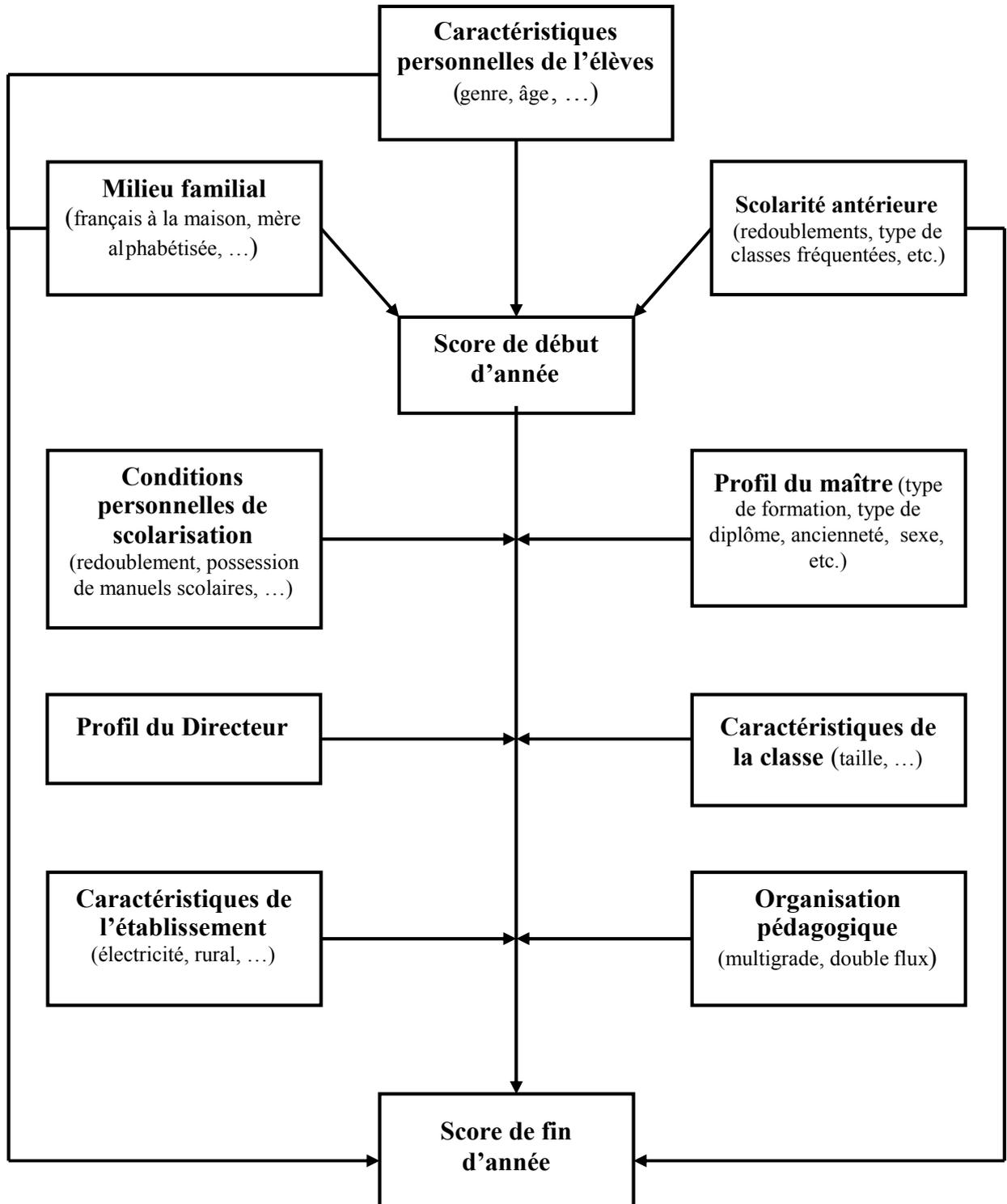
Ces principes méthodologiques évoqués sont au cœur de la méthodologie du PASEC. En effet, le programme observe le niveau d'acquisition des élèves de deuxième et de cinquième année de l'enseignement primaire dans des situations de scolarisation aussi diversifiées que possible. Les acquis des élèves sont mesurés en français, en mathématiques⁹⁵ à deux reprises : au début (pré-test ou profil d'entrée, bâti sur les acquis jugés prioritaires du programme de l'année scolaire précédente 1^{ère} année ou 4^{ème} année) et à la fin de l'année scolaire (post-test ou profil de sortie, fonction du programme de l'année scolaire en cours 2^{ème} année ou 5^{ème} année). L'utilisation d'instruments standardisés (mêmes épreuves, conditions de passation homogènes et correction centralisée) permet d'établir des comparaisons aux niveaux national et international.

⁹⁵ Le PASEC a également introduit des tests d'arabe en Mauritanie, de malgache à Madagascar et d'anglais au Cameroun et à Maurice.

Le protocole d'enquête prévoit le recueil simultané d'un certain nombre de données contextuelles, d'ordre institutionnel, social, économique et culturel. Les données relatives aux élèves sont recueillies en début d'année, tandis que celles relatives aux enseignants, aux classes et aux écoles le sont en fin d'année. Dans l'explication des performances scolaires des élèves, ces données permettent de séparer les effets dus au milieu de ceux dus à la situation pédagogique.

Le schéma causal ci-après, élaboré pour une année scolaire, permet de mettre en évidence cette méthodologie d'analyse.

Schéma n°1 : Schéma d'analyse causale du PASEC



Les instruments

Les évaluations menées par le PASEC, comme nous venons de le voir, n'ont pas pour objectif prioritaire de mesurer le niveau des acquisitions des élèves dans différentes disciplines. Cependant, pour analyser les déterminants des acquisitions des élèves, il est nécessaire de s'appuyer sur des tests dans différentes disciplines. Ces tests ont été conçus spécifiquement pour répondre aux exigences de la méthodologie d'analyse du programme. Ils ont bien sûr été élaborés en référence aux programmes scolaires des pays d'Afrique francophone et, dans la mesure du possible, sur les aspects communs de ces différents programmes.

Toutefois, ces tests ne cherchent pas à mesurer le degré de maîtrise de tel ou tel point du programme. La construction des items qui les composent répond avant tout à la nécessité d'avoir des différences dans les résultats des élèves car il est indispensable, pour les analyses que l'on souhaite mener, qu'il existe une variété de résultats chez les élèves. Une homogénéité des résultats, qu'ils soient bons ou faibles, limiterait les analyses permettant d'identifier les déterminants des apprentissages.

Par conséquent, les tests utilisés ne sauraient être considérés comme une échelle de mesure des acquisitions des élèves aussi précise qu'on le souhaiterait dans le cas d'une évaluation *stricto sensu* des acquis des élèves. Cela dit, ces tests constituent une échelle de mesure, certes imparfaite⁹⁶, mais qui reste une source d'informations précieuses pour les systèmes éducatifs étudiés, ceci dans la mesure où les données fondées sur des épreuves standardisées demeurent relativement rares en Afrique subsaharienne. De plus, le PASEC utilise les mêmes tests⁹⁷ de français et de mathématiques dans les différents pays où il travaille, ce qui offre une base comparative particulièrement intéressante.

Pour recueillir des informations sur les conditions de scolarisation et le contexte social, économique et culturel, des questionnaires ont été administrés aux élèves, aux enseignants et aux directeurs d'école. En ce qui concerne les élèves, le questionnaire est administré en début d'année. Un questionnaire est administré aux enseignants en début d'année et un autre en fin d'année afin de prendre en compte les changements de maître en cours d'année. Enfin, le questionnaire directeur est administré en fin d'année.

Les instruments utilisés par le PASEC sont disponibles auprès du Secrétariat technique permanent de la Confemen (www.confemen.org).

⁹⁶ Sur ce point, il convient de relativiser car les travaux menés par la Banque mondiale et le Pôle de Dakar pour construire un indicateur comparable de qualité (Indice Africain de la Qualité de l'Éducation) ont montré la très grande convergence entre les résultats du MLA, du PASEC et du SACMEQ.

⁹⁷ Des adaptations sont réalisées selon les contextes nationaux (changement des préfixes, des noms d'objets ou d'animaux, etc.), et parfois quelques items sont ajoutés mais il reste une base commune à tous les pays.

Principes d'échantillonnage

L'échantillonnage a pour objectif de permettre de tirer des conclusions valables pour toute une population à partir d'un nombre restreint d'individus de cette population. Dans le cas du PASEC, l'exercice se complexifie pour au moins trois raisons : (i) pour répondre à certaines questions, il s'avère nécessaire de faire des entorses aux principes élémentaires de construction d'un échantillon, (ii) le fait d'administrer des tests en début d'année pose des problèmes de disponibilité d'informations fiables⁹⁸ qui peuvent avoir des conséquences sur l'échantillon final, et (iii) la prise en compte de deux niveaux (2^{ème} et 5^{ème} années) dans l'évaluation est une source de complexité supplémentaire.

La procédure d'échantillonnage retenue est un sondage stratifié à deux degrés ou sondage stratifié en grappes. Le principe de ce sondage est de retenir dans un premier temps un ensemble d'écoles proportionnellement à leurs poids en nombre d'élèves de 2^{ème} année et de 5^{ème} année. Ces écoles sont choisies par stratification, de façon à être représentatives de l'ensemble du système éducatif du pays. Lorsqu'une école est choisie, on procède ensuite au tirage d'un nombre fixe d'élèves dans chacun des deux niveaux d'étude, 2^{ème} année et 5^{ème} année.

L'efficacité de ce type de sondage dépend du degré d'homogénéité ou de ressemblance (sur le plan scolaire) des élèves d'une même école. En effet, le nombre d'élèves à retenir par classe et le nombre total d'écoles à enquêter sont fonction du degré d'homogénéité des élèves d'une même classe. Par exemple, si les élèves d'une même classe présentent tous un même niveau en mathématiques, on pourrait se limiter à n'enquêter qu'un seul élève par classe et enquêter un maximum d'écoles afin de couvrir l'ensemble des spécificités-écoles. Si, par contre, les élèves sont très différents à l'intérieur d'une même classe, pour cerner au mieux toutes leurs caractéristiques on doit alors étudier beaucoup d'élèves dans chaque classe et diminuer le nombre d'écoles à échantillonner. Le degré d'homogénéité des élèves d'une même classe est donc déterminant pour définir la taille de l'échantillon, et par suite pour la précision des estimateurs. Il est mesuré par un indicateur appelé roh ou coefficient de corrélation intraclasse. Il n'est pas connu *a priori* sauf enquête antérieure. En se fixant une valeur probable, on peut déduire le nombre d'élèves à enquêter par école⁹⁹, de même que le nombre d'écoles à retenir comme unités primaires, afin de rendre le sondage par grappes aussi précis qu'un sondage aléatoire simple.

⁹⁸ En effet, on ne dispose au mieux que des données de l'année précédente or des changements peuvent avoir lieu dans les écoles notamment dans le mode d'organisation des classes.

⁹⁹ Cette déduction se fait à partir d'une relation établie par Kish (1965) dans le cadre des sondages en grappe.

Problématique de l'analyse des données hiérarchiques

On ne saurait aborder l'étude des comportements d'un individu sans se référer au groupe social auquel il appartient. De la même manière, pour aborder l'étude des caractéristiques d'un groupe social, il est bon de tenir compte de la société à laquelle appartient ce groupe. On rencontre souvent dans les sciences sociales ce genre de réalité hiérarchisée. Dans notre exemple, les individus sont liés au sein d'un groupe, et les groupes à leur tour sont liés au sein de la société. Le « **groupe** » et la « **société** » sont des niveaux hiérarchiques supérieurs à celui de l'individu qui, ici, est le niveau le plus fin. Par convention, le niveau le plus fin (qui ne sera pas toujours l'individu) est appelé niveau 1, le niveau hiérarchique immédiatement supérieur sera alors le niveau 2, etc.

Une collecte bien menée devra produire des données rendant compte de cette structure hiérarchique du phénomène : certaines variables décriront de façon spécifique les individus pris isolément, d'autres variables serviront à décrire de façon globale les individus pris par groupe, et d'autres encore serviront à décrire les groupes pris dans un contexte social plus large. Les données d'enquêtes scolaires comme celles du PASEC sont précisément un exemple type de données hiérarchiques : les élèves appartiennent au groupe que constitue la classe, la classe appartient à l'école, l'école appartient à une circonscription scolaire, etc. On distinguera alors les variables de niveau élève, les variables de niveau classe, les variables de niveau école, etc.

Le mode de tirage et la nature des variables d'intérêt permettent de considérer que les données du PASEC sont à deux niveaux : niveau *élève*, et niveau *classe*¹⁰¹. Les données se présentent alors comme suit.

¹⁰⁰ Cette annexe est extraite d'un document de travail du PASEC : Bernard, J.M., Hounkannounon, B., Kotchoni, R., & Nkengne Nkengne, A.P. (2005). Techniques d'estimation sur les données hiérarchiques. Note technique PASEC, Dakar : CONFEMEN.

¹⁰¹ Compte tenu du fait qu'on tire une classe de chaque type (CP2 et CM1) par école et qu'à chaque type de classe correspond un modèle, les variables écoles seront attribuées au niveau classe.

Elèves	Classes	Y (variable à expliquer)	V (variables explicatives de niveau "élève")	Z (variables explicatives de niveau "classe")
1	1	Y_{11}	V_{11}	Z_1
2		Y_{21}	V_{21}	
3		Y_{31}	V_{31}	
4		Y_{41}	V_{41}	
5		Y_{51}	V_{51}	
1	2	Y_{12}	V_{12}	Z_2
2		Y_{22}	V_{22}	
3		Y_{32}	V_{32}	
4		Y_{42}	V_{42}	
1	3	Y_{13}	V_{13}	Z_3
2		Y_{23}	V_{23}	
3		Y_{33}	V_{33}	
4		Y_{43}	V_{43}	
5		Y_{53}	V_{53}	
6		Y_{63}	V_{63}	

Pourquoi est-il si important de prendre en compte cette structure hiérarchique des données dans la démarche économétrique ?

En réalité, les influences extérieures communes aux individus d'un groupe donné, ainsi que l'influence du groupe sur chacun de ses membres induisent une certaine liaison entre les observations faites sur les individus de ce groupe. Les modèles de régression basés sur les MCO (moindres carrés ordinaires) considèrent, de façon tacite, à travers l'hypothèse d'indépendance des résidus, que les observations sont indépendantes. Or, les variables décrivant les groupes prennent la même valeur pour tous les individus d'un même groupe, ce qui fait qu'un petit nombre d'observations est répété un grand nombre fois (voir tableau suivant). Par exemple, si on connaît la valeur que prend la variable "âge du maître" pour un élève donné, on connaît également la valeur de cette variable pour tous les élèves de sa classe. L'hypothèse d'indépendance des observations est ainsi invalidée.

De plus, les individus d'un même groupe partagent des conditions en commun dont certaines sont latentes. Ils ont alors tendance à se ressembler et le groupe a tendance à devenir homogène. On peut donc s'attendre à ce que la variance du terme d'erreur ne soit pas la même lorsqu'on passe des observations d'un groupe à celles d'un autre groupe, ce qui invaliderait l'hypothèse d'homoscédasticité.

Elèves	Classes	Y (variable à expliquer)	V (variables explicatives de niveau "élève")	Z (variables explicatives de niveau "classe")
1	1	Y_{11}	V_{11}	Z_1
2		Y_{21}	V_{21}	Z_1
3		Y_{31}	V_{31}	Z_1
4		Y_{41}	V_{41}	Z_1
5		Y_{51}	V_{51}	Z_1
1	2	Y_{12}	V_{12}	Z_2
2		Y_{22}	V_{22}	Z_2
3		Y_{32}	V_{32}	Z_2
4		Y_{42}	V_{42}	Z_2
1	3	Y_{13}	V_{13}	Z_3
2		Y_{23}	V_{23}	Z_3
3		Y_{33}	V_{33}	Z_3
4		Y_{43}	V_{43}	Z_3
5		Y_{53}	V_{53}	Z_3
6		Y_{63}	V_{63}	Z_3

Pour toutes ces raisons, les MCO présentent de sérieuses limites dans l'analyse des données hiérarchiques.

Insuffisances des MCO

L'utilisation des MCO dans l'estimation de ce modèle sur des données hiérarchiques pose deux problèmes. Le premier concerne le biais d'estimation des intervalles de confiance pour les coefficients estimés, et le second concerne le biais d'estimation des coefficients eux-mêmes. Passons en revue chacun de ces deux points.

- **Biais d'estimation des intervalles de confiance**

On considère le modèle linéaire suivant :

$$Y = X\beta + \varepsilon, \text{ avec } \text{var}(\varepsilon) = \Omega$$

Y, variable dépendante, est le score final des élèves. X est un ensemble de variables censées expliquer le score final. X est composé de variables d'au moins deux niveaux. Dans notre exemple, X contient V, Z et éventuellement la variable constante.

L'estimateur MCO de β dans le modèle ci-dessus s'écrit : $b = (X'X)^{-1}X'Y$. Comme les MCO s'appliquent sous l'hypothèse d'homoscédasticité et d'indépendance des résidus ($\Omega = \sigma^2 I$), la variance de l'estimateur b est supposée être égale à : $\Sigma_1 = \sigma^2(X'X)^{-1}$ et les intervalles de confiance sont calculées sur cette base.

Mais compte tenu de la structure hiérarchique des données, les problèmes suivants se posent :

- i) La répétition des valeurs des variables de niveau *classe* pour tous les élèves d'une même classe remet en cause l'hypothèse d'indépendance des observations ;
- ii) Les élèves d'une même classe partagent un certain nombre de variables ensemble dont certaines sont probablement inobservables et par conséquent non prises en compte dans la spécification du modèle. Comme la partie non expliquée par le modèle correspond au résidu, il est souhaitable dans le plus simple des cas de concevoir un terme d'erreur du à chaque niveau de données (*voir tableau suivant*).

Elèves	Classes	Y	V	Erreur de niveau 1	Z	Erreur de niveau 2
1	1	Y_{11}	V_{11}	u_{11}	Z_1	ω_1
2		Y_{21}	V_{21}	u_{21}		
3		Y_{31}	V_{31}	u_{31}		
4		Y_{41}	V_{41}	u_{41}		
5		Y_{51}	V_{51}	u_{51}		
1	2	Y_{12}	V_{12}	u_{12}	Z_2	ω_2
2		Y_{22}	V_{22}	u_{22}		
3		Y_{32}	V_{32}	u_{32}		
4		Y_{42}	V_{42}	u_{42}		
1	3	Y_{13}	V_{13}	u_{13}	Z_3	ω_3
2		Y_{23}	V_{23}	u_{23}		
3		Y_{33}	V_{33}	u_{33}		
4		Y_{43}	V_{43}	u_{43}		
5		Y_{53}	V_{53}	u_{53}		
6		Y_{63}	V_{63}	u_{63}		

La variance de l'erreur due au niveau classe dépend du degré d'homogénéité de chaque groupe et n'est donc pas forcément constante. D'une manière générale, les hypothèses de constance de variance et d'indépendance des résidus ne peuvent être admises *a priori*.

Le modèle étant donc *a priori* hétéroscédastique, la matrice de variance covariance de l'estimateur MCO \mathbf{b} de $\boldsymbol{\beta}$ est alors : $\Sigma_2 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Omega\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}$ avec $\Omega \neq \sigma^2\mathbf{I}$. C'est sur Σ_2 que devraient se baser les inférences et les intervalles de confiances basées sur Σ_1 ne sont pas fiables en réalité.

L'utilisation de Σ_2 permet en fait de corriger a posteriori l'hétéroscédasticité. Les intervalles de confiance ainsi obtenus sont dits robustes car plus fiables. Or Σ_2 est inconnu car Ω est inconnu. Nous présenterons plus loin les principales approches d'estimation robuste des intervalles de confiance.

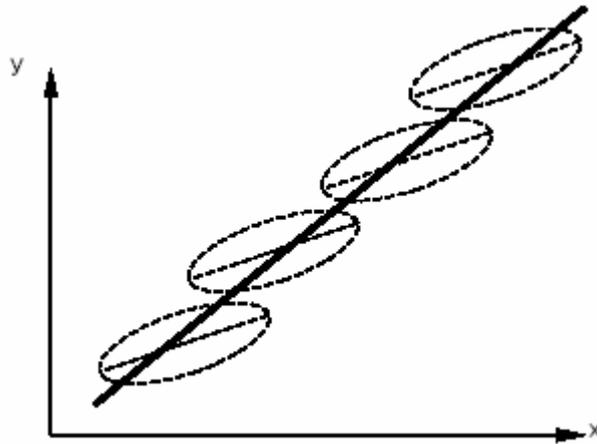
- **Biais d'estimation des coefficients**

Nous allons voir à présent comment l'utilisation des MCO sur des données hiérarchiques peut conduire à une estimation biaisée des coefficients. Examinons les trois situations suivantes.

Situation 1 : Le graphique suivant présente un ajustement linéaire simple entre la variable y et la variable x effectué sur quatre classes, d'abord un ajustement sur le nuage formé par les

quatre classes prises ensemble, ensuite un ajustement partiel pour chacune des classes prises séparément.

Si une droite ajustée à un nuage a pour équation $y = a x + b$, alors a et b sont respectivement les coefficients de la variable x et de la constante dans une régression MCO de y sur x restreinte à l'ensemble des points du nuage.

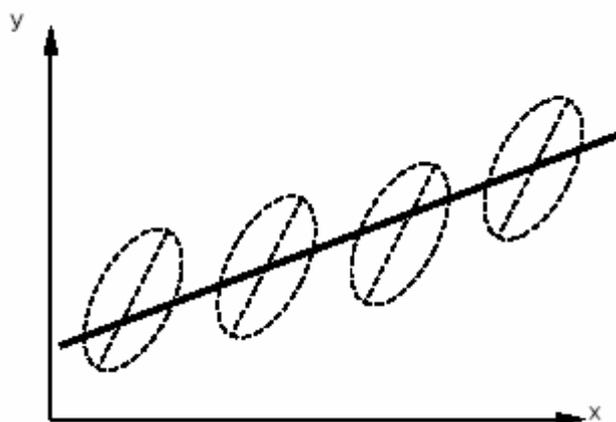


Ce graphique indique que les quatre droites d'ajustement partiel ont la même pente, mais des ordonnées à l'origine différentes. Dans une logique explicative, cela signifie que l'effet de x sur y , c'est-à-dire le coefficient de régression linéaire de y sur x , est le même quelle que soit la classe considérée.

Mais compte tenu de la forme du nuage d'ensemble, la droite d'ajustement globale présente une pente plus grande que la pente commune des droites d'ajustements partielles. Cela signifie qu'une régression linéaire faite sur l'ensemble du nuage sans tenir compte de l'existence de sous nuages formés par les classes attribuera un effet biaisé vers le haut à la variable explicative x .

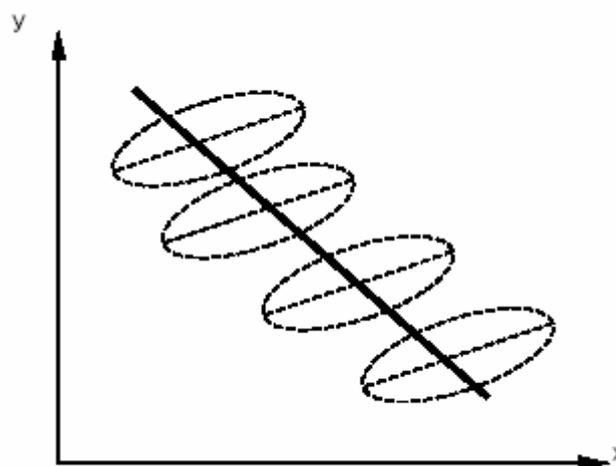
Il s'agit là d'un exemple de surestimation de l'effet d'une variable explicative due à la non prise en compte de la structure hiérarchique des données.

Situation 2 : Le graphique suivant présente une situation légèrement différente de la précédente. Comme précédemment, la variable explicative x a le même effet sur y quelle que soit la classe considérée. Mais compte tenu de la nouvelle forme du nuage d'ensemble, la droite d'ajustement globale présente une pente plus petite que la pente commune des droites d'ajustement partiel. Une régression MCO faite sur l'ensemble du nuage sans tenir compte de l'existence de sous-groupes que constituent les classes attribuera donc un effet biaisé vers le bas à la variable explicative x .



Ce second exemple montre qu'on peut sous-estimer l'effet d'une variable explicative en ne prenant pas en compte la structure hiérarchique des données.

Situation 3 : Dans ce troisième exemple, la variable explicative x continue d'avoir le même effet sur y quelle que soit la classe considérée. Mais la forme du nuage d'ensemble est telle que la droite d'ajustement global présente désormais une pente négative pendant que les droites d'ajustement partiel ont une pente commune positive.



La non-prise en compte de la hiérarchie des données conduit ici à affirmer que la variable x a un effet négatif sur la variable y , alors qu'en réalité son effet est positif. Il ne s'agit donc plus simplement d'une sous-estimation ou d'une surestimation, mais d'une inversion de l'effet.

Ces trois situations mettent en évidence de façon non exhaustive les pièges liés à l'utilisation des MCO sur des données de type hiérarchiques. Les sections suivantes présentent quelques remèdes aux différentes insuffisances des MCO.

- **MCO avec estimation robuste des écarts types**

Les MCO avec estimation robuste des écarts types ont pour but de remédier au biais d'estimation des intervalles de confiance. Ces approches des MCO proposent des intervalles de confiance réputés plus fiables car basés sur une estimation robuste de la variance des estimateurs. Ces méthodes fournissent donc rigoureusement les mêmes estimations ponctuelles que celle fournie par les MCO. Le logiciel STATA intègre ces approches dans ses commandes sous forme d'options dans la commande de régression MCO : il s'agit des options "*robust*" et "*cluster*".

Dans la littérature économétrique, on rencontre généralement trois estimations dites robustes de Σ_2 basées sur les carrés des résidus de régression MCO (e_i^2).

i) **Huber (1967) ; White (1980) :** $\hat{\Sigma}_2 = \frac{n}{n-K} (X'X)^{-1} X' \text{diag}(e_i^2) X (X'X)^{-1}$

Ici, Ω est estimée par $\hat{\Omega} = \frac{n}{n-K} \text{diag}(e_i^2)$ où $\frac{n}{n-K}$ est un facteur correcteur de degrés de liberté (n est le total nombre d'observations et K est le nombre de variables explicatives, constante comprise). Le principe retenu pour l'estimation de Ω indique qu'on a juste supposé que la variance des résidus n'est pas constante. L'hypothèse d'indépendance des erreurs est toujours maintenue.

ii) **MacKinnon et White 1 (1985) :** $\hat{\Sigma}_2 = (X'X)^{-1} X' \text{diag}\left(\frac{e_i^2}{1-h_{ii}}\right) X (X'X)^{-1}$

Cette estimation propose une correction par les effets de levier h_{ii} . h_{ii} est le $i^{\text{ème}}$ terme diagonal de la matrice de projection orthogonale $H = X(X'X)^{-1}X'$ et représente l'influence de l'observation i sur la prédiction de Y_i . On a :

$$\hat{\Omega} = \text{diag}\left(\frac{e_i^2}{1-h_{ii}}\right).$$

Plus l'effet de levier d'une observation est grand, plus la variance estimée du terme d'erreur correspondant est grande.

iii) **MacKinnon et White 2 (1985) :** $\hat{\Sigma}_2 = (X'X)^{-1} X' \text{diag}\left(\frac{e_i^2}{(1-h_{ii})^2}\right) X (X'X)^{-1}$

$$\hat{\Omega} = \text{diag}\left(\frac{e_i^2}{(1-h_{ii})^2}\right)$$

L'idée en divisant e_i^2 par $(1 - h_{ii})^2$ est qu'il est nécessaire de corriger davantage les observations à fort effet de levier. Long et Ervin (2000) ont trouvé après plusieurs simulations que cette estimation de Σ_2 est meilleure lorsqu'on travaille sur de petits échantillons (moins de 250). Pour des échantillons de plus de 500 observations, les deux estimations précédentes peuvent être utilisées pour les inférences.

Procédure de mise en œuvre sous stata

Il existe trois options sous STATA correspondant respectivement à ces trois estimations de Σ_2 .

- i) Pour baser les inférences sur l'estimation de Σ_2 proposée par Huber-White, la commande suivante doit être utilisée:

```
regress Y X, robust
```

Y est la variable endogène et X est l'ensemble des variables explicatives.

- ii) Pour baser les inférences sur la première estimation de Σ_2 proposée par McKinnon-White (voir 4.1.1, ii) *McKinnon et White 1*), la commande suivante doit être utilisée :

```
regress Y X, hc2
```

- iii) Pour baser les inférences sur la première estimation de Σ_2 proposée par McKinnon-White (voir ci-dessus 4.1.1, iii) *McKinnon et White 2*), la commande suivante doit être utilisée :

```
regress Y X, hc3
```

MCO, option cluster

L'option cluster lève l'hypothèse d'indépendance des résidus à l'intérieur d'un même groupe ou strate, mais maintient cette hypothèse d'un groupe à l'autre. Cette option présente donc un avantage par rapport aux variantes précédentes.

L'estimation robuste de la variance des estimateurs passe par le calcul de la contribution des individus au score du modèle (le score d'un modèle est la dérivée de sa log vraisemblance). La formule générale de l'estimation robuste de la variance des estimateurs est la suivante :

$$\hat{\Sigma}_2 = \left(\frac{n-1}{n-K} * \frac{J}{J-1} \right) \hat{V}_{MCO} \left(\sum_{j=1}^J U_j' U_j \right) \hat{V}_{MCO}$$

Avec : n le nombre total d'observations, K le nombre de variables explicatives (constante comprise) ;

J le nombre de clusters ou classes ;

$U'_j, j = 1$ à J , est la contribution de la classe j au score du modèle ;

\hat{V}_{MCO} est la matrice estimée de variance covariance des MCO : $\hat{V}_{MCO} = \hat{\sigma}^2 (X'X)^{-1}$;

$U_j = \sum_{i \in \text{classe } j} u_i$ où u_i est la contribution de l'individu i au score.

Dans le cas limite où chaque groupe contient un seul individu, (c'est-à-dire $J = n$) la formule devient :

$$\hat{\Sigma}_2 = \left(\frac{n}{n-K} \right) \hat{V}_{MCO} \left(\sum_{i=1}^n U'_i U_i \right) \hat{V}_{MCO}$$

Cette formule est équivalente à celle du *robust* proposée au **3.1.1 i)**. (cf. Manuel de STATA, *User's Guide*).

Procédure de mise en œuvre sous Stata

Avant de pouvoir lancer une régression avec l'option cluster sous STATA, il faut disposer d'une variable discrète dont les valeurs déterminent les différentes classes. Soit G cette variable. La commande à saisir est alors :

```
regress Y X, robust cluster(G)
```

Avantages et limites de l'estimation robuste des écarts types

L'avantage incontestable de l'estimation robuste des écarts types est de fournir des intervalles de confiance plus fiables que ceux qu'on obtiendrait en utilisant les MCO simples. Outre cet avantage, le cluster permet de prendre en compte a posteriori les liaisons éventuelles entre les observations d'une même classe. De toutes les approches présentées dans cette section, l'option cluster est celle qui convient le mieux aux données hiérarchiques.

Mais à l'instar des autres options, l'option cluster présente l'inconvénient majeur de se baser sur l'estimation ponctuelle MCO qui, en présence d'hétéroscédasticité, n'est pas de variance minimale bien qu'étant théoriquement sans biais (Théorème de GAUSS-MARKOV). Cela se traduit par des intervalles de confiance relativement larges par rapport à ceux qu'auraient fournis des méthodes prenant en compte l'hétéroscédasticité dans l'estimation même des coefficients. On sait de plus que pour des données hiérarchiques, les MCO peuvent induire dans certaines situations des biais dans l'estimation ponctuelle des coefficients. C'est

pourquoi les modèles hiérarchiques ou multiniveaux deviennent peu à peu la norme dans la modélisation statistique du processus d'acquisition scolaire.

**ANNEXE 1.2.3 : Comparaison des modèles explicatifs des scores de
5^{ème} année en français et mathématiques selon la prise en compte du score
initial**

Dictionnaire des variables utilisées

Variables	
STINI5FM	Score standardisé de français et mathématiques de début d'année
STFIN5FM	Score standardisé de français et mathématiques de fin d'année
FILLE	L'élève est une fille
AGE	Age
AISE	Niveau de vie
REDOUBLE	L'élève a redoublé au moins une fois
LIV_FRMT	L'élève possède les manuels de français et mathématiques
SERVICE	Ancienneté du maître (en années)
NIVCYCLB	Le maître a fréquenté le second cycle secondaire
DIPCYCLB	Le maître est titulaire du BAC
FPIL	Le maître a suivi une formation professionnelle initiale longue (1 an et plus)
DOUBLFLX	La classe fonctionne en double flux
MULTIGRA	La classe est multigrade
TCLASSE	Nombre d'élèves dans la classe
RURAL	L'école est en zone rurale

Madagascar

Modèle sans score initial

Linear regression

Number of obs = 2063
F(13, 117) = 4.07
Prob > F = 0.0000
R-squared = 0.0980
Root MSE = .95336

Number of clusters (NUMECOLE) = 118

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	.0410919	.0443863	0.93	0.356	-.0468127	.1289966
AGE	-.0774165	.018596	-4.16	0.000	-.1142448	-.0405882
AISE	.1104889	.0950328	1.16	0.247	-.0777187	.2986964
REDOUBLE	-.1514113	.060369	-2.51	0.014	-.270969	-.0318536
LIV_FRMT	.1205228	.0782939	1.54	0.126	-.0345342	.2755798
SERVICE	.0158887	.0134358	1.18	0.239	-.0107201	.0424976
NIVCYCLB	.2918294	.1482166	1.97	0.051	-.0017057	.5853646
DIPCYCLB	.414768	.2777006	1.49	0.138	-.1352035	.9647396
FPIL	.1808907	.1742889	1.04	0.301	-.1642794	.5260608
DOUBLFLX	-.2779618	.2245563	-1.24	0.218	-.7226837	.1667602
MULTIGRA	.3208658	.4497313	0.71	0.477	-.5698036	1.211535
TCLASSE	.0008144	.0045617	0.18	0.859	-.0082197	.0098486
RURAL	-.1417182	.1353317	-1.05	0.297	-.4097356	.1262991

_cons	.4624132	.4639079	1.00	0.321	-.4563321	1.381158

Variable	VIF	1/VIF				
SERVICE	1.72	0.580255				
DIPCYCLB	1.61	0.619339				
FPIL	1.30	0.767010				
NIVCYCLB	1.20	0.834769				
AGE	1.16	0.861494				
REDOUBLE	1.13	0.886622				
MULTIGRA	1.11	0.897490				
AISE	1.10	0.913063				
TCLASSE	1.08	0.926793				
LIV_FRMT	1.06	0.940690				
DOUBLFLX	1.06	0.947140				
RURAL	1.05	0.955467				
FILLE	1.02	0.983470				

Mean VIF	1.20					

Le VIF (Variance Inflation Factor) permet de juger du degré de multicolinéarité des différentes variables d'un modèle :

$$VIF(x_j) = 1 / (1 - R^2_j)$$

où R^2_j est le carré du coefficient de corrélation multiple qui est obtenu quand on régresse x_j sur l'ensemble des autres variables explicatives.

Modèle avec score initial

Linear regression

Number of obs = 2063
 F(14, 117) = 10.14
 Prob > F = 0.0000
 R-squared = 0.2611
 Root MSE = .86306

Number of clusters (NUMECOLE) = 118

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.4295877	.0596082	7.21	0.000	.3115367	.5476387
FILLE	.0381097	.0393565	0.97	0.335	-.0398339	.1160533
AGE	-.0794395	.0166504	-4.77	0.000	-.1124147	-.0464643
AISE	.1703945	.0870909	1.96	0.053	-.0020845	.3428735
REDOUBLE	-.078179	.0564579	-1.38	0.169	-.1899909	.0336329
LIV_FRMT	.0995081	.0740518	1.34	0.182	-.0471476	.2461638
SERVICE	.0190673	.012314	1.55	0.124	-.00532	.0434546
NIVCYCLB	.1240313	.1309885	0.95	0.346	-.1353845	.3834471
DIPCYCLB	.3006303	.2351424	1.28	0.204	-.165057	.7663176
FPIL	.0362586	.1423333	0.25	0.799	-.2456251	.3181423
DOUBLFLX	-.3069038	.1628365	-1.88	0.062	-.6293929	.0155854
MULTIGRA	.5531214	.4246498	1.30	0.195	-.2878753	1.394118
TCLASSE	-.0024937	.0042144	-0.59	0.555	-.01084	.0058526
RURAL	.0160428	.1223917	0.13	0.896	-.2263476	.2584333
_cons	.5282562	.4001179	1.32	0.189	-.2641564	1.320669

. vif

Variable	VIF	1/VIF
SERVICE	1.73	0.579258
DIPCYCLB	1.62	0.617083
FPIL	1.33	0.753076
NIVCYCLB	1.24	0.805936
AGE	1.16	0.861450
REDOUBLE	1.14	0.880883
STINI5FM	1.13	0.881890
MULTIGRA	1.13	0.885268
AISE	1.10	0.910176
TCLASSE	1.09	0.916038
RURAL	1.08	0.922018
LIV_FRMT	1.06	0.940133
DOUBLFLX	1.06	0.946830
FILLE	1.02	0.983457
Mean VIF	1.21	

Tests de différence des coefficients des deux modèles :

```

. test FILLE=0.04
      F( 1, 117) =    0.00
      Prob > F =    0.9618

. test AGE=-0.08
      F( 1, 117) =    0.00
      Prob > F =    0.9732

. test AISE=0.11
      F( 1, 117) =    0.48
      Prob > F =    0.4894

. test REDOUBLE=-0.15
      F( 1, 117) =    1.62
      Prob > F =    0.2059

. test LIV_FRMT=0.12
      F( 1, 117) =    0.08
      Prob > F =    0.7825

. test SERVICE=0.02
      F( 1, 117) =    0.01
      Prob > F =    0.9398

. test NIVCYCLB=0.29
      F( 1, 117) =    1.61
      Prob > F =    0.2077

. test DIPCYCLB=0.41
      F( 1, 117) =    0.22
      Prob > F =    0.6427

. test FPIL=0.18
      F( 1, 117) =    1.02
      Prob > F =    0.3146

. test DOUBLFLX=-0.28
      F( 1, 117) =    0.03
      Prob > F =    0.8691

. test MULTIGRA=0.32
      F( 1, 117) =    0.30
      Prob > F =    0.5841

. test TCLASSE=0
      F( 1, 117) =    0.35
      Prob > F =    0.5552

. test RURAL=-0.14
      F( 1, 117) =    1.63
      Prob > F =    0.2049

```

Test de l'endogénéité du score initial

```

. reg STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT DOMFRANC ALPHMERE /*
> */ SERVICE NIVCYCLB DIPCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL , cluster (NUMECOLE)

```

```

Linear regression                               Number of obs =    2063
                                                F( 15, 117) =     4.23
                                                Prob > F       =    0.0000
                                                R-squared     =    0.1318
                                                Root MSE     =    .93692

Number of clusters (NUMECOLE) = 118

```

STINI5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	.0013861	.045644	0.03	0.976	-.0890095	.0917816
AGE	.0073323	.0195078	0.38	0.708	-.0313019	.0459666
AISE	-.1159518	.0899347	-1.29	0.200	-.2940627	.0621591
REDOUBLE	-.1726387	.0579501	-2.98	0.004	-.2874058	-.0578716
LIV_FRMT	.0666987	.0840805	0.79	0.429	-.0998184	.2332158
DOMFRANC	-.120534	.1124193	-1.07	0.286	-.3431746	.1021066
ALPHMERE	.293681	.0767776	3.83	0.000	.141627	.4457349
SERVICE	-.0085024	.0103037	-0.83	0.411	-.0289084	.0119036
NIVCYCLB	.3793189	.1331768	2.85	0.005	.1155692	.6430685
DIPCYCLB	.2147053	.3418629	0.63	0.531	-.4623364	.8917469
FPIL	.3360195	.1697232	1.98	0.050	-.0001084	.6721473
DOUBLFLX	.0607265	.2539266	0.24	0.811	-.442162	.5636149
MULTIGRA	-.519189	.365112	-1.42	0.158	-1.242274	.2038961
TCLASSE	.0072548	.0049301	1.47	0.144	-.002509	.0170187
RURAL	-.3705324	.1263148	-2.93	0.004	-.6206923	-.1203725
_cons	-.3736445	.3841659	-0.97	0.333	-1.134465	.3871759

```
. predict RES, residual
```

```
.
. reg STFIN5FM STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT /*
> */ SERVICE NIVCYCLB DIPCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL RES , cluster (NUMECOLE)
```

```
Linear regression                                Number of obs =    2063
                                                F( 15, 117) =    10.01
                                                Prob > F      =    0.0000
                                                R-squared    =    0.2629
Number of clusters (NUMECOLE) = 118          Root MSE     =    .8622
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.0676158	.1941648	0.35	0.728	-.3169174	.452149
FILLE	.0406226	.0392623	1.03	0.303	-.0371344	.1183795
AGE	-.0777349	.0167079	-4.65	0.000	-.110824	-.0446458
AISE	.1199178	.0847931	1.41	0.160	-.0480105	.2878462
REDOUBLE	-.1398848	.0672789	-2.08	0.040	-.2731271	-.0066424
LIV_FRMT	.1172152	.0744878	1.57	0.118	-.030304	.2647343
SERVICE	.016389	.0123801	1.32	0.188	-.0081291	.0409072
NIVCYCLB	.2654185	.1336601	1.99	0.049	.0007117	.5301254
DIPCYCLB	.3968031	.2398644	1.65	0.101	-.0782357	.8718419
FPIL	.158126	.1631092	0.97	0.334	-.1649031	.4811552
DOUBLFLX	-.2825171	.1654423	-1.71	0.090	-.6101669	.0451326
MULTIGRA	.3574221	.4264591	0.84	0.404	-.4871578	1.202002
TCLASSE	.0002938	.0041484	0.07	0.944	-.007922	.0085095
RURAL	-.1168871	.1312309	-0.89	0.375	-.3767831	.1430088
RES	.3676634	.1813501	2.03	0.045	.008509	.7268178
_cons	.4727767	.4002877	1.18	0.240	-.3199721	1.265525

La significativité de la variable RES confirme l'endogénéité de la variable STINI5FM.

Mali

Modèle sans score initial

Linear regression

Number of obs = 1761
 F(12, 133) = 3.37
 Prob > F = 0.0003
 R-squared = 0.0780
 Root MSE = .96352

Number of clusters (NUMECOLE) = 134

```
-----+-----
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	-.1909108	.0526122	-3.63	0.000	-.2949757	-.0868458
AGE	.0164456	.0245597	0.67	0.504	-.0321325	.0650237
AISE	-.0312737	.0885722	-0.35	0.725	-.2064661	.1439186
REDOUBLE	-.1704411	.0889467	-1.92	0.057	-.3463743	.0054921
LIV_FRMT	.1679447	.0896574	1.87	0.063	-.0093941	.3452835
SERVICE	-.0022271	.0078143	-0.28	0.776	-.0176834	.0132293
NIVCYCLB	.0510917	.1718096	0.30	0.767	-.2887411	.3909245
DIPCYCLB	(dropped)					
FPIL	-.0798464	.1862443	-0.43	0.669	-.4482304	.2885376
DOUBLFLX	-.4479566	.1402006	-3.20	0.002	-.7252679	-.1706453
MULTIGRA	.2497738	.2943544	0.85	0.398	-.3324478	.8319954
TCLASSE	-.0005299	.003427	-0.15	0.877	-.0073083	.0062485
RURAL	-.1117396	.1590759	-0.70	0.484	-.4263856	.2029065
_cons	.1884189	.4339767	0.43	0.665	-.6699702	1.046808

```
-----+-----
```

. vif

```
-----+-----
```

Variable	VIF	1/VIF
SERVICE	1.68	0.595084
FPIL	1.66	0.603291
RURAL	1.41	0.707891
TCLASSE	1.30	0.766995
MULTIGRA	1.24	0.805410
AISE	1.20	0.830474
AGE	1.17	0.851315
REDOUBLE	1.17	0.852012
NIVCYCLB	1.17	0.855611
DOUBLFLX	1.12	0.892005
LIV_FRMT	1.07	0.936847
FILLE	1.03	0.968306

```
-----+-----
```

Mean VIF | 1.27

Modèle avec score initial

Linear regression

Number of obs = 1761
 F(13, 133) = 16.31
 Prob > F = 0.0000
 R-squared = 0.3354
 Root MSE = .81827

Number of clusters (NUMECOLE) = 134

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.524604	.0485786	10.80	0.000	.4285174	.6206906
FILLE	-.0932625	.046028	-2.03	0.045	-.184304	-.0022209
AGE	-.0033742	.0206973	-0.16	0.871	-.0443127	.0375643
AISE	-.0119868	.0738618	-0.16	0.871	-.1580827	.134109
REDOUBLE	.0076078	.0661904	0.11	0.909	-.1233143	.1385298
LIV_FRMT	.0445763	.0665468	0.67	0.504	-.0870507	.1762032
SERVICE	.000955	.0055313	0.17	0.863	-.0099857	.0118957
NIVCYCLB	.0678626	.1366112	0.50	0.620	-.2023492	.3380743
DIPCYCLB	(dropped)					
FPIL	-.1377422	.134606	-1.02	0.308	-.4039877	.1285032
DOUBLFLX	-.3983114	.120252	-3.31	0.001	-.6361651	-.1604577
MULTIGRA	.0386273	.2549241	0.15	0.880	-.4656027	.5428573
TCLASSE	-.0012176	.0026951	-0.45	0.652	-.0065484	.0041132
RURAL	-.1139496	.1293203	-0.88	0.380	-.3697401	.141841
_cons	.3366995	.3491615	0.96	0.337	-.3539284	1.027327

. vif

Variable	VIF	1/VIF
SERVICE	1.68	0.593619
FPIL	1.66	0.602210
RURAL	1.41	0.707889
TCLASSE	1.30	0.766418
MULTIGRA	1.26	0.795875
AISE	1.20	0.830321
REDOUBLE	1.20	0.830639
AGE	1.18	0.848979
NIVCYCLB	1.17	0.855472
DOUBLFLX	1.12	0.890392
LIV_FRMT	1.08	0.929080
STINI5FM	1.07	0.935349
FILLE	1.04	0.959886
Mean VIF	1.26	

Tests de différence des coefficients des deux modèles :

```

. test FILLE=-0.19
      F( 1, 133) = 4.42
      Prob > F = 0.0375
. test AGE=0.02
      F( 1, 133) = 1.28
      Prob > F = 0.2608
. test AISE=-0.03
      F( 1, 133) = 0.06
      Prob > F = 0.8077
. test REDOUBLE=-0.17
      F( 1, 133) = 7.20
      Prob > F = 0.0082
. test LIV_FRMT=0.17
      F( 1, 133) = 3.55
      Prob > F = 0.0616
. test SERVICE=0
      F( 1, 133) = 0.03
      Prob > F = 0.8632
. test NIVCYCLB=0.05
      F( 1, 133) = 0.02
      Prob > F = 0.8962
. test DIPCYCLB=0
      Constraint 1 dropped
      F( 0, 133) = .
      Prob > F = .
. test FPIL=-0.08
      F( 1, 133) = 0.18
      Prob > F = 0.6686
. test DOUBLFLX=-0.45
      F( 1, 133) = 0.18
      Prob > F = 0.6680
. test MULTIGRA=0.25
      F( 1, 133) = 0.69
      Prob > F = 0.4085
. test TCLASSE=0
      F( 1, 133) = 0.20
      Prob > F = 0.6522
. test RURAL=-0.11
      F( 1, 133) = 0.00
      Prob > F = 0.9757

```

Test de l'endogénéité du score initial

```

. reg STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT DOMFRANC REDAN4 NBACTDOM
REPASMOINS/*
> */ SERVICE NIVCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL , cluster (NUMECOLE)

```

```

Linear regression
Number of obs = 1761
F( 16, 133) = 3.70
Prob > F = 0.0000
R-squared = 0.0745
Root MSE = .96644

Number of clusters (NUMECOLE) = 134

```

STINI5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	-.1167489	.0669581	-1.74	0.084	-.2491894	.0156917
AGE	.0379977	.0241541	1.57	0.118	-.0097783	.0857736
AISE	-.0499172	.0789878	-0.63	0.528	-.206152	.1063175
REDOUBLE	-.3819298	.0927652	-4.12	0.000	-.5654158	-.1984439
LIV_FRMT	.2183103	.0904426	2.41	0.017	.0394183	.3972024
DOMFRANC	.1135436	.1050663	1.08	0.282	-.0942735	.3213607
REDAN4	.1444989	.0688383	2.10	0.038	.0083395	.2806584
NBACTDOM	-.0301201	.0195185	-1.54	0.125	-.068727	.0084867
REPASMOINS	-.5805493	.2740029	-2.12	0.036	-1.122516	-.0385822
SERVICE	-.0054802	.0085158	-0.64	0.521	-.0223241	.0113636
NIVCYCLB	-.0171294	.1466936	-0.12	0.907	-.3072837	.2730249
FPIL	.1052941	.1818063	0.58	0.563	-.2543118	.4648999
DOUBLFLX	-.0844323	.1161358	-0.73	0.468	-.3141444	.1452799

MULTIGRA		.4384831	.2112943	2.08	0.040	.0205511	.8564151
TCLASSE		.0011365	.0030638	0.37	0.711	-.0049235	.0071965
RURAL		.0202539	.1516843	0.13	0.894	-.2797717	.3202795
_cons		-.2105696	.4241073	-0.50	0.620	-1.049438	.6282983

. predict RES, residual

. reg STFIN5FM STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT /*
> */ SERVICE NIVCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL RES , cluster (NUMECOLE)

Linear regression

Number of obs = 1761
F(14, 133) = 15.66
Prob > F = 0.0000
R-squared = 0.3401
Root MSE = .81557

Number of clusters (NUMECOLE) = 134

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	-.1676425	.2922124	-0.57	0.567	-.7456274	.4103424
FILLE	-.2221153	.0728535	-3.05	0.003	-.3662166	-.0780139
AGE	.0227793	.0232056	0.98	0.328	-.0231206	.0686791
AISE	-.0374371	.0749894	-0.50	0.618	-.1857632	.110889
REDOUBLE	-.2273384	.118324	-1.92	0.057	-.4613788	.0067019
LIV_FRMT	.2073683	.0845102	2.45	0.015	.0402104	.3745262
SERVICE	-.0032439	.0058557	-0.55	0.581	-.0148262	.0083384
NIVCYCLB	.0457324	.1318064	0.35	0.729	-.2149756	.3064403
FPIL	-.0613452	.1340815	-0.46	0.648	-.3265532	.2038629
DOUBLFLX	-.4638212	.1216992	-3.81	0.000	-.7045375	-.223105
MULTIGRA	.3172478	.2838359	1.12	0.266	-.2441686	.8786641
TCLASSE	-.0003101	.0027773	-0.11	0.911	-.0058035	.0051833
RURAL	-.1110334	.1281092	-0.87	0.388	-.3644285	.1423617
RES	.6995952	.2976075	2.35	0.020	.1109391	1.288251
_cons	.1410343	.3600509	0.39	0.696	-.5711325	.853201

Tchad

Modèle sans score initial

Linear regression

Number of obs = 1235
 F(13, 107) = 3.24
 Prob > F = 0.0004
 R-squared = 0.0969
 Root MSE = .95221

Number of clusters (NUMECOLE) = 108

```
-----+-----
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	-.0398506	.064203	-0.62	0.536	-.1671256	.0874244
AGE	.007833	.0198546	0.39	0.694	-.0315263	.0471924
AISE	-.1065185	.1263236	-0.84	0.401	-.3569404	.1439034
REDOUBLE	-.1046743	.0765619	-1.37	0.174	-.2564494	.0471008
LIV_FRMT	.2058263	.1084295	1.90	0.060	-.0091226	.4207752
SERVICE	-.0060197	.0104288	-0.58	0.565	-.0266935	.0146542
NIVCYCLB	.0407924	.1867096	0.22	0.827	-.3293376	.4109224
DIPCYCLB	.2290232	.2367208	0.97	0.335	-.2402483	.6982946
FPIL	-.2181416	.1897843	-1.15	0.253	-.5943669	.1580836
DOUBLFLX	-.7093371	.3322119	-2.14	0.035	-1.367908	-.0507657
MULTIGRA	.0427948	.1845759	0.23	0.817	-.3231053	.408695
TCLASSE	-.0083488	.0024733	-3.38	0.001	-.0132519	-.0034457
RURAL	-.4496603	.1830143	-2.46	0.016	-.8124648	-.0868559
_cons	.760071	.3721768	2.04	0.044	.0222739	1.497868

```
-----+-----
```

. vif

```
-----+-----
```

Variable	VIF	1/VIF
DIPCYCLB	2.86	0.349333
NIVCYCLB	2.10	0.476347
RURAL	1.78	0.561007
FPIL	1.54	0.648347
MULTIGRA	1.50	0.667178
SERVICE	1.48	0.673693
TCLASSE	1.33	0.752291
AISE	1.16	0.864776
DOUBLFLX	1.12	0.895288
AGE	1.06	0.939708
REDOUBLE	1.04	0.957622
LIV_FRMT	1.03	0.968899
FILLE	1.02	0.982164
Mean VIF	1.46	

```
-----+-----
```

Modèle avec score initial

Linear regression

Number of obs = 1235
 F(14, 107) = 15.36
 Prob > F = 0.0000
 R-squared = 0.3365
 Root MSE = .81649

Number of clusters (NUMECOLE) = 108

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5170697	.051861	9.97	0.000	.4142614	.6198781
FILLE	.0447402	.0579936	0.77	0.442	-.0702253	.1597058
AGE	.003092	.0173007	0.18	0.858	-.0312047	.0373887
AISE	-.1764949	.1039588	-1.70	0.092	-.3825812	.0295914
REDOUBLE	-.0130551	.0613874	-0.21	0.832	-.1347484	.1086382
LIV_FRMT	.1325597	.1087318	1.22	0.225	-.0829883	.3481078
SERVICE	-.0129957	.0100731	-1.29	0.200	-.0329644	.0069731
NIVCYCLB	.0073086	.1725429	0.04	0.966	-.3347375	.3493547
DIPCYCLB	.1333832	.2242819	0.59	0.553	-.3112294	.5779959
FPIL	-.1763772	.1607093	-1.10	0.275	-.4949645	.1422102
DOUBLFLX	-.630405	.3377357	-1.87	0.065	-1.299927	.0391166
MULTIGRA	.0737123	.1592309	0.46	0.644	-.2419444	.389369
TCLASSE	-.0074034	.0020851	-3.55	0.001	-.0115368	-.00327
RURAL	-.2400805	.1405765	-1.71	0.091	-.5187571	.0385961
_cons	.6382237	.3201158	1.99	0.049	.0036315	1.272816

. vif

Variable	VIF	1/VIF
DIPCYCLB	2.87	0.348178
NIVCYCLB	2.10	0.476099
RURAL	1.83	0.546923
FPIL	1.54	0.647629
MULTIGRA	1.50	0.666797
SERVICE	1.49	0.669414
TCLASSE	1.33	0.750586
AISE	1.16	0.863053
DOUBLFLX	1.12	0.894921
STINI5FM	1.08	0.928452
AGE	1.06	0.939425
REDOUBLE	1.05	0.951072
LIV_FRMT	1.03	0.967155
FILLE	1.03	0.975560
Mean VIF	1.44	

Tests de différence des coefficients des deux modèles :

```
. test FILLE=-0.04
      F( 1, 107) = 2.14
      Prob > F = 0.1469

. test AGE=0.01
      F( 1, 107) = 0.16
      Prob > F = 0.6905

. test AISE=-0.11
      F( 1, 107) = 0.41
      Prob > F = 0.5238

. test REDOUBLE=-0.10
      F( 1, 107) = 2.01
      Prob > F = 0.1596

. test LIV_FRMT=0.21
      F( 1, 107) = 0.51
      Prob > F = 0.4779

. test SERVICE=-0.01
      F( 1, 107) = 0.09
      Prob > F = 0.7667

. test NIVCYCLB=0.04
      F( 1, 107) = 0.04
      Prob > F = 0.8501
. test DIPCYCLB=0.23
      F( 1, 107) = 0.19
      Prob > F = 0.6675

. test FPIL=-0.22
      F( 1, 107) = 0.07
      Prob > F = 0.7866

. test DOUBLFLX=-0.71
      F( 1, 107) = 0.06
      Prob > F = 0.8141

. test MULTIGRA=0.04
      F( 1, 107) = 0.04
      Prob > F = 0.8327

. test TCLASSE=-0.01
      F( 1, 107) = 1.55
      Prob > F = 0.2157

. test RURAL=-0.45
      F( 1, 107) = 2.23
      Prob > F = 0.1383
```

Test de l'endogénéité du score initial

```
. reg STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT MERALPHA_ES AIDMAITR_ES /*
> */ SERVICE NIVCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL , cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1235
                                                F( 14, 107) =    3.88
                                                Prob > F       =    0.0000
                                                R-squared     =    0.0807
Number of clusters (NUMECOLE) = 108          Root MSE      =    .9443
```

STINI5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
FILLE	-.1807429	.0544806	-3.32	0.001	-.2887444	-.0727415
AGE	.0123358	.0177752	0.69	0.489	-.0229014	.047573
AISE	.1279065	.1314051	0.97	0.333	-.1325888	.3884019
REDOUBLE	-.1860047	.072685	-2.56	0.012	-.3300942	-.0419151
LIV_FRMT	.0946801	.1340273	0.71	0.481	-.1710135	.3603737
MERALPHA_ES	.2313908	.0909003	2.55	0.012	.0511915	.4115901
AIDMAITR_ES	.5148392	.2392372	2.15	0.034	.0405794	.989099
SERVICE	.0086736	.0140811	0.62	0.539	-.0192405	.0365876
NIVCYCLB	-.0376561	.1698037	-0.22	0.825	-.3742722	.29896
FPIL	-.050424	.1967937	-0.26	0.798	-.4405446	.3396966
DOUBLFLX	-.1631481	.1938807	-0.84	0.402	-.547494	.2211978
MULTIGRA	-.0193989	.1581379	-0.12	0.903	-.3328889	.2940911
TCLASSE	-.0017472	.0030394	-0.57	0.567	-.0077724	.004278
RURAL	-.4069002	.1835374	-2.22	0.029	-.7707416	-.0430588
_cons	.2697333	.4301563	0.63	0.532	-.5830013	1.122468

```
. predict RES, residual
```

```
. reg STFIN5FM STINI5FM FILLE AGE AISE REDOUBLE LIV_FRMT /*
> */ SERVICE NIVCYCLB FPIL /*
> */ DOUBLFLX MULTIGRA TCLASSE RURAL RES , cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1235
                                                F( 14, 107) =   15.89
                                                Prob > F       =    0.0000
                                                R-squared     =    0.3365
Number of clusters (NUMECOLE) = 108          Root MSE      =    .81652
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.8729278	.2237492	3.90	0.000	.4293712	1.316484
FILLE	.1038468	.0708415	1.47	0.146	-.0365882	.2442818
AGE	.0006192	.0186756	0.03	0.974	-.0364029	.0376414
AISE	-.2200418	.1024921	-2.15	0.034	-.4232204	-.0168631
REDOUBLE	.0493612	.0754132	0.65	0.514	-.1001367	.198859
LIV_FRMT	.0790912	.1081231	0.73	0.466	-.1352501	.2934326
SERVICE	-.019525	.0086195	-2.27	0.026	-.0366121	-.0024379
NIVCYCLB	-.0554588	.1335971	-0.42	0.679	-.3202994	.2093818
FPIL	-.1286154	.1391802	-0.92	0.358	-.4045239	.1472931
DOUBLFLX	-.5876655	.37743	-1.56	0.122	-1.335877	.1605455
MULTIGRA	.1010071	.1603043	0.63	0.530	-.2167774	.4187917
TCLASSE	-.0067762	.0021094	-3.21	0.002	-.0109578	-.0025947
RURAL	-.1040547	.1676221	-0.62	0.536	-.436346	.2282366
RES	-.3581773	.2055822	-1.74	0.084	-.7657201	.0493654
_cons	.5944299	.2813523	2.11	0.037	.0366818	1.152178

ANNEXE 1.2.4 : Comparaison des scores moyens de début et de fin d'année selon certaines caractéristiques de l'élève et du contexte

Ci-dessous le programme de création de la variable AISE.

```
/* Etape 1 : accès aux équipements de base */
/* l'eau courante*/
gen EAUCOUR= ROBINET==1| EAUTOIL==1

/* l'electricite */

tab ELECTRIC
tab EAUCOUR ELECTRIC

/* variable d'equipement de base */

gen EQPTBASE= EAUCOUR==1& ELECTRIC==1
tab EQPTBASE

/* Etape 2 : identification des plus aisés : NIVEAUVIE1 */

tab1 VIDEO TV VOITURE FRIGO

gen BIENSLUX=VIDEO+TV+VOITURE+FRIGO>1
tab BIENSLUX

gen NIVEAUVIE1=EQPTBASE==1&BIENSLUX==1
tab NIVEAUVIE1

/* Etape 3 : identification des catégories intermédiaires : NIVEAUVIE2
NIVEAUVIE3 */

gen BIENSFAMILL= VELO+MOBYLETT+RADIO+CUISIGAZ+VIDEO+TV+VOITURE+FRIGO
tab BIENSFAMILL

gen NIVEAUVIE2=BIENSFAMILL>3&NIVEAUVIE1==0
tab NIVEAUVIE2

gen NIVEAUVIE3=BIENSFAMILL==3&NIVEAUVIE1==0
tab NIVEAUVIE3

/* Etape 3 : identification de la catégorie la plus pauvre */

gen NIVEAUVIE4=NIVEAUVIE3==0&NIVEAUVIE1==0&NIVEAUVIE2==0
tab NIVEAUVIE4

/* regroupement des 2 catégories les plus favorisées*/

gen AISE=NIVEAUVIE1==1|NIVEAUVIE2==1
tab AISE
```

Tests de différence des scores moyens selon le niveau socio-économique de l'élève

Dans les tests de différence de moyennes qui suivent, le groupe 1 correspond à la catégorie des élèves les plus « riches » telle que cette dimension peut être appréhendée à travers la variable AISE.

Madagascar

Test de différence de moyenne pour le score initial

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	1657	.0353919	.0235546	.9588202	-.0108081	.0815918
1	406	.0171874	.0576806	1.162232	-.0962033	.1305781
combined	2063	.0318092	.0220571	1.00184	-.0114473	.0750658
diff		.0182045	.0554904		-.0906186	.1270275
diff = mean(0) - mean(1)					t =	0.3281
Ho: diff = 0					degrees of freedom =	2061
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.6286		Pr(T > t) = 0.7429		Pr(T > t) = 0.3714		

Test de différence de moyenne pour le score final

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	1657	-.0501306	.0240031	.9770768	-.0972102	-.0030509
1	406	.2161838	.0529197	1.066303	.1121521	.3202154
combined	2063	.0022803	.0220303	1.000621	-.0409236	.0454842
diff		-.2663143	.055113		-.3743973	-.1582314
diff = mean(0) - mean(1)					t =	-4.8322
Ho: diff = 0					degrees of freedom =	2061
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.0000		Pr(T > t) = 0.0000		Pr(T > t) = 1.0000		

Les élèves identifiés comme appartenant aux familles les plus aisées ont de meilleurs résultats en fin d'année que les autres élèves alors que leur niveau était comparable en début d'année. Le coefficient positif du modèle est donc cohérent avec ce que l'on observe ici.

Mali

Test de différence de moyenne pour le score initial

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	1428	.0139439	.026847	1.01452	-.03872	.0666078
1	333	-.0597953	.0511979	.9342736	-.1605085	.0409178
combined	1761	3.19e-10	.0238298	1	-.0467377	.0467377
diff		.0737392	.0608465		-.0455999	.1930783
diff = mean(0) - mean(1)					t =	1.2119
Ho: diff = 0					degrees of freedom =	1759
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.8871		Pr(T > t) = 0.2257		Pr(T > t) = 0.1129		

Test de différence de moyenne pour le score final

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	1428	.024838	.0266454	1.0069	-.0274304	.0771063
1	333	-.1065124	.0528316	.964087	-.2104393	-.0025854
combined	1761	-6.84e-10	.0238298	1	-.0467377	.0467377
diff		.1313503	.0607913		.0121196	.2505811
diff = mean(0) - mean(1)					t =	2.1607
Ho: diff = 0					degrees of freedom =	1759
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.9846		Pr(T > t) = 0.0309		Pr(T > t) = 0.0154		

Au seuil de 5 %, les élèves de la catégorie AISE obtiennent un score légèrement plus faible (13 % d'écart-type) que leurs camarades.

Tchad

Test de différence de moyenne pour le score initial

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	1078	-.0250819	.0298076	.97867	-.0835694	.0334057
1	159	.2945421	.0766245	.9661984	.1432016	.4458826
combined	1237	.0160016	.0279359	.9825335	-.0388054	.0708086
diff		-.319624	.0830058		-.482472	-.156776
diff = mean(0) - mean(1)					t =	-3.8506
Ho: diff = 0					degrees of freedom =	1235

Test de différence de moyennes selon que la classe fonctionne en double flux ou non à Madagascar

En début d'année :

```
. ttest STINI5FM, by ( DOUBLFLX)

Two-sample t test with equal variances
-----+-----
   Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
       0 |     1913   .0328296   .0231856   1.014089   -.0126422   .0783013
       1 |       150   .0187966   .0679892   .8326943   -.115551   .1531441
-----+-----
combined |     2063   .0318092   .0220571   1.00184   -.0114473   .0750658
-----+-----
   diff |           .014033   .0849664           -.152596   .180662
-----+-----
   diff = mean(0) - mean(1)                                t =    0.1652
Ho: diff = 0                                               degrees of freedom =    2061

   Ha: diff < 0                Ha: diff != 0                Ha: diff > 0
Pr(T < t) = 0.5656            Pr(|T| > |t|) = 0.8688            Pr(T > t) = 0.4344
```

En fin d'année :

```
. ttest STFIN5FM, by ( DOUBLFLX)

Two-sample t test with equal variances
-----+-----
   Group |      Obs      Mean   Std. Err.   Std. Dev.   [95% Conf. Interval]
-----+-----
       0 |     1913   .0320175   .0229424   1.003453   -.0129773   .0770124
       1 |       150  -.3769688   .0720645   .8826066   -.5193693  -.2345684
-----+-----
combined |     2063   .0022803   .0220303   1.000621   -.0409236   .0454842
-----+-----
   diff |           .4089864   .0843841           .2434994   .5744733
-----+-----
   diff = mean(0) - mean(1)                                t =    4.8467
Ho: diff = 0                                               degrees of freedom =    2061

   Ha: diff < 0                Ha: diff != 0                Ha: diff > 0
Pr(T < t) = 1.0000            Pr(|T| > |t|) = 0.0000            Pr(T > t) = 0.0000
```

Les tests montrent qu'alors qu'il n'y a pas de différence significative en début d'année, un écart important existe en fin d'année en défaveur des élèves scolarisés en double flux.

ANNEXE 1.2.5 : Création de la variable basée sur l'évolution du classement de l'élève aux tests standardisés sous STATA

Exemple de Madagascar :

```
Drop if STINI5FM==.| STFIN5FM==.
egen RANG1=rank(STINI5FM) if STINI5FM, field
egen RANG2=rank(STFIN5FM), field
gen PROGRESS=RANG1-RANG2
sum RANG1 RAN2 PROGRESS
```

Variable	Obs	Mean	Std. Dev.	Min	Max
RANG1	2063	1002.266	597.922	1	2054
RANG2	2063	1028.904	595.2975	1	2060
PROGRESS	2063	-26.63839	617.0476	-2056	1912

On constate que la variable PROGRESS ainsi créée comporte des valeurs extrêmes qui pourraient avoir une influence sur les analyses. Un élève qui a perdu 2056 places entre le début et la fin de l'année a connu une évolution extrême. Celle-ci peut être liée à des conditions exceptionnelles, par exemple à une absence longue de l'élève liée à des problèmes de santé, ou encore à une erreur dans le traitement de l'information. En tout état de cause, ce type de cas extrême ne devrait pas perturber les analyses. Il est donc utile de procéder à un cylindrage pour éviter ces valeurs extrêmes et leur influence excessive sur les estimations. Pour cela, on élimine les extrémités de la distribution ainsi que les scores nuls qui paraissent un peu suspects compte tenu de la forme des tests (une bonne partie des tests sont des QCM).

```
gen VALEXTR=PROGRES>2.8|PROGRES<-2.8
replace VALEXTR=1 if SINI5FM==0| SFIN5FM==0
tab VALEXTR
```

On remarque cependant dans le tableau suivant que l'élimination de ces valeurs extrêmes n'a quasiment pas d'incidence sur les estimations effectuées.

Comparaisons des modèles de rang avec et sans valeurs extrêmes (VE)

	Madagascar		Mali		Tchad	
	Avec VE	Sans VE	Avec VE	Sans VE	Avec VE	Sans VE
FILLE	0.04 (0.05)	0.04 (0.05)	-0.03 (0.06)	-0.05 (0.06)	0.12* (0.07)	0.13* (0.06)
AGE	-0.08*** (0.02)	-0.08*** (0.02)	-0.02 (0.02)	-0.02 (0.02)	-0.01 (0.02)	-0.01 (0.02)
AISE	0.22** (0.10)	0.20** (0.10)	0.01 (0.08)	0.05 (0.08)	-0.29** (0.11)	-0.28** (0.11)
REDOUBLE	-0.00 (0.06)	0.00 (0.06)	0.11 (0.07)	0.11 (0.07)	0.05 (0.07)	0.02 (0.07)
LIV FRMT	0.05 (0.08)	0.03 (0.08)	-0.02 (0.07)	-0.03 (0.07)	0.10 (0.16)	0.08 (0.16)
SERVICE	0.02 (0.01)	0.02 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)
NIVCYCLB	-0.13 (0.14)	-0.08 (0.13)	0.09 (0.15)	0.11 (0.14)	-0.04 (0.24)	-0.05 (0.23)
DIPCYCLB	0.04 (0.26)	0.06 (0.26)	/	/	0.10 (0.28)	0.05 (0.27)
FPIL	-0.14 (0.16)	-0.10 (0.15)	-0.14 (0.13)	-0.13 (0.13)	-0.21 (0.18)	-0.16 (0.16)
DOUBLFLX	-0.30 (0.19)	-0.28 (0.18)	-0.34** (0.13)	-0.30** (0.13)	-0.62 (0.48)	-0.63 (0.47)
MULTIGRA	0.80* (0.43)	0.79* (0.43)	-0.20 (0.20)	-0.14 (0.19)	0.12 (0.18)	0.07 (0.17)
TCLASSE	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.00)	-0.00 (0.00)	-0.01** (0.00)	-0.01** (0.00)
RURAL	0.21 (0.13)	0.19 (0.13)	-0.07 (0.14)	-0.11 (0.12)	-0.04 (0.14)	-0.03 (0.14)
Observations	2063	2044	1761	1740	1235	1226
R ²	0.08	0.08	0.03	0.03	0.08	0.08

ANNEXE 1.2.6 : Endogénéité du score de début d'année dans le modèle à valeur ajoutée

Explication du score de début d'année par les variables du modèle explicatif du score final

	Madagascar	Mali	Tchad
L'élève est une fille	0.01 (0.05)	-0.19*** (0.05)	-0.16*** (0.06)
Age	0.00 (0.02)	0.04 (0.02)	0.01 (0.02)
Le niveau de vie de la famille de l'élève est parmi les plus élevés	-0.14 (0.10)	-0.04 (0.08)	0.14 (0.13)
L'élève a redoublé au moins une fois	-0.17*** (0.06)	-0.34*** (0.09)	-0.18** (0.07)
L'élève possède les manuels de français et mathématiques	0.05 (0.09)	0.24*** (0.09)	0.14 (0.13)
Ancienneté du maître	-0.01 (0.01)	-0.01 (0.01)	0.01 (0.02)
Le maître a fréquenté le second cycle secondaire	0.39*** (0.14)	-0.03 (0.15)	0.06 (0.20)
Le maître est titulaire du BAC	0.27 (0.34)	/	0.18 (0.26)
Le maître a suivi une formation professionnelle initiale longue (1 an et plus)	0.34* (0.17)	0.11 (0.18)	-0.08 (0.22)
La classe fonctionne en double flux	0.07 (0.25)	-0.09 (0.12)	-0.15 (0.20)
La classe est multigrade	-0.54 (0.38)	0.40* (0.21)	-0.06 (0.16)
Nombre d'élèves dans la classe	0.01 (0.01)	0.00 (0.00)	-0.00 (0.00)
L'école est en zone rurale	-0.37*** (0.13)	0.00 (0.15)	-0.41** (0.19)
Observations	2063	1761	1235
R ²	0.12	0.06	0.07

Erreurs types robustes entre parenthèses

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

Les coefficients significatifs des différentes variables ainsi que l'importance des R² confirment le caractère endogène du score initial.

ANNEXE 1.2.7 : Comparaison des modèles MCO avec les modèles hiérarchiques

Les traitements ont été réalisés sous STATA avec les commandes regress et xtmixed.

Madagascar

Le modèle MCO (regress option cluster)

```

Linear regression                               Number of obs =      2063
                                                F( 14, 117) =      10.14
                                                Prob > F      =      0.0000
                                                R-squared     =      0.2611
                                                Root MSE     =      .86306

Number of clusters (NUMECOLES) = 118

```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.4295877	.0596082	7.21	0.000	.3115367	.5476387
FILLE	.0381097	.0393565	0.97	0.335	-.0398339	.1160533
AGE	-.0794395	.0166504	-4.77	0.000	-.1124147	-.0464643
AISE	.1703945	.0870909	1.96	0.053	-.0020845	.3428735
REDOUBLE	-.078179	.0564579	-1.38	0.169	-.1899909	.0336329
LIV_FRMT	.0995081	.0740518	1.34	0.182	-.0471476	.2461638
SERVICE	.0190673	.012314	1.55	0.124	-.00532	.0434546
NIVCYCLB	.1240313	.1309885	0.95	0.346	-.1353845	.3834471
DIPCYCLB	.3006303	.2351424	1.28	0.204	-.165057	.7663176
FPIL	.0362586	.1423333	0.25	0.799	-.2456251	.3181423
DOUBLFLX	-.3069038	.1628365	-1.88	0.062	-.6293929	.0155854
MULTIGRA	.5531214	.4246498	1.30	0.195	-.2878753	1.394118
TCLASSE	-.0024937	.0042144	-0.59	0.555	-.01084	.0058526
RURAL	.0160428	.1223917	0.13	0.896	-.2263476	.2584333
_cons	.5282562	.4001179	1.32	0.189	-.2641564	1.320669

Le modèle hiérarchique (xtmixed option mle)

```

Mixed-effects ML regression          Number of obs      =      2063
Group variable: NUMECOLE            Number of groups   =       118

                                     Obs per group: min =        1
                                     avg =      17.5
                                     max =       20

                                     Wald chi2(14)      =      675.80
                                     Prob > chi2       =      0.0000

Log likelihood = -1888.8732

```

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
STINI5FM	.4723734	.0192649	24.52	0.000	.4346148	.510132
FILLE	.0237812	.0254165	0.94	0.349	-.0260342	.0735967
AGE	-.0378773	.0093436	-4.05	0.000	-.0561904	-.0195641
AISE	.0003974	.0362298	0.01	0.991	-.0706118	.0714066
REDOUBLE	-.0981815	.0291249	-3.37	0.001	-.1552654	-.0410977
LIV_FRMT	.0464418	.0305107	1.52	0.128	-.013358	.1062416
SERVICE	.0182201	.0125579	1.45	0.147	-.0063928	.0428331
NIVCYCLB	.0225281	.1429339	0.16	0.875	-.2576172	.3026734
DIPCYCLB	.2316116	.3106516	0.75	0.456	-.3772544	.8404776
FPIL	.0473523	.1693524	0.28	0.780	-.2845723	.3792769
DOUBLFLX	-.3561586	.2527437	-1.41	0.159	-.8515272	.13921
MULTIGRA	.6803912	.3173445	2.14	0.032	.0584074	1.302375
TCLASSE	-.0018268	.0047775	-0.38	0.702	-.0111905	.0075369
RURAL	.0379744	.1343095	0.28	0.777	-.2252675	.3012162
_cons	.126699	.3546447	0.36	0.721	-.5683919	.8217898

```

-----
Random-effects Parameters | Estimate Std. Err. [95% Conf. Interval]
-----+-----
NUMECOLE: Identity
      sd(_cons) | .6996096 .047905 .6117454 .8000936
-----+-----
      sd(Residual) | .5492822 .0088131 .5322777 .56683
-----
LR test vs. linear regression: chibar2(01) = 1454.08 Prob >= chibar2 = 0.0000

```

Mali

Le modèle MCO (regress option cluster)

```

Linear regression          Number of obs      =      1761
                           F( 13, 133) =      16.31
                           Prob > F      =      0.0000
                           R-squared     =      0.3354
                           Root MSE   =      .81827

Number of clusters (NUMECOLE) = 134

```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.524604	.0485786	10.80	0.000	.4285174	.6206906

FILLE	-.0932625	.046028	-2.03	0.045	-.184304	-.0022209
AGE	-.0033742	.0206973	-0.16	0.871	-.0443127	.0375643
AISE	-.0119868	.0738618	-0.16	0.871	-.1580827	.134109
REDOUBLE	.0076078	.0661904	0.11	0.909	-.1233143	.1385298
LIV_FRMT	.0445763	.0665468	0.67	0.504	-.0870507	.1762032
SERVICE	.000955	.0055313	0.17	0.863	-.0099857	.0118957
NIVCYCLB	.0678626	.1366112	0.50	0.620	-.2023492	.3380743
DIPCYCLB	(dropped)					
FPIL	-.1377422	.134606	-1.02	0.308	-.4039877	.1285032
DOUBLFLX	-.3983114	.120252	-3.31	0.001	-.6361651	-.1604577
MULTIGRA	.0386273	.2549241	0.15	0.880	-.4656027	.5428573
TCLASSE	-.0012176	.0026951	-0.45	0.652	-.0065484	.0041132
RURAL	-.1139496	.1293203	-0.88	0.380	-.3697401	.141841
_cons	.3366995	.3491615	0.96	0.337	-.3539284	1.027327

Le modèle hiérarchique (xtmixed option mle)

Computing standard errors:

```

Mixed-effects ML regression          Number of obs      =      1761
Group variable: NUMECOLE            Number of groups   =       134

                                     Obs per group: min =         8
                                     avg =          13.1
                                     max =          15

Log likelihood = -1470.224          Wald chi2(13)      =       746.63
                                     Prob > chi2         =       0.0000

```

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
STINI5FM	.4261568	.0168653	25.27	0.000	.3931015	.4592121
FILLE	-.0943497	.0253521	-3.72	0.000	-.1440389	-.0446605
AGE	-.0194922	.0099908	-1.97	0.049	-.0389115	-.0000728
AISE	.0041205	.0370253	0.11	0.911	-.0684478	.0766887
REDOUBLE	-.0223141	.0293853	-0.76	0.448	-.0799082	.03528
LIV_FRMT	.0291252	.0349702	0.83	0.405	-.0394152	.0976656
SERVICE	.0011634	.0072014	0.16	0.872	-.0129511	.015278
NIVCYCLB	.0744892	.1465106	0.51	0.611	-.2126663	.3616447
FPIL	-.1255977	.1536715	-0.82	0.414	-.4267882	.1755929
DOUBLFLX	-.4010428	.1313231	-3.05	0.002	-.6584313	-.1436543
MULTIGRA	.0721852	.2171598	0.33	0.740	-.3534402	.4978106
TCLASSE	-.0012898	.0028482	-0.45	0.651	-.0068721	.0042925
RURAL	-.110158	.1317214	-0.84	0.403	-.3683272	.1480112
_cons	.5295344	.2798207	1.89	0.058	-.0189041	1.077973

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLE: Identity				
sd(_cons)	.6509642	.04159	.5743465	.7378026
sd(Residual)	.4944558	.0086688	.477754	.5117416

LR test vs. linear regression: chibar2(01) = 1336.62 Prob >= chibar2 = 0.0000

Tchad

Le modèle MCO (regress option cluster)

Linear regression

Number of obs = 1235
 F(14, 107) = 15.36
 Prob > F = 0.0000
 R-squared = 0.3365
 Root MSE = .81649

Number of clusters (NUMECOLE) = 108

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5170697	.051861	9.97	0.000	.4142614	.6198781
FILLE	.0447402	.0579936	0.77	0.442	-.0702253	.1597058
AGE	.003092	.0173007	0.18	0.858	-.0312047	.0373887
AISE	-.1764949	.1039588	-1.70	0.092	-.3825812	.0295914
REDOUBLE	-.0130551	.0613874	-0.21	0.832	-.1347484	.1086382
LIV_FRMT	.1325597	.1087318	1.22	0.225	-.0829883	.3481078
SERVICE	-.0129957	.0100731	-1.29	0.200	-.0329644	.0069731
NIVCYCLB	.0073086	.1725429	0.04	0.966	-.3347375	.3493547
DIPCYCLB	.1333832	.2242819	0.59	0.553	-.3112294	.5779959
FPIL	-.1763772	.1607093	-1.10	0.275	-.4949645	.1422102
DOUBLFLX	-.630405	.3377357	-1.87	0.065	-1.299927	.0391166
MULTIGRA	.0737123	.1592309	0.46	0.644	-.2419444	.389369
TCLASSE	-.0074034	.0020851	-3.55	0.001	-.0115368	-.00327
RURAL	-.2400805	.1405765	-1.71	0.091	-.5187571	.0385961
_cons	.6382237	.3201158	1.99	0.049	.0036315	1.272816

Le modèle hiérarchique (xtmixed option mle)

Computing standard errors:

```
Mixed-effects ML regression      Number of obs      =      1235
Group variable: NUMECOLES       Number of groups   =       108

                                Obs per group: min =        3
                                avg =       11.4
                                max =       15

                                Wald chi2(14)    =     513.41
                                Prob > chi2     =     0.0000

Log likelihood = -1104.9797
```

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
STINI5FM	.552582	.0257228	21.48	0.000	.5021663	.6029977
FILLE	-.0570817	.0339397	-1.68	0.093	-.1236023	.009439
AGE	-.0133849	.0092384	-1.45	0.147	-.0314918	.004722
AISE	-.0415344	.0576151	-0.72	0.471	-.1544579	.0713891
REDOUBLE	-.006462	.036514	-0.18	0.860	-.0780281	.0651041
LIV_FRMT	.0131796	.0568727	0.23	0.817	-.0982888	.1246481
SERVICE	-.0121394	.011467	-1.06	0.290	-.0346144	.0103355
NIVCYCLB	.0986834	.1923392	0.51	0.608	-.2782945	.4756613
DIPCYCLB	.1075611	.2169689	0.50	0.620	-.3176901	.5328123
FPIL	-.120539	.164963	-0.73	0.465	-.4438605	.2027824
DOUBLFLX	-.5980984	.5021985	-1.19	0.234	-1.582389	.3861925
MULTIGRA	.1337573	.1683097	0.79	0.427	-.1961236	.4636382
TCLASSE	-.007644	.0026111	-2.93	0.003	-.0127617	-.0025262
RURAL	-.2226064	.1709037	-1.30	0.193	-.5575715	.1123587
_cons	.7872394	.3138918	2.51	0.012	.1720229	1.402456

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLES: Identity				
sd(_cons)	.6456538	.0472773	.5593342	.7452946
sd(Residual)	.5218096	.0110048	.5006804	.5438305

LR test vs. linear regression: chibar2(01) = 778.96 Prob >= chibar2 = 0.0000

ANNEXE 1.2.8 : les modèles hiérarchiques avec et sans score initial

	Madagascar		Mali		Tchad	
	M1	M2	M3	M4	M5	M6
L'élève						
est une fille	0.02 (0.03)	0.02 (0.03)	-0.09*** (0.03)	-0.17*** (0.03)	-0.06* (0.03)	-0.15*** (0.04)
Age	-0.04*** (0.01)	-0.04*** (0.01)	-0.02** (0.01)	-0.02 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Appartient à une famille de niveau de vie élevé	0.00 (0.04)	0.01 (0.04)	0.00 (0.04)	-0.00 (0.04)	-0.04 (0.06)	-0.03 (0.07)
a redoublé au moins une fois	-0.10*** (0.03)	-0.14*** (0.03)	-0.02 (0.03)	-0.13*** (0.03)	-0.01 (0.04)	-0.09** (0.04)
possède les manuels de français et mathématiques	0.05 (0.03)	0.09** (0.03)	0.03 (0.03)	0.09** (0.04)	0.01 (0.06)	0.04 (0.07)
Le maître						
Années d'ancienneté	0.02 (0.01)	0.02 (0.01)	0.00 (0.01)	-0.00 (0.01)	-0.01 (0.01)	-0.01 (0.01)
niveau second cycle secondaire	0.02 (0.14)	0.20 (0.15)	0.07 (0.15)	0.06 (0.17)	0.10 (0.19)	0.15 (0.23)
titulaire du BAC	0.23 (0.31)	0.40 (0.34)			0.11 (0.22)	0.22 (0.26)
formation professionnelle initiale ≥ 1 an	0.05 (0.17)	0.19 (0.18)	-0.13 (0.15)	-0.09 (0.18)	-0.12 (0.16)	-0.14 (0.20)
La classe						
double flux	-0.36 (0.25)	-0.29 (0.27)	-0.40*** (0.13)	-0.44*** (0.15)	-0.60 (0.50)	-0.71 (0.60)
multigrade	0.68** (0.32)	0.46 (0.34)	0.07 (0.22)	0.24 (0.25)	0.13 (0.17)	0.10 (0.20)
Nombre d'élèves	-0.00 (0.00)	0.00 (0.01)	-0.00 (0.00)	-0.00 (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
L'école						
est en zone rurale	0.04 (0.13)	-0.14 (0.14)	-0.11 (0.13)	-0.10 (0.15)	-0.22 (0.17)	-0.42** (0.20)
Score de début d'année (français et mathématiques)	0.47*** (0.02)	/	0.43*** (0.02)	/	0.55*** (0.03)	/
Observations	2063	2063	1761	1761	1235	1235
Nombre de classes	118	118	134	134	108	108

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

On observe que les résultats changent relativement peu quand on retire le score initial du modèle, comme cela a été observé avec les MCO robustes précédemment. En outre, les changements s'interprètent assez aisément quand on considère le processus cumulatif des apprentissages. Le résultat relatif à l'âge est peut-être le moins intuitif, il est cependant cohérent avec les tendances observées dans les études du PASEC montrant qu'être plus âgé en début de cycle constitue parfois un avantage alors que c'est plutôt l'inverse en fin de cycle.

ANNEXE 2.1.1 : Principaux paramètres du cadre indicatif pour l'initiative de mise en œuvre accélérée de l'Education pour tous

Les valeurs de référence mentionnées dans le cadre indicatif correspondent à des valeurs moyennes observées dans des pays qui ont réussi la scolarisation primaire universelle.

Indicateurs	Valeurs de référence
Mobilisation des ressources Revenus publics générés à l'intérieur du pays comme % du PIB	14-18
Part de l'éducation dans le budget (%)	20
Part de l'enseignement primaire dans le budget de l'éducation (%)	42-64
Ratio enseignant/élèves dans les écoles primaires sur financement public	1:40
Salaire annuel moyen des enseignants des écoles primaire en PIB/tête	3,5
Dépenses récurrentes autres que la rémunération des enseignants en % des dépenses récurrentes totales de l'enseignement primaire	33
Heures d'instruction annuelles	850-1000
% d'élèves inscrits dans les écoles primaires exclusivement privées	10 ou moins

Source : Education pour tous - Initiative de mise en œuvre accélérée (2004)

ANNEXE 2.1.2 : Proportion d'élèves scolarisés dans le privé à l'école primaire dans 12 pays d'Afrique francophone

PAYS	% d'élèves dans le privé à l'école primaire
Bénin (2003/2004)	11 %
Burkina Faso (2003/2004)	13 %
Cameroun (2002/2003)	23 %
Congo (2003/2004)	25 %
Côte d'Ivoire (2002/2003)	11 %
Guinée (2003/2004)	20 %
Madagascar (2002/2003)	20 %
Mali (2003/2004)	18 %
Niger (2002/2003)	4 %
Sénégal (2002/2003)	13 %
Tchad (2003/2004)	27 %
Togo (2001)	41 %
Moyenne	18,8 %

Source : Pôle de Dakar (2005)

On observe que la part des élèves scolarisés dans l'enseignement privé est assez élevée, presque 19 %, dans les pays considérés. Il y a aussi une variabilité importante selon les pays puisque cette part oscille entre 11 % au Bénin et en Côte d'Ivoire et 41 % au Togo. En tout état de cause, on ne saurait négliger la part de l'enseignement privé dans la scolarisation des élèves à l'école primaire en Afrique.

ANNEXE 2.2.1 : Les résultats des variables maître dans les modèles avec et sans score initial dans l'étude de Michaelowa et Wechtler (2006)

Résultats pour les variables maître des modèles
de Michaelowa et Wechtler (2006) en Français

Variables	2 ^{ème} année		5 ^{ème} année	
	Sans score initial	Avec score initial	Sans score initial	Avec score initial
Formation académique (0=primaire non achevé- 6 : au moins trois années de supérieur)	-0,42/-0,44	-0,19/-0,25	0,28/0,30	0,03/0,02
L'enseignant a suivi une formation professionnelle initiale	1,03/1,37	0,77/0,99	-0,28/-0,39	0,06/0,09
Nombre moyen de formations continues par an au cours des 5 dernières années	1,03/1,37	0,77/0,99	0,98**/1,09***	0,64*/0,67**
Années d'ancienneté	-0,04/-0,04	-0,04/-0,04	0,14***/0,14***	0,11***/0,11**
Enseignant non fonctionnaire	0/-0,08	-1,69/-2,27*	2,1**/2,55***	1,14/1,3
Le maître est une femme	0,62/0,77	0,65 / 0,76	0,41/0,70	-0,10/0,28

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

Résultats pour les variables maître des modèles
de Michaelowa et Wechtler (2006) en Mathématiques

Variables	2 ^{ème} année		5 ^{ème} année	
	Sans score initial	Avec score initial	Sans score initial	Avec score initial
Formation académique (0=primaire non achevé- 6 : au moins trois années de supérieur)	-0,26 / -0,26	-0,01/0,00	0,42/0,49*	0,34/0,42
L'enseignant a suivi une formation professionnelle initiale	0,14/0,23	-0,27/-0,40	-0,59/-0,60	-0,64/-0,69
Nombre moyen de formations continues par an au cours des 5 dernières années	0,46/0,48	0,49/0,52	0,52/0,61	0,37/0,40
Années d'ancienneté	-0,08/-0,08	-0,08/-0,07	0,10**/0,11**	0,09**/0,10**
Enseignant non fonctionnaire	-0,45/-0,54	-1,03/-1,24	1,44/1,99*	0,70/0,98
Le maître est une femme	1,70*/1,78*	1,85**/1,87**	-0,78/-0,99	-1,54**/-1,99**

* significatif à 10 %; ** significatif à 5 %; *** significatif à 1 %

ANNEXE 3.1.1 : Echantillon théorique de l'évaluation PASEC Mali

Représentation proportionnelle des enseignants contractuels par région

Régions	Contractuels	Nb contractuels région / Nb contractuels total	Echantillon Théorique de contractuels	Echantillon complet
Kayes	591	15.6 %	11	22
Koulikoro	664	17.5 %	12	24
Sikasso	793	20.9 %	15	30
Ségou	458	12.1 %	9	18
Mopti	325	8.6 %	6	12
Tombouctou	195	5.2 %	4	8
Gao	176	4.7 %	3	6
Bamako	561	14.8 %	10	20
Kidal	24	0.6 %	0	0
Total	3787	100 %	70	140

La région de Kidal n'a pas été retenue dans l'échantillon final faute d'un effectif suffisant d'enseignants contractuels.

A partir de la répartition par région, on procède à une répartition des enseignants par cellule d'animation pédagogique (CAP) en fonction des effectifs de contractuels.

On constate dans le tableau qui suit que seuls 140 enseignants, soit la moitié de l'échantillon théorique, sont choisis en fonction de leur statut, ce qui permet de sélectionner 140 écoles. Les autres enseignants seront choisis dans ces 140 écoles sans qu'on ne maîtrise au départ leur statut. Si l'école a été sélectionnée pour un maître de 2^{ème} année, le maître présent en 5^{ème} année sera aussi enquêté. Chaque enquêteur avait ainsi un profil d'enseignant basé sur le statut pour identifier une école. Ainsi, seulement 35 enseignants de chaque catégorie, contractuel et fonctionnaire, sont sûrs d'être enquêtés à chacun des niveaux visés.

Au final, seules deux écoles ont été « perdues », une dans la CAP de Bafoulabé, l'autre dans la CAP de Nara.

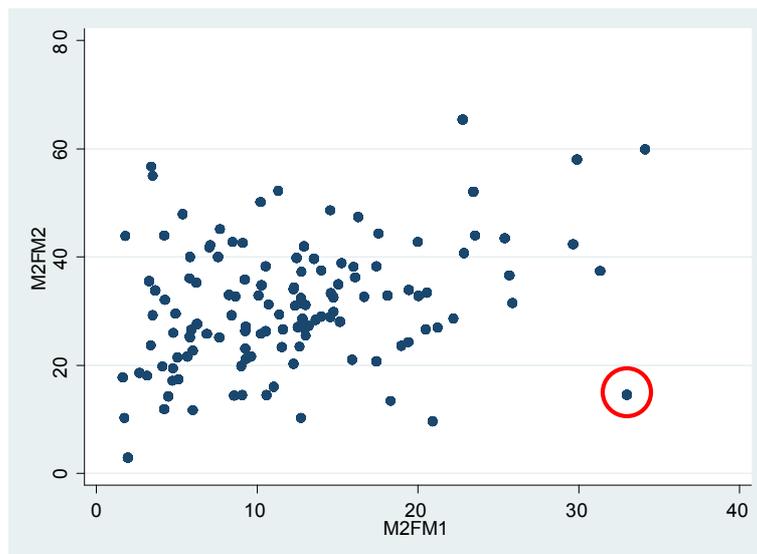
Répartition des enseignants par CAP

Région	CAP	contractuels 2A	contractuels 5A	titulaires 2A	titulaires 5A	Nb d'écoles
Kayes	Bafoulabé	1	1	1	1	4
	Kayes RD	1	1	1	1	4
	Kayes RG	1	1	1	1	4
	Kita	2	1	2	1	6
	Nioro	0	1	0	1	2
	Yélimané	0	1	0	1	2
	Total	5	6	5	6	22
Koulikoro	Kati	2	1	2	1	6
	Koulikoro	1	1	1	1	4
	Dioila	1	2	1	2	6
	Banamba	1	1	1	1	4
	Kolokani	1	1	1	1	4
	Nara	1	1	1	1	4
	Baguineda	2	2	2	2	8
	Total	9	9	9	9	36
Sikasso	Sikasso 1	1	1	1	1	4
	Sikasso 2	1	1	1	1	4
	Bougouni 1	1	1	1	1	4
	Bougouni 2	1	1	1	1	4
	Koutiala	2	2	2	2	8
	Total	6	6	6	6	24
Ségou	Bla	1	1	1	1	4
	Niono	1	1	1	1	4
	San	2	2	2	2	8
	Total	4	4	4	4	16
Mopti	Sévaré	1	1	1	1	4
	Bandiagara	1	1	1	1	4
	Koro	1	1	1	1	4
	Total	3	3	3	3	12
Tombouctou	Tombouctou	1	1	1	1	4
	Diré	1	1	1	1	4
	Total	2	2	2	2	8
Gao	Gao	2	1	2	1	6
	Total	2	1	2	1	6
Bamako	CAP1	0	1	0	1	2
	CAP2	0	1	0	1	2
	CAP3	1	0	1	0	2
	CAP4	0	1	0	1	2
	CAP5	1	1	1	1	4
	CAP6	1	0	1	0	2
	CAP7	1	0	1	0	2
	Total	4	4	4	4	16
Total général		35	35	35	35	140

ANNEXE 3.1.2 : Traitement des valeurs extrêmes

En 2^{ème} année

On recherche les classes avec des évolutions extrêmes entre prétest et post test en 2^{ème} année. Le graphique montre la relation entre les scores moyens d'une classe au prétest (M2FM1) et au post test (M2FM2).



On observe une classe avec une évolution très atypique, il s'agit de l'école 78, il s'agit d'un enseignant contractuel. Conserver cette classe pourrait perturber les analyses.

Pour éviter que les analyses soient perturbées par des cas extrêmes d'évolution du score entre le début et la fin de l'année, qui peuvent être imputés à des situations exceptionnelles, comme la maladie ou encore une erreur d'enquête où deux élèves différents auraient été enquêtés entre le début et la fin de l'année sous le même identifiant¹⁰², ceux-ci ont été identifiés afin d'être exclus des analyses.

Pour cela, une variable a été créée afin de mesurer les évolutions entre début et fin d'année. Il s'agit de la valeur absolue de la différence entre les scores standardisés de début et fin d'année (gen DIFF=abs(STFIN2FM- STINI2FM)). Au total, 26 observations, soit 1,5 % des

¹⁰² A priori, la procédure d'enquête ne permet pas ce genre de confusion mais on n'est pas à l'abri de quelques enquêteurs zélés qui souhaitaient avoir 15 cahiers par classe. La base de données ne comprenant pas le nom des élèves, les contrôles a posteriori ne sont pas possibles.

observations ont été retenues comme valeurs extrêmes, elles sont identifiées dans le tableau suivant.

NUMECOLE	Freq.	Percent	Cum.
10	1	3.85	3.85
12	2	7.69	11.54
15	2	7.69	19.23
30	2	7.69	26.92
36	3	11.54	38.46
38	2	7.69	46.15
65	1	3.85	50.00
78	8	30.77	80.77
95	1	3.85	84.62
97	1	3.85	88.46
102	1	3.85	92.31
112	1	3.85	96.15
127	1	3.85	100.00
Total	26	100.00	

On note que l'école 78 est fortement représentée avec 8 élèves ce qui confirme le caractère atypique de cette classe qu'il est préférable d'écartier de nos analyses. Les autres observations identifiées seront également écartées des analyses. Au total, ce seront 32 élèves qui seront exclus des analyses pour la 2^{ème} année.

Les modèles ci-après montrent que lorsqu'on ne prend pas en compte ces 32 observations (soit 1,9 % de l'échantillon), le R² passe de 0,19 à 0,26 soit un gain de près de 7 points qui montre bien l'influence de ces cas extrêmes sur la relation moyenne entre score de début d'année et score de fin d'année.

```
. reg STFIN2FM STINI2FM, cluster (NUMECOLE)
```

Linear regression

Number of obs = **1694**
 F(1, 128) = 55.50
 Prob > F = 0.0000
R-squared = **0.1896**
 Root MSE = .90046

Number of clusters (NUMECOLE) = 129

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.4311392	.0578705	7.45	0.000	.3166325	.545646
_cons	-.0105207	.0594235	-0.18	0.860	-.1281002	.1070588

```
. reg STFIN2FM STINI2FM if VALEX==0, cluster (NUMECOLE)
```

Linear regression

Number of obs = **1662**

F(1, 127) = 133.68

Prob > F = 0.0000

R-squared = **0.2570**

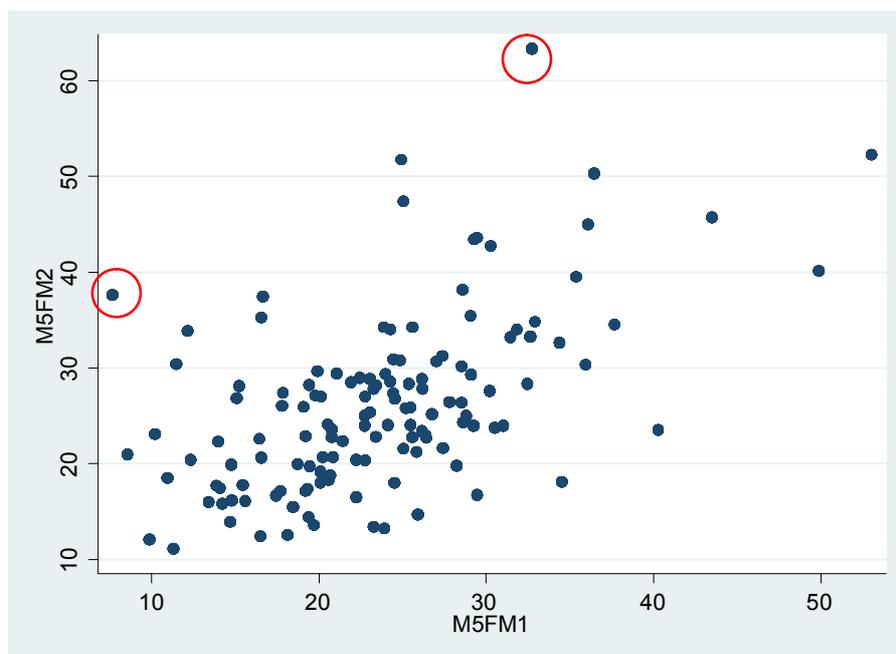
Root MSE = .85237

Number of clusters (NUMECOLE) = 128

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.5133171	.0443966	11.56	0.000	.4254642	.60117
_cons	.0070931	.0557873	0.13	0.899	-.1032998	.1174861

En 5^{ème} année

Le graphique montre la relation entre les scores moyens d'une classe au pré-test (M5FM1) et au post-test (M5FM2).



Le graphique est assez différent de celui de la 2^{ème} année puisque globalement moins dispersé mais avec quelques établissements qui ont des performances assez différentes de la moyenne. Deux cas plus marqués ont été entourés, il s'agit des écoles 98 et 104.

La même procédure qu'en deuxième année a été retenue pour l'identification des valeurs extrêmes au niveau des élèves. Au total, ce sont 27 élèves (1,5 % des observations) qui sont exclus des analyses.

NUMECOLE	Freq.	Percent	Cum.
8	1	3.70	3.70
11	3	11.11	14.81
31	2	7.41	22.22
37	1	3.70	25.93
44	2	7.41	33.33
98	5	18.52	51.85
103	2	7.41	59.26
104	5	18.52	77.78
109	2	7.41	85.19
112	2	7.41	92.59
121	1	3.70	96.30
128	1	3.70	100.00
Total	27	100.00	

On constate que lorsqu'on ne prend pas en compte ces 27 observations, le R^2 caractérisant la régression du score final sur le score initial progresse d'environ 5 points soit légèrement moins qu'en 2^{ème} année mais le gain reste important.

```
. reg STFIN5FM STINI5FM, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1771
                                                F( 1, 133) = 126.28
                                                Prob > F      = 0.0000
                                                R-squared    = 0.2947
                                                Root MSE    = .84007

Number of clusters (NUMECOLE) = 134
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5401249	.0480656	11.24	0.000	.4450529	.6351969
_cons	-.021217	.0597028	-0.36	0.723	-.1393068	.0968727

```
. reg STFIN5FM STINI5FM if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1744
                                                F( 1, 133) = 181.14
                                                Prob > F      = 0.0000
                                                R-squared    = 0.3460
                                                Root MSE    = .7882

Number of clusters (NUMECOLE) = 134
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5789059	.0430126	13.46	0.000	.4938286	.6639832
_cons	-.040789	.0528563	-0.77	0.442	-.1453367	.0637586

ANNEXE 3.1.3 : Comparaison des statistiques descriptives selon le statut de l'enseignant

2^{ème} année

Tests de différence de moyennes au pré-test et au post-test selon le statut du maître

```
. ttest SINI2FM100 if VALEX==0, by(CONTRACTUEL)unequal
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	656	26.0813	.7658259	19.61471	24.57753	27.58507
1	1006	22.80384	.5635284	17.87371	21.69802	23.90967
combined	1662	24.09747	.4573131	18.64357	23.2005	24.99444
diff		3.277457	.9508173		1.412162	5.142753
diff = mean(0) - mean(1)				t =	3.4470	
Ho: diff = 0			Satterthwaite's degrees of freedom = 1306.68			
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.9997		Pr(T > t) = 0.0006		Pr(T > t) = 0.0003		

```
. ttest SFIN2FM100 if VALEX==0, by(CONTRACTUEL)unequal
```

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	656	36.17284	.7788824	19.94912	34.64343	37.70225
1	1006	41.84684	.6505592	20.63411	40.57023	43.12345
combined	1662	39.60728	.5040416	20.54858	38.61866	40.59591
diff		-5.674	1.014832		-7.664716	-3.683283
diff = mean(0) - mean(1)				t =	-5.5911	
Ho: diff = 0			Satterthwaite's degrees of freedom = 1433.11			
Ha: diff < 0		Ha: diff != 0		Ha: diff > 0		
Pr(T < t) = 0.0000		Pr(T > t) = 0.0000		Pr(T > t) = 1.0000		

Tests de Chi2

DEUXPARENT	CONTRACTUEL		Total
	0	1	
0	134	319	453
1	522	687	1,209
Total	656	1,006	1,662

Pearson chi2(1) = 25.4948 Pr = 0.000

REDAN1	CONTRACTUEL		Total
	0	1	
0	547	812	1,359
1	109	194	303
Total	656	1,006	1,662

Pearson chi2(1) = 1.8966 Pr = 0.168

ttest TCLASSE, by(CONTRACTUEL)unequal

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	49	71.81633	3.246096	22.72267	65.28961	78.34304
1	79	66.78481	2.382907	21.17974	62.04081	71.52881
combined	128	68.71094	1.929782	21.83299	64.89225	72.52963
diff		5.031516	4.026833		-2.961221	13.02425

diff = mean(0) - mean(1) t = 1.2495
 Ho: diff = 0 Satterthwaite's degrees of freedom = 96.4376

Ha: diff < 0 Pr(T < t) = 0.8927
 Ha: diff != 0 Pr(|T| > |t|) = 0.2145
 Ha: diff > 0 Pr(T > t) = 0.1073

(mean)	PROPCP2		Total
CONTRACTUE L	0	1	
0	32	17	49
1	49	30	79
Total	81	47	128

Pearson chi2(1) = 0.1401 Pr = 0.708

5^{ème} année

Tests de différence de moyennes au pré-test et au post-test selon le statut du maître

ttest SINI5FM100 if VALEX==0, by(CONTRACTUEL)unequal

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	979	30.68589	.4238646	13.26229	29.8541	31.51768
1	765	32.20391	.5332013	14.74762	31.1572	33.25062
combined	1744	31.35176	.3340322	13.9496	30.69662	32.00691
diff		-1.51802	.6811496		-2.854091	-.1819482

diff = mean(0) - mean(1) t = -2.2286
 Ho: diff = 0 Satterthwaite's degrees of freedom = 1550.88

Ha: diff < 0 Pr(T < t) = 0.0130
 Ha: diff != 0 Pr(|T| > |t|) = 0.0260
 Ha: diff > 0 Pr(T > t) = 0.9870

ttest SFIN5FM100 if VALEX==0, by(CONTRACTUEL)unequal

Two-sample t test with unequal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	979	30.35477	.3898034	12.19655	29.58983	31.11972
1	765	33.43834	.5203924	14.39334	32.41677	34.45991
combined	1744	31.70737	.3182312	13.28972	31.08322	32.33153
diff		-3.083568	.6501961		-4.358962	-1.808174

diff = mean(0) - mean(1) t = -4.7425
 Ho: diff = 0 Satterthwaite's degrees of freedom = 1494.35

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
 Pr(T < t) = **0.0000** Pr(|T| > |t|) = **0.0000** Pr(T > t) = 1.0000

Coefficients de corrélation intraclasse

-> CONTRACTUEL = 0 (enseignants fonctionnaires)

One-way Analysis of Variance for STINI5FM: Standardized values of (SINI5FM)

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	420.6476	75	5.6086347	11.36	0.0000
Within NUMECOLE	445.9885	903	.49389646		
Total	866.6361	978	.88613099		

Number of obs = 979
R-squared = **0.4854**

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.44572	0.04639	0.35480	0.53665

Estimated SD of NUMECOLE effect .6302103
 Estimated SD within NUMECOLE .7027777
 Est. reliability of a NUMECOLE mean 0.91194
 (evaluated at n=12.88)

-> CONTRACTUEL = 1 (enseignants contractuels)

One-way Analysis of Variance for STINI5FM: Standardized values of (SINI5FM)

Number of obs = 765
R-squared = 0.5442

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	455.57621	57	7.9925651	14.81	0.0000
Within NUMECOLE	381.56367	707	.53969401		
Total	837.13988	764	1.0957328		

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.51155	0.05250	0.40864	0.61445
Estimated SD of NUMECOLE effect			.7518048
Estimated SD within NUMECOLE			.7346387
Est. reliability of a NUMECOLE mean (evaluated at n=13.19)			0.93248

One-way Analysis of Variance for STINI5FM: Standardized values of (SINI5FM)

Number of obs = 1744
R-squared = 0.5157

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	881.20937	133	6.6256343	12.89	0.0000
Within NUMECOLE	827.55217	1610	.51400756		
Total	1708.7615	1743	.98035659		

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.47745	0.03476	0.40933	0.54558
Estimated SD of NUMECOLE effect			.6853107
Estimated SD within NUMECOLE			.7169432
Est. reliability of a NUMECOLE mean (evaluated at n=13.01)			0.92242

-> CONTRACTUEL = 0 (enseignants fonctionnaires)

One-way Analysis of Variance for STFIN5FM: Standardized values of (SFIN5FM)

Number of obs = 979
R-squared = 0.6148

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	480.82278	75	6.4109704	19.22	0.0000
Within NUMECOLE	301.26037	903	.33362167		
Total	782.08315	978	.79967602		

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.58584	0.04386	0.49988	0.67179
Estimated SD of NUMECOLE effect			.686959
Estimated SD within NUMECOLE			.5775999
Est. reliability of a NUMECOLE mean			0.94796
(evaluated at n=12.88)			

-> CONTRACTUEL = 1 (enseignants contractuels)

One-way Analysis of Variance for STFIN5FM: Standardized values of (SFIN5FM)

Number of obs = 765
R-squared = 0.6991

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	594.82329	57	10.435496	28.82	0.0000
Within NUMECOLE	256.03418	707	.3621417		
Total	850.85748	764	1.1136878		

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.67841	0.04430	0.59157	0.76524
Estimated SD of NUMECOLE effect			.8740381
Estimated SD within NUMECOLE			.6017821
Est. reliability of a NUMECOLE mean			0.96530
(evaluated at n=13.19)			

```
. loneway STFIN5FM NUMECOLE if VALEX==0
```

One-way Analysis of Variance for STFIN5FM: Standardized values of (SFIN5FM)

Number of obs = 1744
R-squared = 0.6632

Source	SS	df	MS	F	Prob > F
Between NUMECOLE	1097.5966	133	8.252606	23.84	0.0000
Within NUMECOLE	557.29456	1610	.34614569		
Total	1654.8911	1743	.94944988		

Intraclass correlation	Asy. S.E.	[95% Conf. Interval]	
0.63706	0.03103	0.57624	0.69788

Estimated SD of NUMECOLE effect .7794713
Estimated SD within NUMECOLE .5883415
Est. reliability of a NUMECOLE mean 0.95806
(evaluated at n=13.01)

Tests de Chi2

```
. tab PAUVR CONTRACTUEL if VALEX==0, chi2
```

PAUVR	CONTRACTUEL		Total
	0	1	
0	821	540	1,361
1	158	225	383
Total	979	765	1,744

Pearson chi2(1) = 44.1430 Pr = 0.000

```
. tab RICH CONTRACTUEL if VALEX==0, chi2
```

RICH	CONTRACTUEL		Total
	0	1	
0	738	642	1,380
1	241	123	364
Total	979	765	1,744

Pearson chi2(1) = 18.9573 Pr = 0.000

```
. tab REDAN5 CONTRACTUEL if VALEX==0, chi2
```

REDAN5	CONTRACTUEL		Total
	0	1	
0	657	578	1,235
1	322	187	509
Total	979	765	1,744

Pearson chi2(1) = 14.8230 Pr = 0.000

```
. tab LIVR_FRMT CONTRACTUEL if VALEX==0, chi2
```

LIVR_FRMT	CONTRACTUEL		Total
	0	1	
0	776	643	1,419
1	203	122	325
Total	979	765	1,744

Pearson chi2(1) = 6.4921 Pr = 0.011

```
. tab DOUBLFLX CONTRACTUEL if VALEX==0, chi2
```

DOUBLFLX	CONTRACTUEL		Total
	0	1	
0	650	558	1,208
1	329	207	536
Total	979	765	1,744

Pearson chi2(1) = 8.6463 Pr = 0.003

```
. tab RURAL CONTRACTUEL if VALEX==0, chi2
```

RURAL	CONTRACTUEL		Total
	0	1	
0	641	390	1,031
1	338	375	713
Total	979	765	1,744

Pearson chi2(1) = 37.3296 Pr = 0.000

Test de différence de la taille de classe

```
. ttest TCLASSE, by(CONTRACTUEL)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
0	76	67.23684	2.670164	23.27795	61.9176	72.55608
1	58	59.53448	2.965441	22.58412	53.59629	65.47268
combined	134	63.90299	2.005264	23.21261	59.93665	67.86932
diff		7.702359	4.006812		-.2235103	15.62823

diff = mean(0) - mean(1) t = 1.9223
 Ho: diff = 0 degrees of freedom = 132

Ha: diff < 0 Pr(T < t) = 0.9716
 Ha: diff != 0 Pr(|T| > |t|) = 0.0567
 Ha: diff > 0 Pr(T > t) = 0.0284

ANNEXE 3.1.4 : Construction des variables pour mesurer le niveau de vie de la famille de l'élève

Il a été relevé dans le chapitre 1 que les informations relatives au niveau de vie des familles qui étaient collectées auprès des élèves n'offraient pas une précision satisfaisante mais constituaient dans les pays africains la seule source d'information disponible. Pour essayer d'exploiter le plus efficacement possible cette information, la méthode préconisée par Filmer et Pritchett (1998) a été utilisée pour créer des variables relatives au niveau de vie des familles. Cette dernière se fonde sur une analyse en composante principale pour composer un index.

```
pca FAUTEUIL FRIGO ROBINET LAMPETROL ELECTRIC VOITURE VELO MOBYLETT VIDEO TV /*  
*/ RADIO CUISIGAZ EAUTOIL , components(5)
```

```
predict C1, score  
ereturn list  
matrix list e(L)  
matrix W=e(L)
```

```
gen w1=e1(W,1,1)  
sum w1  
egen X1=mean(FAUTEUIL)  
tab X1  
egen X2=sd(FAUTEUIL)  
tab X2  
gen X3=FAUTEUIL-X1  
sum X3  
gen I1=w1*X3/X2  
sum I1
```

On suit la même procédure pour les autres biens (réfrigérateur, robinet, lampe à pétrole, électricité à la maison, voiture, vélo, mobylette, magnétoscope, télévision, radio, cuisinière à gaz, des toilettes avec eau courante). On peut ensuite créer différentes variables :

```
gen IRICH=I1+I2+I3+I4+I5+I6+I7+I8+I9+I10+I11+I12+I13  
sum IRICH  
xtile QRICH=IRICH, nq (5)  
tab QRICH  
gen RICH=QRICH==5 /*quintile des élèves appartenant à la catégorie la plus riche */  
gen PAUVR=QRICH==1 /*quintile des élèves appartenant à la catégorie la plus pauvre*/
```

ANNEXE 3.2.1 : Modèles explicatifs des scores de 2^{ème} année

Dictionnaire des variables utilisées

Variables	
STINI2FM	Score standardisé de français et mathématiques de début d'année
STFIN2FM	Score standardisé de français et mathématiques de fin d'année
FILLE	L'élève est une fille
AGEPLUS	L'élève est plus âgé que l'âge normal pour la classe (ici plus de 8 ans)
PAUVR	L'élève appartient au quintile le plus pauvre
REDAN1	L'élève a redoublé sa 1 ^{ère} année
REDAN2	L'élève a redoublé sa 2 ^{ème} année
BACPLUS	Le maître est titulaire du BAC voire d'un diplôme du supérieur
RESTENS_EST	Si c'était à refaire l'enseignant choisirait à nouveau le même métier (avec imputation des valeurs manquantes)
COLLEGUES	Le maître discute régulièrement avec ses collègues des difficultés rencontrées en classe
SERVICE	Ancienneté du maître (en années)
FPIIANPL	Le maître a suivi une formation professionnelle initiale longue (1 an et plus)
DOUBLFLX	La classe fonctionne en double flux
SALDUR	La salle de classe est construite en dur
COVAR	Coefficient de variation par classe (écart-type /score moyen)
ANCDIR	Ancienneté du directeur d'école dans la fonction de directeur (en années)
APEACTIV	Association de parents d'élèves jugée active ou très active par le directeur
PROPC_ES	Proportion d'enseignants contractuels dans l'école
DEPER	% d'élèves présents au test de début d'année et non soumis au test de fin d'année
RURAL	Le directeur déclare que l'école est en zone rurale
FONCURB	Enseignant fonctionnaire travaillant en zone urbaine
FONCRURAL	Enseignant fonctionnaire travaillant en zone rurale
CONTRURB	Enseignant contractuel travaillant en zone urbaine
CONTRURAL	Enseignant contractuel travaillant en zone rurale

Modèle 1

```
. reg STFIN2FM STINI2FM CONTRACTUEL if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1662
                                                F( 2, 127) =    74.57
                                                Prob > F      =    0.0000
                                                R-squared    =    0.2891
Number of clusters (NUMECOLE) = 128          Root MSE     =    .83399
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.5289756	.0439603	12.03	0.000	.4419861	.6159652
CONTRACTUEL	.3637444	.106229	3.42	0.001	.1535364	.5739524
_cons	-.2130431	.0761378	-2.80	0.006	-.3637061	-.0623801

Modèle 2

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1662
                                                F( 17, 127) =    37.70
                                                Prob > F      =    0.0000
                                                R-squared    =    0.3922
Number of clusters (NUMECOLE) = 128          Root MSE     =    .77467
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.6173424	.0346409	17.82	0.000	.5487943	.6858905
FILLE	-.0689512	.0422269	-1.63	0.105	-.1525106	.0146082
AGEPLUS	.1231395	.0524274	2.35	0.020	.0193951	.2268839
PAUVR	-.1818865	.064796	-2.81	0.006	-.310106	-.053667
REDAN1	-.1504861	.0614922	-2.45	0.016	-.272168	-.0288043
REDAN2	-.2146564	.0562579	-3.82	0.000	-.3259806	-.1033322
CONTRACTUEL	.3052225	.1005519	3.04	0.003	.1062485	.5041966
BACPLUS	-.1723012	.1290774	-1.33	0.184	-.4277222	.0831197
RESTENS_EST	.1954033	.0961283	2.03	0.044	.0051828	.3856238
COLLEGUES	.354705	.1230343	2.88	0.005	.1112423	.5981678
SALDUR	-.2202354	.1089733	-2.02	0.045	-.4358738	-.004597
DOUBLFLX	-.3711776	.1208277	-3.07	0.003	-.6102739	-.1320814
COVAR	.6342898	.2072308	3.06	0.003	.2242174	1.044362
ANCDIR	.0150683	.0053916	2.79	0.006	.0043993	.0257373
APEACTIV	-.1660351	.100855	-1.65	0.102	-.365609	.0335388
PROPC_ES	-.0548904	.2444585	-0.22	0.823	-.5386296	.4288489
RURAL	.0398744	.1083029	0.37	0.713	-.1744375	.2541862
_cons	-.7308598	.222491	-3.28	0.001	-1.171129	-.2905902

```
. vif
```

Variable	VIF	1/VIF
PROPC_ES	1.71	0.586140
RURAL	1.43	0.699127

STINI2FM	1.39	0.718477
COVAR	1.39	0.721336
CONTRACTUEL	1.19	0.839956
DOUBLFLX	1.19	0.842919
PAUVR	1.16	0.863165
AGEPLUS	1.12	0.894930
COLLEGUES	1.11	0.900698
ANCDIR	1.11	0.901846
REDAN2	1.09	0.921178
APEACTIV	1.07	0.933722
REDAN1	1.06	0.944730
RESTENS_EST	1.05	0.948503
BACPLUS	1.05	0.956810
SALDUR	1.04	0.962071
FILLE	1.01	0.987314

Mean VIF	1.19	

. ovtest

Ramsey RESET test using powers of the fitted values of STFIN2FM

Ho: model has no omitted variables
 F(3, 1641) = 1.09
 Prob > F = 0.3516

. linktest

Source	SS	df	MS	Number of obs =	1662

Model	638.125522	2	319.062761	F(2, 1659) =	537.37
Residual	985.034216	1659	.593751788	Prob > F =	0.0000

Total	1623.15974	1661	.977218385	R-squared =	0.3931
				Adj R-squared =	0.3924
				Root MSE =	.77055

STFIN2FM	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	

_hat	1.026565	.0346803	29.60	0.000	.9585428	1.094586
_hatsq	-.0527586	.0326337	-1.62	0.106	-.1167662	.011249
_cons	.0200525	.0226082	0.89	0.375	-.024291	.064396

Modèle 3 : Modèle multiniveau

```
. xtmixed STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0 ||NUMECOLE:, variance
```

Mixed-effects REML regression
 Group variable: NUMECOLE

Number of obs = 1662
 Number of groups = 128

Obs per group: min = 6
 avg = 13.0
 max = 15

Wald chi2(17) = 961.48
 Prob > chi2 = 0.0000

Log restricted-likelihood = -1651.6551

STFIN2FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

STINI2FM	.6060305	.0212181	28.56	0.000	.5644436	.6476173

FILLE	-.0592509	.0303498	-1.95	0.051	-.1187354	.0002336
AGEPLUS	.0663173	.036915	1.80	0.072	-.0060347	.1386693
PAUVR	-.1029926	.0438393	-2.35	0.019	-.188916	-.0170691
REDAN1	-.0972197	.0418793	-2.32	0.020	-.1793016	-.0151379
REDAN2	-.1812402	.0453419	-4.00	0.000	-.2701088	-.0923717
CONTRACTUEL	.3175228	.1102368	2.88	0.004	.1014627	.533583
BACPLUS	-.1943886	.1869808	-1.04	0.299	-.5608642	.172087
RESTENS_EST	.1990762	.1041575	1.91	0.056	-.0050688	.4032211
COLLEGUES	.3559293	.1656544	2.15	0.032	.0312526	.6806061
SALDUR	-.2067713	.1138685	-1.82	0.069	-.4299495	.0164068
DOUBLFLX	-.3897299	.1251872	-3.11	0.002	-.6350923	-.1443676
COVAR	.6530542	.182566	3.58	0.000	.2952315	1.010877
ANCDIR	.0154893	.0066978	2.31	0.021	.0023617	.0286168
APEACTIV	-.1808968	.1022806	-1.77	0.077	-.381363	.0195694
PROPC_ES	-.1080416	.2679622	-0.40	0.687	-.6332379	.4171546
RURAL	.0354103	.1175141	0.30	0.763	-.194913	.2657336
_cons	-.7326252	.2672974	-2.74	0.006	-1.256518	-.208732

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLE: Identity				
var(_cons)	.2857346	.0412071	.2153809	.3790691
var(Residual)	.3441234	.0124513	.3205645	.3694137

LR test vs. linear regression: chibar2(01) = 630.75 Prob >= chibar2 = 0.0000

Modèle 4 : Modèle avec interactions entre RURAL et CONTRACTUEL

```
. xtmixed STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV CONTRURB CONTRURAL FONCRURAL /*
> */if VALEX==0 ||NUMECOLE:, variance
```

Mixed-effects REML regression
Group variable: NUMECOLE

Number of obs = 1662
Number of groups = 128
Obs per group: min = 6
avg = 13.0
max = 15

Log restricted-likelihood = -1650.8882
Wald chi2(17) = 964.69
Prob > chi2 = 0.0000

STFIN2FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
STINI2FM	.6051752	.0212186	28.52	0.000	.5635875	.6467629
FILLE	-.0590491	.0303425	-1.95	0.052	-.1185194	.0004212
AGEPLUS	.0662613	.0369034	1.80	0.073	-.006068	.1385907
PAUVR	-.103097	.0438128	-2.35	0.019	-.1889684	-.0172256
REDAN1	-.0973165	.0418728	-2.32	0.020	-.1793857	-.0152473
REDAN2	-.1805561	.0453363	-3.98	0.000	-.2694136	-.0916985
BACPLUS	-.1591226	.1870086	-0.85	0.395	-.5256527	.2074074
RESTENS_EST	.2010932	.1032681	1.95	0.051	-.0013085	.4034949
COLLEGUES	.3606429	.1619983	2.23	0.026	.043132	.6781537
SALDUR	-.2058676	.1122919	-1.83	0.067	-.4259556	.0142204
DOUBLFLX	-.4151981	.12315	-3.37	0.001	-.6565678	-.1738285
COVAR	.7006703	.1844622	3.80	0.000	.339131	1.06221

ANCDIR	.0168025	.0065403	2.57	0.010	.0039836	.0296213
APEACTIV	-.1718053	.1015692	-1.69	0.091	-.3708773	.0272666
CONTRURB	.1803733	.1353594	1.33	0.183	-.0849263	.4456728
CONTRURAL	.3014776	.1355216	2.22	0.026	.0358601	.5670951
FONCRURAL	-.2001592	.1845568	-1.08	0.278	-.5618839	.1615655
_cons	-.7492757	.252557	-2.97	0.003	-1.244278	-.2542732

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
NUMECOLE: Identity			
var(_cons)	.2808109	.0405469	.2115958 .3726668
var(Residual)	.344115	.0124507	.3205572 .3694041

LR test vs. linear regression: chibar2(01) = 622.24 Prob >= chibar2 = 0.0000

La variable PROP_EC a été retirée du modèle car elle est colinéaire avec la variable CONTRURAL.

Modèle 5 : Second modèle avec interactions entre RURAL et CONTRACTUEL

```
. xtmixed STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV CONTRURAL FONCURB FONCRURAL /*
> */if VALEX==0 ||NUMECOLE:, variance
```

Mixed-effects REML regression
Group variable: NUMECOLE
Number of obs = 1662
Number of groups = 128
Obs per group: min = 6
 avg = 13.0
 max = 15

Log restricted-likelihood = -1650.8882
Wald chi2(17) = 964.69
Prob > chi2 = 0.0000

STFIN2FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
STINI2FM	.6051752	.0212186	28.52	0.000	.5635875 .6467629
FILLE	-.0590491	.0303425	-1.95	0.052	-.1185194 .0004212
AGEPLUS	.0662613	.0369034	1.80	0.073	-.006068 .1385907
PAUVR	-.103097	.0438128	-2.35	0.019	-.1889684 -.0172256
REDAN1	-.0973165	.0418728	-2.32	0.020	-.1793857 -.0152473
REDAN2	-.1805561	.0453363	-3.98	0.000	-.2694136 -.0916985
BACPLUS	-.1591226	.1870086	-0.85	0.395	-.5256527 .2074074
RESTENS_EST	.2010932	.1032681	1.95	0.051	-.0013085 .4034949
COLLEGUES	.3606429	.1619983	2.23	0.026	.043132 .6781537
SALDUR	-.2058676	.1122919	-1.83	0.067	-.4259556 .0142204
DOUBLFLX	-.4151981	.12315	-3.37	0.001	-.6565678 -.1738285
COVAR	.7006703	.1844622	3.80	0.000	.339131 1.06221
ANCDIR	.0168025	.0065403	2.57	0.010	.0039836 .0296213
APEACTIV	-.1718053	.1015692	-1.69	0.091	-.3708773 .0272666
CONTRURAL	.1211044	.1291848	0.94	0.349	-.1320931 .3743018
FONCURB	-.1803733	.1353594	-1.33	0.183	-.4456728 .0849263
FONCRURAL	-.3805325	.1741805	-2.18	0.029	-.72192 -.0391449
_cons	-.5689025	.2698546	-2.11	0.035	-1.097808 -.0399971

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
-----+-----				
NUMECOLE: Identity				
var(_cons)	.2808109	.0405469	.2115958	.3726668
-----+-----				
var(Residual)	.344115	.0124507	.3205572	.3694041
-----+-----				
LR test vs. linear regression: chibar2(01) = 622.24 Prob >= chibar2 = 0.0000				

ANNEXE 3.2.2 : Modèles explicatifs des scores de 5^{ème} année

Dictionnaire des variables utilisées

Variables	
STINI5FM	Score standardisé de français et mathématiques de début d'année
STFIN5FM	Score standardisé de français et mathématiques de fin d'année
FILLE	L'élève est une fille
AGEPLUS	L'élève est plus âgé que l'âge normal pour la classe
RICH	L'élève appartient au quintile le plus riche
REDAN4	L'élève a redoublé sa 4 ^{ème} année
PROCHECOLE	L'enseignant habite à proximité de l'école
BACPLUS	Le maître est titulaire du BAC voire d'un diplôme du supérieur
DOUBLFLX	La classe fonctionne en double flux
REUMOIS	L'ensemble des enseignants et le directeur se réunissent une fois par mois
ANCDIR	Ancienneté du directeur d'école dans la fonction de directeur (en années)
APEACTIV	Association de parents d'élèves jugée active ou très active par le directeur
PROPC_ES	Proportion d'enseignants contractuels par école
DEPER	% d'élèves présents au test de début d'année et non soumis au test de fin d'année
RURAL	Le directeur déclare que l'école est en zone rurale
FONCURB	Enseignant fonctionnaire travaillant en zone urbaine
CONTRURB	Enseignant contractuel travaillant en zone urbaine
CONTRURAL	Enseignant contractuel travaillant en zone rurale
FONCRURAL	Enseignant fonctionnaire travaillant en zone rurale

Modèle 1

```
. reg STFIN5FM STINI5FM if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1730
                                                F( 1, 133) = 178.83
                                                Prob > F      = 0.0000
                                                R-squared    = 0.3506
Number of clusters (NUMECOLE) = 134          Root MSE     = .78173
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5828091	.0435824	13.37	0.000	.4966049	.6690134
_cons	-.0239211	.0524772	-0.46	0.649	-.1277189	.0798767

Modèle 2

```
. reg STFIN5FM STINI5FM CONTRACTUEL if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1730
                                                F( 2, 133) = 92.51
                                                Prob > F      = 0.0000
                                                R-squared    = 0.3577
Number of clusters (NUMECOLE) = 134          Root MSE     = .77763
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.578748	.0426776	13.56	0.000	.4943333	.6631627
CONTRACTUEL	.1654983	.1067584	1.55	0.123	-.0456657	.3766624
_cons	-.0966349	.0631633	-1.53	0.128	-.2215694	.0282996

Modèle 3

```
. reg STFIN5FM STINI5FM FILLE AGEPLUS RICH /*
> */ REDAN4 /*
> */ PROCHECOLE CONTRACTUEL BACPLUS /*
> */ DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =    1730
                                                F( 13, 133) = 29.36
                                                Prob > F      = 0.0000
                                                R-squared    = 0.4362
Number of clusters (NUMECOLE) = 134          Root MSE     = .7309
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5451126	.037286	14.62	0.000	.4713624	.6188629
FILLE	-.0628191	.0385849	-1.63	0.106	-.1391386	.0135003
AGEPLUS	-.0850244	.0574252	-1.48	0.141	-.1986092	.0285603
RICH	.0000687	.0609225	0.00	0.999	-.1204337	.120571
REDAN4	-.1420178	.0712858	-1.99	0.048	-.2830183	-.0010173
PROCHECOLE	.335568	.0896844	3.74	0.000	.1581756	.5129603
CONTRACTUEL	.2209731	.0978889	2.26	0.026	.0273527	.4145934
BACPLUS	-.2237592	.1198082	-1.87	0.064	-.4607352	.0132167
DOUBLFLX	-.3262699	.1031282	-3.16	0.002	-.5302535	-.1222864
REUMOIS	.2095979	.0995516	2.11	0.037	.0126887	.406507

ANCDIR		.0134175	.0059074	2.27	0.025	.0017329	.025102
APEACTIV		-.175313	.094805	-1.85	0.067	-.3628336	.0122076
RURAL		-.0980306	.1000908	-0.98	0.329	-.2960063	.099945
_cons		-.2825719	.1329373	-2.13	0.035	-.5455167	-.0196271

Modèle 4 : modèle multiniveau

```
. xtmixed STFIN5FM STINI5FM FILLE AGEPLUS RICH REDAN4 /*
> */ PROCHECOLE CONTRACTUEL BACPLUS DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL if VALEX=0 ||NUMECOLE:, variance
```

```
Mixed-effects REML regression                    Number of obs      =      1730
Group variable: NUMECOLE                        Number of groups   =       134

Obs per group: min =          8
                  avg =       12.9
                  max =       15

Wald chi2(13) =      852.51
Prob > chi2   =      0.0000

Log restricted-likelihood = -1414.7154
```

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
STINI5FM	.4535853	.0166407	27.26	0.000	.4209702 .4862004
FILLE	-.0958052	.0248929	-3.85	0.000	-.1445943 -.0470161
AGEPLUS	-.058692	.0280522	-2.09	0.036	-.1136733 -.0037107
RICH	.0358562	.0352987	1.02	0.310	-.033328 .1050405
REDAN4	-.0396463	.033244	-1.19	0.233	-.1048034 .0255108
PROCHECOLE	.3381619	.1688928	2.00	0.045	.0071381 .6691856
CONTRACTUEL	.2002532	.1137279	1.76	0.078	-.0226495 .4231558
BACPLUS	-.2113614	.1783564	-1.19	0.236	-.5609335 .1382107
DOUBLFLX	-.3665751	.1258066	-2.91	0.004	-.6131514 -.1199988
REUMOIS	.2253684	.1103015	2.04	0.041	.0091815 .4415554
ANCDIR	.0147565	.007204	2.05	0.041	.0006369 .0288762
APEACTIV	-.1565423	.1086205	-1.44	0.150	-.3694345 .0563499
RURAL	-.0890384	.1187644	-0.75	0.453	-.3218122 .1437355
_cons	-.3180846	.2057413	-1.55	0.122	-.7213301 .0851609

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
NUMECOLE: Identity			
var(_cons)	.3659166	.0490349	.281395 .4758257
var(Residual)	.2316023	.0082167	.2160451 .2482798

```
LR test vs. linear regression: chibar2(01) = 1054.03 Prob >= chibar2 = 0.0000
```

Modèle 5 : modèle avec interaction Contractuel/RURAL

```
. xtmixed STFIN5FM STINI5FM FILLE AGEPLUS RICH REDAN4 /*
> */ PROCHECOLE BACPLUS DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV FONCURB CONTRURB CONTRURAL if VALEX=0 ||NUMECOLE:, variance
```

```
Mixed-effects REML regression                    Number of obs      =      1730
Group variable: NUMECOLE                        Number of groups   =       134

Obs per group: min =          8
                  avg =       12.9
                  max =       15
```

Log restricted-likelihood = -1415.261 Wald chi2(14) = 851.67
 Prob > chi2 = 0.0000

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
STINI5FM	.4535145	.0166437	27.25	0.000	.4208935	.4861355
FILLE	-.0958273	.0248933	-3.85	0.000	-.1446174	-.0470373
AGEPLUS	-.0586751	.0280537	-2.09	0.036	-.1136594	-.0036907
RICH	.0358648	.0353033	1.02	0.310	-.0333285	.105058
REDAN4	-.0396032	.0332472	-1.19	0.234	-.1047665	.0255601
PROCHECOLE	.3375962	.1710416	1.97	0.048	.0023609	.6728315
BACPLUS	-.2107154	.1808647	-1.17	0.244	-.5652037	.1437728
DOUBLFLX	-.3669657	.1271452	-2.89	0.004	-.6161657	-.1177657
REUMOIS	.2258574	.1123664	2.01	0.044	.0056233	.4460915
ANCDIR	.0147873	.0073276	2.02	0.044	.0004254	.0291491
APEACTIV	-.1564784	.1090855	-1.43	0.151	-.3702821	.0573253
FONCURB	.0915834	.1559767	0.59	0.557	-.2141252	.397292
CONTRURB	.2892444	.1779204	1.63	0.104	-.0594731	.6379619
CONTRURAL	.203534	.1728448	1.18	0.239	-.1352357	.5423036
_cons	-.4087546	.2226915	-1.84	0.066	-.8452219	.0277127

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLES: Identity				
var(_cons)	.3690673	.0496324	.2835539	.4803697
var(Residual)	.2315999	.0082165	.216043	.248277

LR test vs. linear regression: chibar2(01) = 1056.26 Prob >= chibar2 = 0.0000

ANNEXE 3.2.3 : Tentative d'instrumentation de la variable contractuelle dans les modèles explicatifs des scores de 2^{ème} et 5^{ème} années

L'instrumentation consiste à instrumenter la variable contractuelle avec les variables d'ancienneté et de formation professionnelle, toutes deux très corrélées au statut (Cf. annexe 3.2.1). Toutefois, l'estimation classique avec variables instrumentales repose sur l'hypothèse d'indépendance et de normalité des résidus qui est remise en cause dans le cas de nos données. Aussi, suivant Baum (2006), un estimateur spécifique¹⁰³ est utilisé ici pour le modèle avec variables instrumentales qui prend également en compte les effets de groupement (Cf. Baum, 2006, p.194 et suivantes).

Un test d'endogénéité a été réalisé avant de procéder à l'estimation des modèles avec variables instrumentales.

2^{ème} année

Test d'endogénéité de la variable contractuelle.

```
. regress CONTRACTUEL STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ BACPLUS RESTENS_EST COLLEGUES SERVICE FPI1ANPL/*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0, cluster (NUMECOLE)
```

Linear regression	Number of obs = 1662
	F(18, 127) = 74.91
	Prob > F = 0.0000
	R-squared = 0.8594
Number of clusters (NUMECOLE) = 128	Root MSE = .18433

CONTRACTUEL	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	-.0222716	.0115057	-1.94	0.055	-.0450392	.0004961
FILLE	.0002395	.0087263	0.03	0.978	-.0170282	.0175072
AGEPLUS	-.0221587	.0165874	-1.34	0.184	-.0549822	.0106648
PAUVR	-.0185545	.0131512	-1.41	0.161	-.0445784	.0074694
REDAN1	-.0184528	.0130092	-1.42	0.159	-.0441957	.0072902
REDAN2	-.0199456	.017322	-1.15	0.252	-.0542228	.0143316
BACPLUS	.0357819	.1065406	0.34	0.738	-.1750428	.2466065
RESTENS_EST	.103253	.0323442	3.19	0.002	.0392496	.1672563
COLLEGUES	.0210404	.0402906	0.52	0.602	-.0586873	.1007682
SERVICE	-.0258761	.0029835	-8.67	0.000	-.0317799	-.0199723
FPI1ANPL	-.5097446	.0639634	-7.97	0.000	-.6363167	-.3831726
SALDUR	-.009159	.0365088	-0.25	0.802	-.0814033	.0630853
DOUBLFLX	-.0120898	.0479222	-0.25	0.801	-.1069193	.0827396
COVAR	.0229918	.0465497	0.49	0.622	-.0691217	.1151052
ANCDIR	-.0004253	.0021341	-0.20	0.842	-.0046482	.0037976

¹⁰³ Generalized Method of Moments estimator ou GMM estimator.

```

APEACTIV | -.0176275 .0338931 -0.52 0.604 -.0846958 .0494408
PROPC_ES | -.0469861 .0967979 -0.49 0.628 -.2385317 .1445595
RURAL | .0072534 .0370683 0.20 0.845 -.066098 .0806048
_cons | 1.022612 .0912737 11.20 0.000 .8419977 1.203226
-----

```

```
. predict RES , residual
```

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ RES CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0, cluster (NUMECOLE)
```

```

Linear regression                               Number of obs =    1662
                                                F( 18, 127) =    36.07
                                                Prob > F      =    0.0000
                                                R-squared    =    0.3961
                                                Root MSE    =    .77241

Number of clusters (NUMECOLE) = 128

```

```
-----
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
STINI2FM	.6176134	.0340506	18.14	0.000	.5502335 .6849933
FILLE	-.0698362	.0423377	-1.65	0.102	-.1536149 .0139426
AGEPLUS	.1261618	.0521482	2.42	0.017	.02297 .2293537
PAUVR	-.1750499	.0649642	-2.69	0.008	-.3036023 -.0464976
REDAN1	-.1533308	.0605852	-2.53	0.013	-.2731952 -.0334208
REDAN2	-.2173626	.0553151	-3.93	0.000	-.3268213 -.107904
RES	-.3691205	.242366	-1.52	0.130	-.848719 .1104781
CONTRACTUEL	.367008	.1089889	3.37	0.001	.1513387 .5826774
BACPLUS	-.1764542	.1243267	-1.42	0.158	-.4224744 .0695659
RESTENS_EST	.1885148	.0956407	1.97	0.051	-.0007409 .3777704
COLLEGUES	.3548581	.1189446	2.98	0.003	.1194882 .5902279
SALDUR	-.2212604	.1066748	-2.07	0.040	-.4323506 -.0101703
DOUBLFLX	-.3738209	.1190134	-3.14	0.002	-.6093269 -.1383148
COVAR	.6282753	.1999516	3.14	0.002	.2326072 1.023944
ANCDIR	.0147617	.0052231	2.83	0.005	.0044262 .0250972
APEACTIV	-.1660497	.100347	-1.65	0.100	-.3646182 .0325189
PROPC_ES	-.1015703	.2505748	-0.41	0.686	-.5974127 .3942721
RURAL	.0364935	.1088011	0.34	0.738	-.1788042 .2517911
_cons	-.7328898	.2148767	-3.41	0.001	-1.158092 -.3076876

```
-----
```

La non significativité de la variable RES, qui correspond au résidu du premier modèle, permet de rejeter l’hypothèse de l’endogénéité de la variable CONTRACTUEL (Wooldridge, 2006). Toutefois, la fragilité du test qui se fonde sur un estimation d’une variable dichotomique avec des MCO invite à estimer un modèle avec variables instrumentales pour confirmer les résultats.

Modèle avec variables instrumentales

```
. ivreg2 STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL (CONTRACTUEL= SERVICE
FPI1ANPL) /*
> */if VALEX==0, gmm cluster(NUMECOLE)
```

```
GMM estimation
-----
```

Number of clusters (NUMECOLE) = 128
 Total (centered) SS = 1623.159737
 Total (uncentered) SS = 1623.217617
 Residual SS = 987.9055392

Number of obs = 1662
 F(17, 110) = 2.50
 Prob > F = 0.0023
 Centered R2 = 0.3914
 Uncentered R2 = 0.3914
 Root MSE = .77

STFIN2FM	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
CONTRACTUEL	.3529976	.1052791	3.35	0.001	.1466544	.5593408
STINI2FM	.6152823	.0345736	17.80	0.000	.5475192	.6830454
FILLE	-.0690391	.0420411	-1.64	0.101	-.151438	.0133599
AGEPLUS	.1343122	.0517177	2.60	0.009	.0329474	.235677
PAUVR	-.1712066	.0643459	-2.66	0.008	-.2973222	-.045091
REDAN1	-.1577636	.0607808	-2.60	0.009	-.2768918	-.0386353
REDAN2	-.2182181	.0565407	-3.86	0.000	-.3290357	-.1074005
BACPLUS	-.1771435	.1322681	-1.34	0.180	-.4363841	.0820972
RESTENS_EST	.2076242	.0943046	2.20	0.028	.0227906	.3924579
COLLEGUES	.36985	.1190372	3.11	0.002	.1365414	.6031586
SALDUR	-.2299683	.107739	-2.13	0.033	-.4411329	-.0188036
DOUBLFLX	-.3748804	.1211416	-3.09	0.002	-.6123137	-.1374471
COVAR	.6335012	.2026801	3.13	0.002	.2362555	1.030747
ANCDIR	.0147622	.0053089	2.78	0.005	.0043569	.0251676
APEACTIV	-.1585657	.0997306	-1.59	0.112	-.354034	.0369027
PROPC_ES	-.1564971	.2412651	-0.65	0.517	-.629368	.3163737
RURAL	.0499738	.1066592	0.47	0.639	-.1590744	.259022
_cons	-.7240323	.220133	-3.29	0.001	-1.155485	-.2925796

Hansen J statistic (overidentification test of all instruments): 1.455
 Chi-sq(1) P-val = 0.22772

Instrumented: CONTRACTUEL
 Instruments: SERVICE FP1IANPL STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2
 BACPLUS RESTENS_EST COLLEGUES SALDUR DOUBLFLX COVAR ANCDIR
 APEACTIV PROPC_ES RURAL

Le résultat du modèle avec variable instrumentale est très proche des résultats obtenus avec les autres modèles et tend à renforcer notre confiance dans les résultats obtenus.

5^{ème} année

Test d'endogénéité de la variable contractuelle.

```
. regress CONTRACTUEL STINI5FM FILLE AGEPLUS RICH /*
> */ REDAN4 /*
> */ PROCHECOLE BACPLUS SERVICE FP1IANPL /*
> */ DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL if VALEX==0, cluster (NUMECOLE)
```

Linear regression

Number of obs = 1730
 F(14, 133) = 122.34
 Prob > F = 0.0000
 R-squared = 0.7973
 Root MSE = .2244

Number of clusters (NUMECOLE) = 134

Robust

CONTRACTUEL	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.0160514	.0159311	1.01	0.316	-.0154598	.0475626
FILLE	-.0032742	.0126875	-0.26	0.797	-.0283696	.0218213
AGEPLUS	.0053497	.0208801	0.26	0.798	-.0359503	.0466497
RICH	-.0221392	.0266391	-0.83	0.407	-.0748303	.0305519
REDAN4	-.0081563	.0195848	-0.42	0.678	-.0468942	.0305816
PROCHECOLE	.0132419	.081475	0.16	0.871	-.1479125	.1743962
BACPLUS	.1462538	.0985505	1.48	0.140	-.0486752	.3411828
SERVICE	-.0270094	.002923	-9.24	0.000	-.032791	-.0212277
FPILANPL	-.4990651	.056202	-8.88	0.000	-.6102304	-.3878997
DOUBLFLX	-.057099	.0447933	-1.27	0.205	-.1456984	.0315004
REUMOIS	.043266	.0398558	1.09	0.280	-.0355672	.1220993
ANCDIR	.0022438	.0028783	0.78	0.437	-.0034492	.0079369
APEACTIV	-.0226489	.0411068	-0.55	0.583	-.1039565	.0586588
RURAL	-.0086463	.0440718	-0.20	0.845	-.0958187	.0785261
_cons	1.04647	.0863303	12.12	0.000	.8757118	1.217228

```
. predict RES, residual
```

```
. reg STFIN5FM STINI5FM FILLE AGEPLUS RICH /*
> */ REDAN4 /*
> */ PROCHECOLE CONTRACTUEL BACPLUS /*
> */ DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL RES if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                                Number of obs =    1730
                                                F( 14,   133) =    27.41
                                                Prob > F       =    0.0000
                                                R-squared     =    0.4366
Number of clusters (NUMECOLE) = 134          Root MSE      =    .73088
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5445831	.0373797	14.57	0.000	.4706475	.6185188
FILLE	-.0624467	.0387181	-1.61	0.109	-.1390296	.0141361
AGEPLUS	-.0841724	.0571614	-1.47	0.143	-.1972355	.0288906
RICH	.0017792	.0612985	0.03	0.977	-.1194668	.1230252
REDAN4	-.1427232	.0713631	-2.00	0.048	-.2838767	-.0015698
PROCHECOLE	.3359882	.0893792	3.76	0.000	.1591997	.5127768
CONTRACTUEL	.2422816	.113853	2.13	0.035	.0170847	.4674785
BACPLUS	-.2330944	.1209661	-1.93	0.056	-.4723607	.0061719
DOUBLFLX	-.3241226	.1029835	-3.15	0.002	-.52782	-.1204253
REUMOIS	.2109045	.0997168	2.12	0.036	.0136685	.4081406
ANCDIR	.013428	.0059318	2.26	0.025	.0016951	.0251608
APEACTIV	-.1769571	.0948679	-1.87	0.064	-.3646021	.0106878
RURAL	-.1014397	.1005001	-1.01	0.315	-.3002251	.0973456
RES	-.0935308	.169573	-0.55	0.582	-.4289396	.241878
_cons	-.2912176	.1371868	-2.12	0.036	-.5625677	-.0198674

On fait le même constat que pour la 2ème année, l'endogénéité n'apparaît pas avérée.

Modèle avec variables instrumentales

```
. ivreg2 STFIN5FM STINI5FM FILLE AGEPLUS RICH REDAN4 /*
> */ PROCHECOLE BACPLUS DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL (CONTRACTUEL= SERVICE FPILANPL) /*
> */if VALEX==0, gmm cluster(NUMECOLE)
```

GMM estimation

Number of clusters (NUMECOLE) = 134

Number of obs = 1730

Total (centered) SS = 1625.955653

F(13, 120) = 2.07

Total (uncentered) SS = 1627.05906

Prob > F = 0.0204

Residual SS = 917.0227958

Centered R2 = 0.4360

Uncentered R2 = 0.4364

Root MSE = .73

STFIN5FM	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
CONTRACTUEL	.2466559	.1124766	2.19	0.028	.0262057	.4671061
STINI5FM	.5439489	.036913	14.74	0.000	.4716008	.616297
FILLE	-.0620379	.0382743	-1.62	0.105	-.1370541	.0129782
AGEPLUS	-.0826659	.0566672	-1.46	0.145	-.1937316	.0283997
RICH	.0009377	.0609114	0.02	0.988	-.1184465	.1203219
REDAN4	-.1398559	.0702201	-1.99	0.046	-.2774848	-.002227
PROCHECOLE	.3311572	.087841	3.77	0.000	.158992	.5033223
BACPLUS	-.229212	.1186001	-1.93	0.053	-.4616639	.0032398
DOUBLFLX	-.3269562	.1018962	-3.21	0.001	-.526669	-.1272434
REUMOIS	.211849	.0988311	2.14	0.032	.0181435	.4055545
ANCDIR	.0134318	.0058648	2.29	0.022	.001937	.0249267
APEACTIV	-.1810592	.0932636	-1.94	0.052	-.3638524	.0017341
RURAL	-.0922522	.0955312	-0.97	0.334	-.2794899	.0949856
_cons	-.2941259	.1359177	-2.16	0.030	-.5605197	-.0277322

Hansen J statistic (overidentification test of all instruments): 0.098

Chi-sq(1) P-val = 0.75451

Instrumented: CONTRACTUEL

Instruments: SERVICE FPILANPL STINI5FM FILLE AGEPLUS RICH REDAN4 PROCHECOLE

BACPLUS DOUBLFLX REUMOIS ANCDIR APEACTIV RURAL

Les résultats de 2^{ème} et 5^{ème} années sont convergents. Ils ne permettent pas de conclure à une endogénéité et, de façon cohérente, les modèles avec variables instrumentales fournissent des résultats très proches des autres modèles. Toutefois, deux problèmes se posent avec cette instrumentation. Le premier tient au fait qu'il est difficile d'affirmer que l'ancienneté et la formation du maître affecte le statut mais pas les apprentissages des élèves. Certes, la plupart des études montrent que le lien est très modéré mais pas toujours inexistant. Le second problème est que ces deux variables fournissent quasiment la même information que la variable contractuelle compte tenu de leur proximité puisqu'un contractuel possède une faible ancienneté et rarement une formation professionnelle. Toutefois, ce point est discutable dans la mesure où il ne s'agit pas d'une identification parfaite et qu'une variable instrumentale de qualité doit être fortement liée à la variable instrumentée. Pour ce qui est du premier point des traitements complémentaires ont été faits uniquement sur les enseignants titulaires pour voir si l'ancienneté influençait les acquis des élèves dans les modèles estimés. Les résultats des modèles qui suivent ne révèlent pas d'effet significatif de l'ancienneté. La même démarche a ensuite été adoptée pour la formation professionnelle initiale longue, mais cette fois en se basant sur les enseignants contractuels. La conclusion est identique et montre, aussi frustrant

que cela puisse être, qu'il n'y a pas de lien avéré entre formation professionnelle et acquis scolaires des élèves. Ces résultats tendent donc à justifier l'utilisation de ces variables comme instruments.

Test de l'impact de l'ancienneté chez les titulaires en 2ème année au Mali

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ SERVICE BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =      656
                                                F( 17,      48) =    25.94
                                                Prob > F          =    0.0000
                                                R-squared         =    0.4563
Number of clusters (NUMECOLE) = 49           Root MSE         =    .71703
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.615237	.0577792	10.65	0.000	.4990641	.73141
FILLE	.0032642	.0519693	0.06	0.950	-.1012271	.1077555
AGEPLUS	.1669377	.0704102	2.37	0.022	.0253685	.3085069
PAUVR	-.067432	.0833986	-0.81	0.423	-.2351161	.1002521
REDAN1	-.2039348	.0945614	-2.16	0.036	-.3940632	-.0138064
REDAN2	-.2226228	.0768113	-2.90	0.006	-.3770623	-.0681833
SERVICE	.0086292	.0082356	1.05	0.300	-.0079295	.0251879
BACPLUS	.0911405	.1736939	0.52	0.602	-.2580945	.4403755
RESTENS_EST	.2006015	.1357802	1.48	0.146	-.0724029	.4736059
COLLEGUES	.6482568	.1942921	3.34	0.002	.2576062	1.038907
SALDUR	-.2787357	.1266825	-2.20	0.033	-.5334479	-.0240235
DOUBLFLX	-.3006123	.1741747	-1.73	0.091	-.6508139	.0495894
COVAR	.8808453	.3493258	2.52	0.015	.1784787	1.583212
ANCDIR	.0073361	.0092541	0.79	0.432	-.0112706	.0259428
APEACTIV	-.153296	.1465222	-1.05	0.301	-.4478985	.1413066
PROPC_ES	-.2929311	.3166331	-0.93	0.360	-.9295646	.3437025
RURAL	-.157173	.1675479	-0.94	0.353	-.4940506	.1797046
_cons	-1.142776	.3104126	-3.68	0.001	-1.766902	-.5186492

```
. sum SERVICE if CONTRACTUEL==0
```

Variable	Obs	Mean	Std. Dev.	Min	Max
SERVICE	733	20.19782	8.864826	1	36

Test de l'impact de l'ancienneté chez les titulaires en 5ème année au Mali

```
. reg STFIN5FM STINI5FM FILLE AGEPLUS RICH /*
> */ REDAN4 /*
> */ PROCHECOLE SERVICE BACPLUS /*
> */ DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =      970
                                                F( 13,      75) =    13.87
                                                Prob > F          =    0.0000
                                                R-squared         =    0.4061
Number of clusters (NUMECOLE) = 76           Root MSE         =    .68849
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5289334	.055095	9.60	0.000	.4191784	.6386883
FILLE	-.1191544	.043391	-2.75	0.008	-.2055937	-.0327151
AGEPLUS	-.0491738	.061147	-0.80	0.424	-.1709848	.0726372
RICH	-.0255975	.0749799	-0.34	0.734	-.1749652	.1237701
REDAN4	-.1081599	.0707403	-1.53	0.130	-.2490818	.0327621
PROCHECOLE	.3340805	.1161293	2.88	0.005	.102739	.565422
SERVICE	-.005054	.0063428	-0.80	0.428	-.0176895	.0075814
BACPLUS	-.4131077	.251449	-1.64	0.105	-.9140198	.0878045
DOUBLFLX	-.3435568	.126263	-2.72	0.008	-.5950855	-.092028
REUMOIS	.1145457	.1184561	0.97	0.337	-.121431	.3505223
ANCDIR	.0182347	.0086454	2.11	0.038	.0010123	.0354572
APEACTIV	-.2842381	.1132028	-2.51	0.014	-.5097496	-.0587265
RURAL	-.1089148	.122905	-0.89	0.378	-.3537541	.1359244
_cons	-.1241451	.1637575	-0.76	0.451	-.4503668	.2020767

. sum SERVICE

Variable	Obs	Mean	Std. Dev.	Min	Max
SERVICE	987	20.1925	8.131365	0	38

Test de l'impact de la formation professionnelle initiale longue chez les contractuels en 2ème année au Mali

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ FPI1ANPL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0&CONTRACTUEL==1, cluster (NUMECOLE)
```

Linear regression

Number of obs = 1006
F(17, 78) = 20.08
Prob > F = 0.0000
R-squared = 0.3643
Root MSE = .79826

Number of clusters (NUMECOLE) = 79

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.6189153	.0465426	13.30	0.000	.5262561	.7115745
FILLE	-.1174956	.0610674	-1.92	0.058	-.2390714	.0040803
AGEPLUS	.1297407	.0715617	1.81	0.074	-.0127278	.2722092
PAUVR	-.2594171	.0858191	-3.02	0.003	-.4302697	-.0885645
REDAN1	-.1449715	.083464	-1.74	0.086	-.3111356	.0211926
REDAN2	-.1974338	.0724664	-2.72	0.008	-.3417034	-.0531643
FPI1ANPL	-.2632908	.2208257	-1.19	0.237	-.7029211	.1763394
BACPLUS	-.1016093	.1794924	-0.57	0.573	-.4589513	.2557327
RESTENS_EST	.2520473	.1393897	1.81	0.074	-.0254564	.529551
COLLEGUES	.246149	.1386352	1.78	0.080	-.0298525	.5221505
SALDUR	-.2018074	.166037	-1.22	0.228	-.5323616	.1287468
DOUBLFLX	-.4507422	.1711348	-2.63	0.010	-.7914454	-.110039
COVAR	.5958115	.2638011	2.26	0.027	.0706239	1.120999
ANCDIR	.0177474	.0073505	2.41	0.018	.0031136	.0323812
APEACTIV	-.2316051	.1443782	-1.60	0.113	-.5190401	.0558299
PROPC_ES	-.0537807	.3658766	-0.15	0.884	-.782185	.6746236
RURAL	.1181083	.1488751	0.79	0.430	-.1782793	.4144958
_cons	-.3279547	.357873	-0.92	0.362	-1.040425	.3845158

```
. tab CONTRACTUEL FPI1ANPL
```

CONTRACTUEL L	FPI1ANPL		Total
	0	1	
0	90	643	733
1	1,095	90	1,185
Total	1,185	733	1,918

Test de l'impact de la formation professionnelle initiale longue chez les contractuels en 5ème année au Mali

```
. reg STFIN5FM STINI5FM FILLE AGEPLUS RICH /*
> */ REDAN4 /*
> */ PROCHECOLE FPI1ANPL BACPLUS /*
> */ DOUBLFLX REUMOIS ANCDIR /*
> */ APEACTIV RURAL if VALEX==0&CONTRACTUEL==1, cluster (NUMECOLE)
```

```
Linear regression                               Number of obs =      760
                                                F( 13,    57) =    19.54
                                                Prob > F      =    0.0000
                                                R-squared     =    0.4807
                                                Root MSE     =    .7656

Number of clusters (NUMECOLE) = 58
```

STFIN5FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI5FM	.5865311	.0502206	11.68	0.000	.485966	.6870962
FILLE	.0278863	.065732	0.42	0.673	-.1037398	.1595124
AGEPLUS	-.1611539	.0885187	-1.82	0.074	-.3384096	.0161017
RICH	.0689672	.0989787	0.70	0.489	-.1292341	.2671686
REDAN4	-.1836277	.1196136	-1.54	0.130	-.4231499	.0558945
PROCHECOLE	.433609	.1766923	2.45	0.017	.0797888	.7874292
FPI1ANPL	-.0891023	.1923131	-0.46	0.645	-.4742027	.2959982
BACPLUS	-.1457644	.1913633	-0.76	0.449	-.5289628	.2374339
DOUBLFLX	-.2917555	.1751981	-1.67	0.101	-.6425837	.0590726
REUMOIS	.3855427	.1660235	2.32	0.024	.0530863	.7179991
ANCDIR	.0091244	.0088781	1.03	0.308	-.0086538	.0269025
APEACTIV	-.0213151	.1527264	-0.14	0.889	-.3271444	.2845142
RURAL	-.0599904	.1634187	-0.37	0.715	-.3872306	.2672498
_cons	-.2708849	.1862776	-1.45	0.151	-.6438992	.1021295

```
.
end of do-file
```

```
. tab CONTRACTUEL FPI1ANPL
```

CONTRACTUEL L	FPI1ANPL		Total
	0	1	
0	117	870	987
1	633	141	774
Total	750	1,011	1,761

ANNEXE 3.3.1 : Les variances inter-élèves et inter-classes dans les modèles multiniveau

```
. xtmixed STFIN2FM if VALEX==0 ||NUMECOLE:, variance
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log restricted-likelihood = -2004.5194
 Iteration 1: log restricted-likelihood = -2004.5194

Computing standard errors:

Mixed-effects REML regression	Number of obs	=	1662	
Group variable: NUMECOLE	Number of groups	=	128	
	Obs per group: min	=	6	
	avg	=	13.0	
	max	=	15	
Log restricted-likelihood = -2004.5194	Wald chi2(0)	=	.	
	Prob > chi2	=	.	

STFIN2FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_cons	.017775	.0615821	0.29	0.773	-.1029237 .1384738

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
NUMECOLE: Identity			
var(_cons)	.4428537	.0609969	.3380798 .5800979
var(Residual)	.5402844	.0195101	.5033671 .5799093

LR test vs. linear regression: chibar2(01) = 673.81 Prob >= chibar2 = 0.0000

```
. xtmixed STFIN5FM if VALEX==0 ||NUMECOLE:, variance
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log restricted-likelihood = -1736.8894
 Iteration 1: log restricted-likelihood = -1736.8894

Computing standard errors:

Mixed-effects REML regression	Number of obs	=	1730	
Group variable: NUMECOLE	Number of groups	=	134	
	Obs per group: min	=	8	
	avg	=	12.9	
	max	=	15	
Log restricted-likelihood = -1736.8894	Wald chi2(0)	=	.	
	Prob > chi2	=	.	

STFIN5FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_cons	-.0147114	.0703218	-0.21	0.834	-.1525396	.1231167

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLE: Identity				
var(_cons)	.6357643	.0814983	.4945168	.817356
var(Residual)	.3394472	.0120192	.3166889	.363841

LR test vs. linear regression: chibar2(01) = 1334.12 Prob >= chibar2 = 0.0000

ANNEXE 4.1.1 : Mesure de la variance au niveau classe et école

La part de variance expliquée par les variables muettes, des différentes classes et/ou écoles de l'échantillon, est mesurée dans le cadre d'un modèle MCO avec le R^2 ajusté. Pour chaque pays, le coefficient de corrélation intra-classe a également été calculé pour s'assurer de la fiabilité de la mesure. Ci-dessous l'exemple du Burkina Faso. Il y a 120 écoles dans l'échantillon, le PASEC n'enquêtant qu'une seule classe par école, on ne peut mesurer qu'un effet classe.

reg STFIN5M ECOLE1-ECOLE120

Source	SS	df	MS	Number of obs =	2276
Model	837.827687	119	7.0405688	F(119, 2156) =	10.56
Residual	1437.17227	2156	.666591962	Prob > F =	0.0000
				R-squared =	0.3683
				Adj R-squared =	0.3334
Total	2274.99996	2275	.999999981	Root MSE =	.81645

STFIN5M	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ECOLE1	-.4465133	.2938709	-1.52	0.129	-1.022813 .1297866
ECOLE2	-1.225313	.3008116	-4.07	0.000	-1.815224 -.6354019
ECOLE3	-.4283994	.2908709	-1.47	0.141	-.9988162 .1420174
ECOLE4	-.4144938	.2971686	-1.39	0.163	-.9972608 .1682732
ECOLE5	-1.241877	.2908709	-4.27	0.000	-1.812294 -.6714601
ECOLE6	-.7099878	.2908709	-2.44	0.015	-1.280405 -.139571
ECOLE7	.3603774	.2938709	1.23	0.220	-.2159225 .9366772
ECOLE8	-.2146008	.2971686	-0.72	0.470	-.7973678 .3681661
ECOLE9	1.159446	.2908709	3.99	0.000	.5890293 1.729863
ECOLE10	-.5100948	.2971686	-1.72	0.086	-1.092862 .0726722
ECOLE11	.1504211	.2908709	0.52	0.605	-.4199957 .7208379
ECOLE12	1.534897	.2908709	5.28	0.000	.9644805 2.105314
ECOLE13	-.0348344	.2938709	-0.12	0.906	-.6111342 .5414655
ECOLE14	-.3806446	.2938709	-1.30	0.195	-.9569445 .1956552
ECOLE15	.1678031	.2971686	0.56	0.572	-.4149638 .7505701
ECOLE16	-.0060168	.2908709	-0.02	0.983	-.5764336 .5643999
ECOLE17	-.3394767	.2938709	-1.16	0.248	-.9157766 .2368231
ECOLE18	-.2641395	.2908709	-0.91	0.364	-.8345563 .3062773
ECOLE19	.4320095	.2908709	1.49	0.138	-.1384073 1.002426
ECOLE20	-.3362748	.2971686	-1.13	0.258	-.9190418 .2464922
ECOLE21	-.9240608	.2938709	-3.14	0.002	-1.500361 -.3477609
ECOLE22	-1.088732	.2938709	-3.70	0.000	-1.665032 -.5124325
ECOLE23	-1.325962	.3048578	-4.35	0.000	-1.923808 -.7281163
ECOLE24	.4806791	.2971686	1.62	0.106	-.1020879 1.063446
ECOLE25	-.4127556	.2908709	-1.42	0.156	-.9831724 .1576612
ECOLE26	-.1702767	.2908709	-0.59	0.558	-.7406935 .4001401
ECOLE27	.3768445	.2938709	1.28	0.200	-.1994553 .9531444
ECOLE28	.0722021	.2908709	0.25	0.804	-.4982146 .6426189
ECOLE29	-.3065424	.2938709	-1.04	0.297	-.8828423 .2697574
ECOLE30	.158243	.2908709	0.54	0.586	-.4121738 .7286598
ECOLE31	-.7334534	.2908709	-2.52	0.012	-1.30387 -.1630367
ECOLE32	.9284943	.2938709	3.16	0.002	.3521944 1.504794
ECOLE33	-.6111848	.2938709	-2.08	0.038	-1.187485 -.034885
ECOLE34	-.0373044	.2908709	-0.13	0.898	-.6077212 .5331123
ECOLE35	-.3267147	.2908709	-1.12	0.261	-.8971315 .2437021
ECOLE36	.6979541	.2908709	2.40	0.017	.1275373 1.268371
ECOLE37	-1.222757	.2971686	-4.11	0.000	-1.805524 -.6399897
ECOLE38	.2634041	.2971686	0.89	0.376	-.3193628 .8461711
ECOLE39	.3616124	.2908709	1.24	0.214	-.2088044 .9320292
ECOLE40	-.7017542	.2938709	-2.39	0.017	-1.278054 -.1254543
ECOLE41	-.6770534	.2938709	-2.30	0.021	-1.253353 -.1007536
ECOLE42	1.620938	.2908709	5.57	0.000	1.050521 2.191355

ECOLE43	.1550222	.3008116	0.52	0.606	-.4348887	.7449332
ECOLE44	.578325	.3008116	1.92	0.055	-.011586	1.168236
ECOLE45	.2912153	.2908709	1.00	0.317	-.2792015	.8616321
ECOLE46	1.44538	.2971686	4.86	0.000	.862613	2.028147
ECOLE47	-.6395907	.2908709	-2.20	0.028	-1.210007	-.0691739
ECOLE48	-.9098807	.2971686	-3.06	0.002	-1.492648	-.3271138
ECOLE49	.0804357	.2938709	0.27	0.784	-.4958641	.6567356
ECOLE50	-.4629804	.2938709	-1.58	0.115	-1.03928	.1133194
ECOLE51	-1.692071	.2971686	-5.69	0.000	-2.274838	-1.109304
ECOLE52	-.6926058	.2971686	-2.33	0.020	-1.275373	-.1098388
ECOLE53	.6119132	.2908709	2.10	0.036	.0414964	1.18233
ECOLE54	-.4049337	.2908709	-1.39	0.164	-.9753505	.1654831
ECOLE55	-.5691936	.2908709	-1.96	0.050	-1.13961	.0012232
ECOLE56	.3537905	.2908709	1.22	0.224	-.2166263	.9242073
ECOLE57	-.5926593	.2908709	-2.04	0.042	-1.163076	-.0222425
ECOLE58	.8543921	.2908709	2.94	0.003	.2839753	1.424809
ECOLE59	-.2324402	.2938709	-0.79	0.429	-.8087401	.3438596
ECOLE60	.5258723	.2908709	1.81	0.071	-.0445445	1.096289
ECOLE61	.2051744	.2908709	0.71	0.481	-.3652424	.7755912
ECOLE62	-.1830388	.2938709	-0.62	0.533	-.7593386	.3932611
ECOLE63	.4059365	.3093797	1.31	0.190	-.2007771	1.01265
ECOLE64	.1738868	.2908709	0.60	0.550	-.39653	.7443036
ECOLE65	-.2406738	.2971686	-0.81	0.418	-.8234408	.3420931
ECOLE66	.3068591	.2908709	1.05	0.292	-.2635577	.8772759
ECOLE67	.0310343	.2938709	0.11	0.916	-.5452656	.6073341
ECOLE68	.7226548	.2938709	2.46	0.014	.1463549	1.298955
ECOLE69	.9638986	.2908709	3.31	0.001	.3934818	1.534315
ECOLE70	-.2719614	.2908709	-0.93	0.350	-.8423782	.2984554
ECOLE71	.1660649	.2908709	0.57	0.568	-.4043518	.7364817
ECOLE72	-1.171068	.2938709	-3.98	0.000	-1.747368	-.5947683
ECOLE73	.845701	.2971686	2.85	0.004	.2629341	1.428468
ECOLE74	-.1189998	.2971686	-0.40	0.689	-.7017668	.4637671
ECOLE75	-.5100948	.2971686	-1.72	0.086	-1.092862	.0726722
ECOLE76	.0252707	.2908709	0.09	0.931	-.545146	.5956875
ECOLE77	-.2736081	.2938709	-0.93	0.352	-.849908	.3026917
ECOLE78	.0548201	.2971686	0.18	0.854	-.5279468	.6375871
ECOLE79	-.7412754	.2908709	-2.55	0.011	-1.311692	-.1708586
ECOLE80	-.5123819	.2938709	-1.74	0.081	-1.088682	.063918
ECOLE81	-.2653745	.2938709	-0.90	0.367	-.8416744	.3109253
ECOLE82	.2981681	.2971686	1.00	0.316	-.2845988	.8809351
ECOLE83	-.3065424	.2938709	-1.04	0.297	-.8828423	.2697574
ECOLE84	-.6474126	.2908709	-2.23	0.026	-1.217829	-.0769958
ECOLE85	-.0842358	.3008116	-0.28	0.779	-.6741468	.5056752
ECOLE86	.9247891	.2908709	3.18	0.001	.3543724	1.495206
ECOLE87	.0956678	.2908709	0.33	0.742	-.4747489	.6660846
ECOLE88	-.6787002	.2908709	-2.33	0.020	-1.249117	-.1082834
ECOLE89	.2562468	.3008116	0.85	0.394	-.3336642	.8461578
ECOLE90	.5649818	.2908709	1.94	0.052	-.005435	1.135399
ECOLE91	-.5222622	.2908709	-1.80	0.073	-1.092679	.0481546
ECOLE92	.1156571	.2971686	0.39	0.697	-.4671098	.6984241
ECOLE93	.1051365	.2938709	0.36	0.721	-.4711634	.6814363
ECOLE94	.4945847	.2908709	1.70	0.089	-.0758321	1.065001
ECOLE95	-.1389891	.2908709	-0.48	0.633	-.7094059	.4314277
ECOLE96	-.7273697	.2971686	-2.45	0.014	-1.310137	-.1446028
ECOLE97	.3929	.2908709	1.35	0.177	-.1775168	.9633168
ECOLE98	-.5848374	.2908709	-2.01	0.044	-1.155254	-.0144206
ECOLE99	-.8586038	.2908709	-2.95	0.003	-1.429021	-.2881871
ECOLE100	.0145671	.2938709	0.05	0.960	-.5617328	.590867
ECOLE101	-.8898914	.2908709	-3.06	0.002	-1.460308	-.3194746
ECOLE102	-.8020101	.3008116	-2.67	0.008	-1.391921	-.2120991
ECOLE103	.4789409	.2908709	1.65	0.100	-.0914759	1.049358
ECOLE104	.0330927	.2908709	0.11	0.909	-.5373241	.6035094
ECOLE105	-.1780986	.2908709	-0.61	0.540	-.7485154	.3923182
ECOLE106	-1.146367	.2938709	-3.90	0.000	-1.722667	-.5700675
ECOLE107	-.0183672	.2938709	-0.06	0.950	-.5946671	.5579327
ECOLE108	.2981681	.2971686	1.00	0.316	-.2845988	.8809351
ECOLE109	.1738868	.2908709	0.60	0.550	-.39653	.7443036
ECOLE110	(dropped)					

ECOLE111	.6119132	.2908709	2.10	0.036	.0414964	1.18233
ECOLE112	-.3241074	.3093797	-1.05	0.295	-.930821	.2826062
ECOLE113	.0804357	.2938709	0.27	0.784	-.4958641	.6567356
ECOLE114	.5649818	.2908709	1.94	0.052	-.005435	1.135399
ECOLE115	-.1885278	.3093797	-0.61	0.542	-.7952414	.4181858
ECOLE116	-.0668538	.2971686	-0.22	0.822	-.6496208	.5159131
ECOLE117	-.1016178	.2971686	-0.34	0.732	-.6843848	.4811491
ECOLE118	.2677496	.2908709	0.92	0.357	-.3026672	.8381664
ECOLE119	-.5361678	.2971686	-1.80	0.071	-1.118935	.0465992
ECOLE120	-.1670559	.3008116	-0.56	0.579	-.7569669	.4228551
_cons	.0871914	.2264427	0.39	0.700	-.3568775	.5312603

```
. xtmixed STFIN5M || NUMECOLE: , variance
```

Performing EM optimization:

Performing gradient-based optimization:

Iteration 0: log restricted-likelihood = -2910.5774
 Iteration 1: log restricted-likelihood = -2910.5774

Computing standard errors:

Mixed-effects REML regression	Number of obs	=	2276
Group variable: NUMECOLE	Number of groups	=	120
	Obs per group: min	=	13
	avg	=	19.0
	max	=	20

Log restricted-likelihood = -2910.5774	Wald chi2(0)	=	.
	Prob > chi2	=	.

STFIN5M	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_cons	-.0056267	.0555662	-0.10	0.919	-.1145345 .1032812

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]
NUMECOLE: Identity			
var(_cons)	.3352004	.0480256	.2531333 .4438741
var(Residual)	.6665771	.0203016	.627951 .7075792

LR test vs. linear regression: chibar2(01) = 642.75 Prob >= chibar2 = 0.0000

```
. xtmrho
```

Levels: NUMECOLE

Intra class correlation on level 1 rho = .33460562

Le R^2 ajusté des MCO avec variables muettes et le rho calculé à partir des modèles multiniveaux constituent tous les deux une mesure de la variabilité au niveau classe. On observe dans le tableau suivant relatif aux données PASEC que les résultats obtenus sont très proches.

Part de la variance du score de mathématiques de 5^{ème} année expliquée au niveau classe avec les données PASEC :

Pays	R ² ajusté	Rho	Nombre d'écoles et de classe	Nombre d'élèves
Burkina Faso (1996)	33,3 %	33,4 %	120	2392
Cameroun (1996)	47,1 %	46,4 %	117	2660
Côte d'Ivoire (1996)	40 %	39,9 %	120	2289
Madagascar (1998)	48,3 %	50,1 %	119	2878
Mali (2002)	63,4 %	63,4 %	134	1761
Niger (2002)	33,5 %	33,8 %	119	1503
Sénégal (1996)	26 %	26,3 %	93	1722
Tchad (2004)	55,2 %	56,8 %	110	1589
Mauritanie (2004)	59,6 %	59,5 %	121	1713
Moyenne	45,2 %	45,5 %		

Compte tenu de la proximité de la mesure, on s'est contenté du R^2 ajusté pour les autres sources de données

**Données SACMEQ II (2000)
(6^{ème} année, mathématiques)**

	R ² ajusté niveau classe	R ² ajusté niveau école	Nombre de classes	Nombre d'écoles	Nombre d'élèves
Botswana	25,1 %	21,9 %	473	170	3321
Kenya	38,1 %	35,6 %	320	185	3296
Lesotho	30,9 %	29,9 %	250	177	3144
Malawi	15 %	15,1 %	197	140	2323
Maurice	37,4 %	24,8 %	525	153	2870
Mozambique	24,9 %	20,4 %	1058	176	3136
Namibie	55,7 %	54,8 %	606	270	4990
Seychelles	61,9 %	7,5 %	60	24	1480
Afrique du sud	64,8 %	63,7 %	399	169	3135
Swaziland	25,8 %	25,2 %	295	168	3138
Tanzanie	26,4 %	24,9 %	313	181	2849
Ouganda	62,4 %	/	163	163	2619
Zambie	23 %	20,3 %	430	173	2590
Zanzibar	34,4 %	33 %	360	145	2459
moyenne	37,6 %	29 %			

Données LLECE 1997
(4^{ème} année, mathématiques)

	R² ajusté niveau école	Nombre d'écoles	Nombre d'élèves
Argentine	46,4 %	113	2119
Bolivie	39,2 %	67	2411
Brésil	35,4 %	118	2162
Chili	24,1 %	129	2310
Colombie	28,4 %	181	2146
Cuba	37,7 %	100	2001
Honduras	39,1 %	109	1831
Mexique	28 %	133	2545
Paraguay	38,1 %	124	2329
Pérou	46,4 %	116	2099
République dominicaine	31,4 %	93	1949
Vénézuéla	26,6 %	112	1777
moyenne	35,1 %		

TIMSS 2003
(4^{ème} année mathématiques)

	R² ajusté niveau classe	R² ajusté niveau école	Nombre de classes	Nombre d'écoles	Nombre d'élèves
Australie	26,4 %	25,1 %	222	204	4321
Belgique	15,2 %	14,1 %	258	149	4712
Angleterre	25 %	23 %	149	123	3585
Hong Kong	27,8 %	24,5 %	145	132	4608
Italie	34,7 %	29,8 %	237	171	4282
Japon	5,8 %	/	150	150	4535
Hollande	15,3 %	14,5 %	141	130	2935
Nouvelle-Zélande	36 %	31,2 %	332	220	4308
Norvège	11,3 %	9 %	228	139	4342
Etats Unis	33,4 %	28,8 %	479	248	9829
Ecosse	19,2 %	17,8 %	173	125	3936
Moyenne	22,7 %	21,8 %			
Maroc	34,6 %	/	197	197	4264
Tunisie	35,2 %	/	150	150	4334

ANNEXE 4.1.2 : La mesure de l'effet classe

La part de la variance totale expliquée par les variables muettes « classe » est calculée en faisant la différence du R² entre le modèle global et le même modèle avec les variables muettes. Cette différence est faite en utilisant les R² ajusté pour éviter les surestimations du R² liées à l'introduction des nombreuses variables dont les variables muettes « classe ». On observe également la part expliquée par les variables caractérisant les enseignants en retirant ces dernières du modèle. Les scores utilisés (début et fin d'année) sont des scores standardisés agrégés de français et mathématiques. Ces traitements ont été réalisés pour 11 pays en 2^{ème} et 5^{ème} années.

Ci-dessous un extrait du programme utilisé pour la 2^{ème} année au Burkina Faso.

1) Estimation du modèle global

```
reg STFIN2FM STINI2FM FILLE AGEMOINS AGEPLUS NIVEAUVI REDOUBLE LIV_FRMT DOMFRANC/*
*/ AIDE TRAVAUX DOUBLFLX MULTIGRA TCLASSE RURAL MAITRFEM SERVICE FPI1_3M FPI1AN /*
*/ FPI1ANPL FORCON1 FORCONPL INST_ADJT INST_ADJT_CERT INST_PRINC INST_CERT /*
*/ NIVCYCLB DIPCYCLB NIVCYCLA
```

Source	SS	df	MS	Number of obs =	2091
Model	993.606321	28	35.48594	F(28, 2062) =	68.63
Residual	1066.14936	2062	.517046244	Prob > F =	0.0000
				R-squared =	0.4824
				Adj R-squared =	0.4754
Total	2059.75568	2090	.985529032	Root MSE =	.71906

STFIN2FM	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
STINI2FM	.6665322	.0170596	39.07	0.000	.6330764 .699988
FILLE	.0114317	.035408	0.32	0.747	-.0580076 .0808709
AGEMOINS	.0300942	.0718	0.42	0.675	-.1107138 .1709022
AGEPLUS	-.0643967	.0404269	-1.59	0.111	-.1436786 .0148852
NIVEAUVI	.0451653	.0236567	1.91	0.056	-.0012282 .0915588
REDOUBLE	-.2074484	.0394418	-5.26	0.000	-.2847983 -.1300984
LIV_FRMT	.0863378	.0665001	1.30	0.194	-.0440766 .2167521
DOMFRANC	.0689213	.0426868	1.61	0.107	-.0147924 .1526349
AIDE	.0088552	.0344978	0.26	0.797	-.0587989 .0765093
TRAVAUX	-.0017601	.0094653	-0.19	0.852	-.0203226 .0168024
DOUBLFLX	-.0180558	.04663	-0.39	0.699	-.1095027 .0733911
MULTIGRA	-.0305297	.1011485	-0.30	0.763	-.2288935 .1678341
TCLASSE	-.0024884	.0006806	-3.66	0.000	-.0038232 -.0011536
RURAL	-.0096747	.0450207	-0.21	0.830	-.0979655 .078616
MAITRFEM	.1040461	.0348954	2.98	0.003	.0356122 .17248
SERVICE	.0101732	.0068129	1.49	0.136	-.0031876 .023534
FPI1_3M	-.0025085	.1352021	-0.02	0.985	-.2676554 .2626383
FPI1AN	.0226182	.0461506	0.49	0.624	-.0678886 .1131249
FPI1ANPL	.177549	.0688178	2.58	0.010	.0425893 .3125087
FORCON1	-.1851365	.0573192	-3.23	0.001	-.297546 -.072727
FORCONPL	-.0085062	.0531467	-0.16	0.873	-.112733 .0957207
INST_ADJT	.4984457	.1280739	3.89	0.000	.2472781 .7496133
INST_ADJT~T	.3901295	.1206378	3.23	0.001	.1535448 .6267142
INST_PRINC	.4898807	.1868825	2.62	0.009	.1233825 .8563788
INST_CERT	.3015546	.1224767	2.46	0.014	.0613638 .5417454
NIVCYCLB	.2963956	.1150173	2.58	0.010	.0708333 .5219578
DIPCYCLB	.3545349	.1239697	2.86	0.004	.1114161 .5976536
NIVCYCLA	.397747	.1214716	3.27	0.001	.1595273 .6359668
_cons	-.5351958	.1792204	-2.99	0.003	-.8866675 -.183724

2) Estimation du modèle global avec les variables muettes « classe »

```
reg STFIN2FM STINI2FM FILLE AGEMOINS AGEPLUS NIVEAUVI REDOUBLE LIV_FRMT /*DOMFRANC
*/AIDE TRAVAUX DOUBLFLX MULTIGRA TCLASSE RURAL MAITRFEM SERVICE FPI1_3M FPI1AN /*
*/FPI1ANPL FORCON1 FORCONPL INST_ADJT INST_ADJT_CERT INST_PRINC INST_CERT /*
*/ NIVCYCLB DIPCYCLB NIVCYCLA ECOLE*
```

Source	SS	df	MS	Number of obs =	2091
Model	1302.84824	126	10.3400654	F(126, 1964) =	26.83
Residual	756.907438	1964	.385390753	Prob > F =	0.0000
				R-squared =	0.6325
				Adj R-squared =	0.6090
Total	2059.75568	2090	.985529032	Root MSE =	.6208

STFIN2FM	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
STINI2FM	.7198353	.0196833	36.57	0.000	.681233 .7584375
FILLE	-.0088844	.0323499	-0.27	0.784	-.0723282 .0545593
AGEMOINS	.0068011	.0666635	0.10	0.919	-.1239376 .1375397
AGEPLUS	-.0479989	.0368516	-1.30	0.193	-.1202712 .0242733
NIVEAUVI	.0112214	.0218154	0.51	0.607	-.0315624 .0540053
REDOUBLE	-.2050026	.0353924	-5.79	0.000	-.2744131 -.1355921
LIV_FRMT	.0048034	.064526	0.07	0.941	-.1217432 .1313501
DOMFRANC	-.0219933	.0424691	-0.52	0.605	-.1052825 .0612958
AIDE	.0474752	.032223	1.47	0.141	-.0157196 .1106701
TRAVAUX	.0012094	.0092375	0.13	0.896	-.016907 .0193257
DOUBLFLX	.4624523	.2118688	2.18	0.029	.0469409 .8779636
MULTIGRA	-.827968	.1929335	-4.29	0.000	-1.206344 -.4495921
TCLASSE	-.0186462	.002093	-8.91	0.000	-.022751 -.0145415
RURAL	.9388921	.1346701	6.97	0.000	.6747808 1.203003
MAITRFEM	.0109425	.1094297	0.10	0.920	-.203668 .225553
SERVICE	-.0624802	.0145102	-4.31	0.000	-.0909371 -.0340233
FPI1_3M	.2495687	.3631955	0.69	0.492	-.4627202 .9618577
FPI1AN	-.2438752	.1179201	-2.07	0.039	-.475137 -.0126135
FPI1ANPL	.8489379	.1336862	6.35	0.000	.5867563 1.111112
FORCON1	-.3180888	.1401881	-2.27	0.023	-.5930219 -.0431557
FORCONPL	-.4004742	.1353935	-2.96	0.003	-.6660041 -.1349442
INST_ADJT	-1.180187	.2815236	-4.19	0.000	-1.732304 -.628071
INST_ADJT~T	.0119681	.1868646	0.06	0.949	-.3545057 .3784419
INST_PRINC	.6322147	.3265047	1.94	0.053	-.0081173 1.272547
INST_CERT	.1398744	.1887768	0.74	0.459	-.2303495 .5100982
NIVCYCLB	-.9058291	.2226229	-4.07	0.000	-1.342431 -.4692272
DIPCYCLB	-1.322548	.2513834	-5.26	0.000	-1.815554 -.8295419
NIVCYCLA	-.6643517	.3164931	-2.10	0.036	-1.285049 -.0436542
ECOLE1	-.2023319	.2946408	-0.69	0.492	-.7801734 .3755097
ECOLE2	-.3595741	.3118125	-1.15	0.249	-.9710923 .2519441
ECOLE3	-.109834	.3131618	-0.35	0.726	-.7239984 .5043304
ECOLE4	.2992031	.3074503	0.97	0.331	-.30376 .9021663
ECOLE5	-1.566788	.2874993	-5.45	0.000	-2.130623 -1.002952
ECOLE6	-.6856571	.2187544	-3.13	0.002	-1.114672 -.2566419
ECOLE7	1.064392	.245998	4.33	0.000	.5819479 1.546837
ECOLE8	.4389262	.2383078	1.84	0.066	-.0284365 .906289
ECOLE9	.446562	.2301724	1.94	0.053	-.0048458 .8979699
ECOLE10	-.1596771	.2861214	-0.56	0.577	-.7208105 .4014564
ECOLE11	.3559152	.2669373	1.33	0.183	-.1675949 .8794253
ECOLE12	1.406262	.2885124	4.87	0.000	.8404391 1.972084
ECOLE13	.9569761	.2984506	3.21	0.001	.3716628 1.542289
ECOLE14	1.037533	.2209019	4.70	0.000	.604306 1.47076
ECOLE15	.0951495	.2666032	0.36	0.721	-.4277054 .6180043
ECOLE16	1.103113	.2840156	3.88	0.000	.5461096 1.660117

...

3) Estimation du modèle global sans les variables « maître »

```
reg STFIN2FM STINI2FM FILLE AGEMOINS AGEPLUS NIVEAUVI REDOUBLE LIV_FRMT /*
*/ DOMFRANC AIDE TRAVAUX DOUBLFLX MULTIGRA TCLASSE RURAL
```

Source	SS	df	MS	Number of obs =	2091
Model	952.889972	14	68.0635694	F(14, 2076) =	127.66
Residual	1106.8657	2076	.533172305	Prob > F =	0.0000
				R-squared =	0.4626
				Adj R-squared =	0.4590
Total	2059.75568	2090	.985529032	Root MSE =	.73019

STFIN2FM	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
STINI2FM	.6549075	.0167349	39.13	0.000	.6220885 .6877264
FILLE	.0028208	.0356859	0.08	0.937	-.067163 .0728045
AGEMOINS	.0285174	.072292	0.39	0.693	-.113255 .1702898
AGEPLUS	-.0519468	.0409086	-1.27	0.204	-.132173 .0282793
NIVEAUVI	.0464186	.0237039	1.96	0.050	-.0000673 .0929046
REDOUBLE	-.2201596	.039921	-5.51	0.000	-.2984491 -.1418702
LIV_FRMT	.1176167	.0671752	1.75	0.080	-.014121 .2493544
DOMFRANC	.0713646	.0426637	1.67	0.095	-.0123034 .1550326
AIDE	.022353	.0346026	0.65	0.518	-.0455064 .0902123
TRAVAUX	.0050319	.0094304	0.53	0.594	-.0134621 .0235258
DOUBLFLX	-.0188878	.0415013	-0.46	0.649	-.1002762 .0625007
MULTIGRA	-.2720846	.0912167	-2.98	0.003	-.4509703 -.0931989
TCLASSE	-.0031372	.0006776	-4.63	0.000	-.0044661 -.0018082
RURAL	-.0245958	.0409331	-0.60	0.548	-.1048701 .0556784
_cons	.2975507	.0728597	4.08	0.000	.154665 .4404364

Le calcul de l'effet classe pour les 11 pays du PASEC

	Burkina Faso		Cameroun		Côte d'Ivoire		Guinée		Madagascar		Mali		Mauritanie		Niger		Sénégal		Tchad		Togo	
	2A	5A	2A	5A	2A	5A	2A	5A	2A	5A	2A	5A	2AM	5FM	2A	5A	2A	5A	2A	5A	2A	5A
R ² score initial (A'')	43,9	45,8	28,1	34	41,9	46,5	35,8	31,8	34	22,4	18,3	29,1	34,5	22,1	28,6	32,2	50,7	53,1	20,8	26,4	41,4	47,8
R ² score initial et Caractéristiques élève, classe et école (A')	45,9	49,1	33,2	35,7	45,5	49,7	40	35,4	37,6	29	27,5	35,8	39,2	34,4	33,1	33	52,7	55,7	31	37,7	41,3	55,1
R ² modèle global (A)	47,5	52	36,2	40,9	47	51,6	43,6	39,8	42,1	31,6	31	38,3	40,9	40,2	38,1	36,6	53,5	57	39,4	41,9	43,6	56,6
R ² modèle global avec les variables muettes classe (B)	60,9	67	55,8	65,8	62,9	63,1	68,9	80,5	69,8	68,3	62,8	75,4	53,3	71,6	63,5	58,9	66,7	63,1	69,2	71,6	60,2	68
Effet classe (B-A)	13,4	15	19,6	24,9	15,9	11,5	25,3	40,7	27,7	36,7	31,8	37,1	12,4	31,4	25,4	22,3	13,2	6,1	29,8	29,7	16,6	11,4
% de variance expliquée par les variables « maître » (B-A')	1,6	3,6	3,0	5,2	1,5	1,9	3,6	4,4	4,5	2,6	3,5	2,5	1,7	5,8	5,0	3,6	0,8	1,3	8,4	4,2	2,3	1,5
% de variance expliquée par l'ensemble des variables explicatives (A-A'')	3,6	6,2	8,1	6,9	5,1	5,1	7,8	8	8,1	9,2	12,1	9,2	6,4	18,1	9,5	4,4	2,8	3,9	18,6	15,5	2,2	8,8

Part de variance expliquée par les variables caractérisant les enseignants dans les modèles à valeur ajoutée sur les données du PASEC

Pays	2^{ème} année	5^{ème} année	Ensemble
Burkina Faso	1,6 %	2,9 %	2,3 %
Cameroun	3 %	5,2 %	4,1 %
Côte d'Ivoire	1,5 %	1,9 %	1,7 %
Guinée	3,6 %	4,4 %	4,0 %
Madagascar	4,5 %	2,6 %	3,6 %
Mali	3,5 %	2,5 %	1,7 %
Mauritanie	1,7 %	5,8 %	3,8 %
Niger	5 %	3,6 %	4,3 %
Sénégal	0,8 %	1,3 %	1,1 %
Tchad	8,4 %	4,2 %	6,3 %
Togo	2,3 %	1,5 %	1,9 %
Moyenne	3,1 %	3,2 %	3,1 %

ANNEXE 4.1.3 : Test de la relation entre l'effet-classe et le nombre d'élèves par classe retenus dans les enquêtes

La possibilité d'une sensibilité de la mesure de l'effet-classe à la variabilité de l'effectif d'élèves testés selon les classes est à considérer. Les vérifications opérées à partir des données PASEC et SACMEQ montrent que cette hypothèse n'est pas avérée. Le graphique qui suit illustre l'absence de relation entre l'effet-classe et la variabilité du nombre d'élèves testés.

Données PASEC

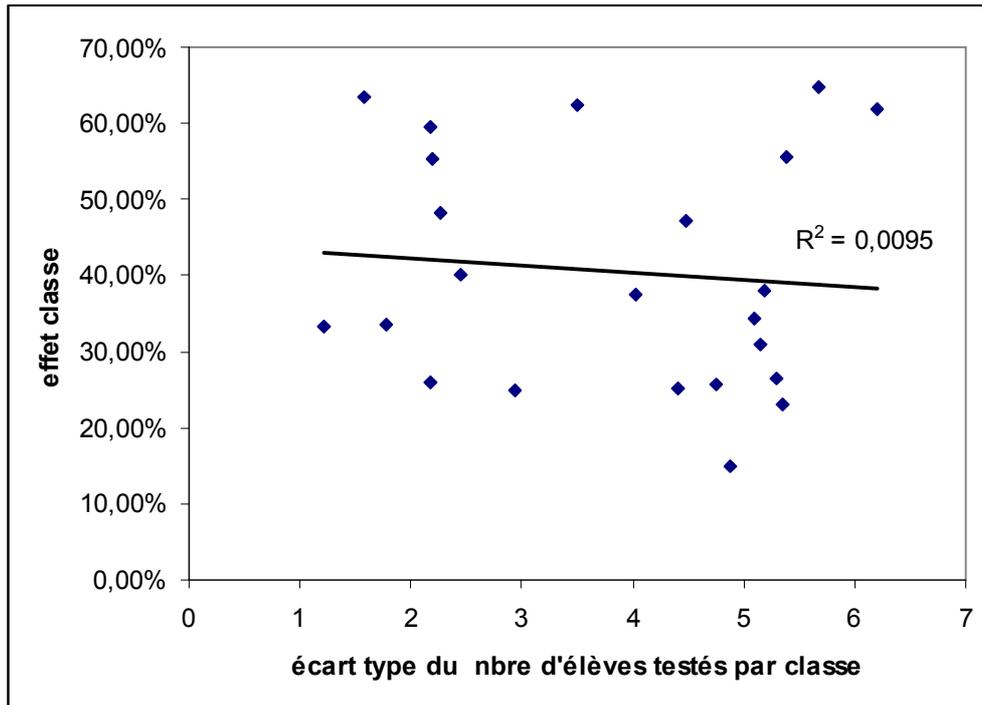
Pays	Ecart type du nbre d'élèves testés par classe	Effet classe
Burkina Faso (1996)	1,21	33,3 %
Cameroun (1996)	4,48	47,1 %
Côte d'Ivoire (1996)	2,45	40 %
Madagascar (1998)	2,26	48,3 %
Mali (2002)	1,58	63,4 %
Niger (2002)	1,77	33,5 %
Sénégal (1996)	2,18	26 %
Tchad (2004)	2,19	55,2 %
Mauritanie (2004)	2,18	59,6 %
Moyenne	2,26	45,2 %

Données SACMEQ

	Ecart type du Nb d'élèves testés par classe	effet classe
Botswana	4,40	25,1 %
Kenya	5,18	38,1 %
Lesotho	5,15	30,9 %
Malawi	4,87	15 %
Maurice	4,03	37,4 %
Mozambique	2,94	24,9 %
Namibie	5,38	55,7 %
Seychelles	6,21	61,9 %
Afrique du sud	5,67	64,8 %
Swaziland	4,76	25,8 %
Tanzanie	5,30	26,4 %
Ouganda	3,50	62,4 %
Zambie	5,35	23 %
Zanzibar	5,09	34,4 %
moyenne	4,85	37,6 %

La plus grande variabilité observée dans les données SACMEQ vient du fait que les élèves sont tirés au sort dans toutes les classes d'une école. Il y a donc nécessairement des différences liées au tirage au sort d'une part, et au fait que quand il y a plusieurs classes dans une école moins d'élèves par classe sont représentés dans l'étude d'autre part.

La relation entre l'effet classe et la variabilité du nombre d'élèves enquêtés par classe pour les données PASEC et SACMEQ



ANNEXE 4.1.4 : La mise en évidence d'un effet école : Le cas du Mali

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL /*
> */if VALEX==0&M5FM2!=., cluster (NUMECOLE)
```

```
Linear regression                                Number of obs =    1609
                                                F( 17, 123) =    35.50
                                                Prob > F      =    0.0000
                                                R-squared    =    0.3831
                                                Root MSE    =    .78096

Number of clusters (NUMECOLE) = 124
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.6099207	.0348421	17.51	0.000	.5409529	.6788884
FILLE	-.0656581	.0434064	-1.51	0.133	-.1515783	.0202622
AGEPLUS	.1206121	.054491	2.21	0.029	.0127505	.2284737
PAUVR	-.1869253	.0678432	-2.76	0.007	-.3212167	-.0526339
REDAN1	-.1592832	.0638778	-2.49	0.014	-.2857254	-.032841
REDAN2	-.2184684	.0574277	-3.80	0.000	-.3321431	-.1047937
CONTRACTUEL	.2883829	.1040208	2.77	0.006	.0824802	.4942856
BACPLUS	-.1989632	.1386027	-1.44	0.154	-.4733189	.0753924
RESTENS_EST	.1845682	.1005289	1.84	0.069	-.0144225	.3835589
COLLEGUES	.3474388	.1386201	2.51	0.014	.0730488	.6218288
SALDUR	-.2050818	.1124486	-1.82	0.071	-.4276669	.0175033
DOUBLFLX	-.3730446	.1211257	-3.08	0.003	-.6128055	-.1332836
COVAR	.6085393	.2191376	2.78	0.006	.1747698	1.042309
ANCDIR	.0149585	.0054092	2.77	0.007	.0042513	.0256656
APEACTIV	-.171188	.1033373	-1.66	0.100	-.3757379	.033362
PROPC_ES	-.0440021	.2575738	-0.17	0.865	-.5538536	.4658494
RURAL	.0542947	.1129341	0.48	0.632	-.1692514	.2778408
_cons	-.7001887	.2248239	-3.11	0.002	-1.145214	-.2551636

```
. reg STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL N5FM2 /*
> */if VALEX==0, cluster (NUMECOLE)
```

```
Linear regression                                Number of obs =    1609
                                                F( 18, 123) =    29.72
                                                Prob > F      =    0.0000
                                                R-squared    =    0.4293
                                                Root MSE    =    .75142

Number of clusters (NUMECOLE) = 124
```

STFIN2FM	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
STINI2FM	.56211	.0385088	14.60	0.000	.4858842	.6383359
FILLE	-.0719657	.0419057	-1.72	0.088	-.1549155	.0109841
AGEPLUS	.101249	.0529377	1.91	0.058	-.003538	.2060359
PAUVR	-.220723	.0621683	-3.55	0.001	-.3437812	-.0976648
REDAN1	-.1532199	.0580112	-2.64	0.009	-.2680494	-.0383904
REDAN2	-.2174485	.0516482	-4.21	0.000	-.3196829	-.1152141
CONTRACTUEL	.2922369	.0940137	3.11	0.002	.1061426	.4783311
BACPLUS	-.2108525	.1197013	-1.76	0.081	-.4477939	.0260889
RESTENS_EST	.1538992	.0874679	1.76	0.081	-.0192381	.3270364
COLLEGUES	.2967459	.1293688	2.29	0.023	.0406683	.5528235

SALDUR	-.085882	.1007296	-0.85	0.396	-.2852702	.1135061
DOUBLFLX	-.3132428	.1070602	-2.93	0.004	-.5251619	-.1013236
COVAR	.5276424	.1928702	2.74	0.007	.1458677	.9094171
ANCDIR	.0082773	.0050108	1.65	0.101	-.0016412	.0181959
APEACTIV	-.1572303	.0956486	-1.64	0.103	-.3465609	.0321004
PROPC_ES	-.155783	.2369403	-0.66	0.512	-.6247918	.3132258
RURAL	.040387	.1006802	0.40	0.689	-.1589033	.2396772
N5FM2	.2398721	.0579454	4.14	0.000	.1251727	.3545716
_cons	-.5498537	.2117473	-2.60	0.011	-.9689945	-.1307129

. vif

Variable	VIF	1/VIF
PROPC_ES	1.75	0.572540
STINI2FM	1.44	0.692386
RURAL	1.44	0.696553
COVAR	1.40	0.714314
N5FM2	1.20	0.835594
DOUBLFLX	1.19	0.837669
CONTRACTUEL	1.19	0.838946
ANCDIR	1.16	0.862414
PAUVR	1.15	0.868314
COLLEGUES	1.15	0.869916
AGEPLUS	1.11	0.901196
SALDUR	1.11	0.904092
REDAN2	1.09	0.917251
RESTENS_EST	1.07	0.932163
APEACTIV	1.07	0.933307
REDAN1	1.06	0.943743
BACPLUS	1.04	0.958768
FILLE	1.01	0.987348
Mean VIF	1.20	

Le VIF montre qu'il n'y a pas de colinéarité dans ce modèle.

```
. xtmixed STFIN2FM STINI2FM FILLE AGEPLUS PAUVR REDAN1 REDAN2 /*
> */ CONTRACTUEL BACPLUS RESTENS_EST COLLEGUES /*
> */ SALDUR DOUBLFLX COVAR ANCDIR APEACTIV PROPC_ES RURAL N5FM2 /*
> */if VALEX==0 ||NUMECOLE:, variance
```

```
Mixed-effects REML regression      Number of obs      =      1609
Group variable: NUMECOLE           Number of groups   =       124

Obs per group: min =                6
                  avg =              13.0
                  max =              15
```

```
Log restricted-likelihood = -1600.9197      Wald chi2(18)      =      951.11
                                           Prob > chi2        =      0.0000
```

STFIN2FM	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
STINI2FM	.5955589	.0215591	27.62	0.000	.5533037 .637814
FILLE	-.0542115	.0310066	-1.75	0.080	-.1149833 .0065603
AGEPLUS	.0649812	.0375508	1.73	0.084	-.0086169 .1385793
PAUVR	-.1103683	.045516	-2.42	0.015	-.199578 -.0211585
REDAN1	-.1039598	.0427542	-2.43	0.015	-.1877565 -.0201631
REDAN2	-.1839856	.0461562	-3.99	0.000	-.27445 -.0935212
CONTRACTUEL	.306615	.1049817	2.92	0.003	.1008547 .5123752
BACPLUS	-.2088994	.1836801	-1.14	0.255	-.5689059 .1511107
RESTENS_EST	.1585205	.1002443	1.58	0.114	-.0379548 .3549958

COLLEGUES	.3181199	.1631671	1.95	0.051	-.0016816	.6379215
SALDUR	-.0856988	.1109935	-0.77	0.440	-.3032422	.1318445
DOUBLFLX	-.3366454	.1179456	-2.85	0.004	-.5678144	-.1054763
COVAR	.5989896	.1758879	3.41	0.001	.2542556	.9437235
ANCDIR	.0087203	.006443	1.35	0.176	-.0039077	.0213484
APEACTIV	-.1683929	.0971359	-1.73	0.083	-.3587758	.02199
PROPC_ES	-.2012965	.2584417	-0.78	0.436	-.7078329	.3052399
RURAL	.0370054	.1127457	0.33	0.743	-.1839722	.257983
N5FM2	.2305129	.0511247	4.51	0.000	.1303104	.3307155
_cons	-.6191781	.2523098	-2.45	0.014	-1.113696	-.12466

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
NUMECOLE: Identity				
var(_cons)	.2461815	.0368854	.1835354	.3302106
var(Residual)	.3476213	.0127855	.3234441	.3736057

LR test vs. linear regression: chibar2(01) = 516.02 Prob >= chibar2 = 0.0000

Bibliographie

Akerlof, G. A., & Kranton, R. E. (2002). Identity and Schooling: Some Lessons for the Economics of Education. *Journal of Economic Literature*, 40, 1167-1201.

Amelewonou, K., Brossard, M., & Gacougnolle, L.C. (2004). *La question enseignante dans la perspective de la scolarisation primaire universelle dans les pays CEDEAO, CEMAC et PALOPs*. Mimeo. Dakar : Pôle de Dakar / UNESCO-BREDA.

American Psychological Association (1995). *Intelligence: knowns and unknowns*. Washington, DC : American Psychological Association

Angrist, J.D., & Lavy, V. (2001). Does Teacher Training Affect Pupil Learning ? Evidence from Matched Comparisons in Jerusalem Public Schools. *Journal of Labor Economics*, 19, 2, 343-369.

Banque mondiale (2007). *L'éducation au Mali. Diagnostic pour le renouvellement de la politique éducative en vue d'atteindre les objectifs du Millénaire*. Washington, DC : World Bank.

Baum, C. F. (2006). *An Introduction to Modern Econometrics Using Stata*. Lakeway Drive [Etats-Unis] : Stata Press.

Becker, G.S. (1964). *Human Capital : A Theoretical and Empirical Analysis*. New York : Columbia University Press.

Behaghel, L., Coustère, P., & Lepla, F. (1999). *Les facteurs de l'efficacité dans l'enseignement primaire : les résultats du programme PASEC sur neuf pays d'Afrique et de l'Océan Indien*. Dakar : PASEC/CONFEMEN.

Bernard, J.-M. (1999a). *Les Enseignants du primaire dans Cinq Pays du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN : le Rôle du Maître dans le Processus d'Acquisition des Élèves*. Rapport pour le groupe de travail sur la fonction enseignante. Paris : ADEA.

Bernard, J.-M. (1999b). *Les enseignants de l'enseignement primaire dans cinq pays du Programme d'Analyse des Systèmes éducatifs de la CONFEMEN : caractéristiques, conditions de travail et représentations*. Rapport pour le groupe de travail sur la fonction enseignante. Paris : ADEA.

Bernard, J.-M. (2003, mars). *Scolarisation primaire universelle et pilotage par les résultats dans le contexte africain : quels indicateurs ?* Papier présenté au colloque international « le droit à l'éducation : quelle effectivité au sud et au nord ? », Ouagadougou, Burkina Faso.

Bernard, J.-M., Nkengne Nkengne, A.P., & Robert, F. (2007). Réformes des programmes scolaires et acquisitions à l'école primaire en Afrique : Mythes et réalités. *International Review of Education*. A venir.

Bernard, J.-M., Tiyab, B., & Vianou, K. (2004). *Profils enseignants et qualité de l'éducation primaire en Afrique subsaharienne francophone: Bilan et perspectives de dix années de recherche du PASEC*. Mimeo. Dakar : PASEC/CONFEMEN.

Bernard, J.-M., Simon, O., & Vianou, K. (2005). *Le redoublement : mirage de l'école africaine ?* Dakar : PASEC/CONFEMEN.

Bourdon, J., Frölich, M.-F., & Michaelowa, K. (2006). Broadening Access to Primary Education: Contract Teacher Programs and Their Impact on Education Outcomes in Africa an Econometric Evaluation for the Republic of Niger. In L. Menkhoff (ed.), *Pro-Poor Growth: Issues, Policies, and Evidence (pp. 119-142)*, Berlin : Duncker & Humblot.

Bressoux, P. (1994). Les recherches sur les effets-école et les effets-maître. *Revue française de pédagogie*, 108, 91-137.

Bressoux (1995). Les effets du contexte scolaire sur les acquisitions des élèves : effet-école et effet-classe en lecture. *Revue française de sociologie*, XXXVI (2), 273-294.

Bressoux, P. (1996). The Effects of Teachers' Training on Pupils' Achievement: the Case of Elementary Schools in France. *School Effectiveness and School Improvement*, 7, 3, 252-279.

Bressoux, P. (2000). Modélisation et évaluation des environnements et des pratiques d'enseignement. Habilitation à diriger des recherches, Université Pierre Mendès France, Grenoble.

Bressoux, P., Kramarz, F., & Prost, C. (2006, mai). *Teachers' Training, Class Size and Students' Outcomes: Evidence from Third Grade Classes in France*. In Team for Advanced Research on Globalization, Education, and Technology (TARGET) and the National Research Data Centre Program. Paper presented at the Conference on Education, Training and the Evolving Workplace, Vancouver, Canada

Bruns, B., Mingat, A., & Rakotomalala, R. (2003). *Achieving Universal Primary Education by 2015: A Chance for Every Child*. Washington, D.C. : The World Bank.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Application and Data Analysis Models*. London : Sage.

Caner, K. (2000). *La dynamique qualitative du produit de l'éducation*. Thèse de doctorat non publiée, Université de Montpellier I, Montpellier.

Cogneau, D. (2003). *Colonisation, School and Development in Africa An empirical analysis*. Document de travail, DT/2003/01, DIAL. [en ligne].

http://www.dial.prd.fr/dial_publications/dial_publi_doctrav.htm#2003 (page consultée le 6 décembre 2006)

Cohn, E., & Geske, T.G. (1990). *The Economics of Education*. Oxford : Pergamon Press.

Coleman, J.S., & al. (1966). *Equality of Educational Opportunity*. Washington : U.S. Government Printing Office.

Cour des comptes (2003). *La gestion du système éducatif*. Rapport au président de la République. [en ligne]. <http://www.ccomptes.fr/FramePrinc/frame-rapports.htm> (page consultée le 21 septembre 2006).

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

De Ketele, J.-M. (1993). L'évaluation conjuguée en paradigmes, *Revue française de pédagogie*, 103, 59-80.

De Ketele, J.-M., & Delory, C. (2000). Quel héritage pédagogique pour le nouveau millénaire ? Forum des pédagogies. [en ligne]. <http://www.henac.be/publications/profs/heritagepedagogique.pdf> (page consultée le 30 octobre 2006).

De Ketele, J.-M., & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*. A paraître.

Demeuse, M. (2004). *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation*. Liège : les éditions de l'Université de Liège.

Duflo, E. (2006). Field Experiments in Development Economics. Prepared for the World Congress of the Econometric Society. [en ligne]. http://econ-www.mit.edu/faculty/?prof_id=eduflo&type=paper (page consultée le 9 septembre 2006).

Duflo, E., & Hanna, R. (2005). Monitoring Works: Getting Teachers to Come to Schools. *NBER working paper*, n° 11 880.

Duru-Bellat, M., Danner, M., Le Bastard-Landrier, S., & Piquée, C. (2004). Les effets de la composition scolaire et sociale du public d'élèves sur leur réussite et leurs attitudes : évaluation externe et explorations qualitatives. *Les cahiers de l'Irédu*, n° 65 [En ligne]. <http://www.u-bourgogne.fr/LABO-IREDU/cahier65.pdf> (page consultée le 8 octobre 2007)

Education pour Tous – Initiative de Mise en Oeuvre Accélérée (2004). *Accélérer les progrès vers une éducation primaire universelle de qualité*. Document cadre. [en ligne]. <http://www.fasttrackinitiative.org/education/efafti/documents/FrameworkFrench.pdf> (page consultée le 5 décembre 2006).

Farrell, J.P., & Oliveira, J. (1993). Teacher Costs and Teacher Effectiveness in Developing Countries. In J.P. Farrell & J. Oliveira (eds.), *Teachers in Developing Countries: Improving Effectiveness and Managing Costs* (pp. 7-23). Washington, DC : World Bank.

Filmer, D., & Pritchett, L. (1998). *Estimating Wealth Effects without Expenditure Data - or Tears: An Application to Educational Enrollments in States of India*. Policy Research Working Paper, no. WPS1994. Washington, DC : World Bank. [en ligne]. http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/1998/11/17/000178830_98111703524522/Rendered/PDF/multi_page.pdf (page consultée le 15 mars 2007)

Glewwe, P. (2002). Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes. *Journal of Economic Literature*, 40, 2, 436-482.

Glewwe, P., Ilias, N., & Kremer, M. (2003). *Teacher Incentives*. Cambridge : National Bureau of Economic Research, (Working paper, 9671) [en ligne] <http://www.nber.org/papers/w9671> (page consultée le 10 octobre 2007)

Glewwe, P., & Kremer, M. (2006). Schools, Teachers, and Education Outcomes in Developing Countries. In E.A. Hanushek & F. Welch (eds.), *Handbook on the Economics of Education*. Amsterdam : Elsevier.

Goldstein, H. (1995). *Multilevel Statistical Models*. London : Arnold.

- Greene, W. H. (2005). *Econométrie*. 5^{ème} édition. Paris : Pearson Education France.
- Gundlach, E., Woessmann, L., & Gmelin, J. (2001). The Decline of Schooling Productivity in OECD Countries. *Economic Journal*, 111, 471, 135-47.
- Hanushek, E. A. (1971). Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data. *American Economic Review*, 61, 2, 280-288.
- Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *The Journal of Human Resources*, 14, 3, 351-388.
- Hanushek, E. A. (1997). Assessing the effect of school resources on student performance: An Update. *Educational Evaluation and Policy Analysis*, 19, 2, 141-164.
- Hanushek, E. A. (2003). The Failure of Input-based Schooling Policies. *The Economic Journal*, 113, 485, F64-F98.
- Hanushek, E. A., & Kimko, D. (2000). Schooling, Labor Force Quality, and the Growth of Nations. *American Economic Review*, 90, 5, 1184-1208.
- Hanushek, E. A., & Woessman, L. (2007). *Education Quality and Economic Growth*. Washington, DC : World Bank.
- Hox, J. (2002). *Multilevel Analysis: Techniques and Applications*. Mahwa [Etats-Unis] : Lawrence Erlbaum Associates.
- Hoxby C. M. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of Economics*, 115, 4, 1239-1285.
- Huber P.J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1. (pp. 221–233). Berkeley : University of California Press

Jacob, B.A., & Lefgren, L. (2004). The Impact of Teacher Training on Student Achievement: Quasi-Experimental Evidence from School Reform Efforts in Chicago. *Journal of Human Resources*, 39, 1, 50-79.

Jarousse, J.P., & Mingat, A. (1993). *L'école primaire en Afrique : analyse pédagogique et économique*. Paris : L'Harmattan.

Kish, L. (1965). *Survey Sampling*. New York : John Wiley & Sons.

Konstantopoulos, S. (2007). *How Long Do Teacher Effects Persist ?* Discussion Paper n°2893, IZA. [en ligne]. <http://ftp.iza.org/dp2893.pdf> (page consultée le 12 septembre 2007)

Kreft, I.G.G., & Deleuw, J. (1998). *Introducing Multilevel Modelling*. London : Sage.

Kremer, M. (1995). Research on Schooling: What We Know and What We Don't: A Comment on Hanushek. *World Bank Research Observer*, 10, 2, 247-54.

Krueger, A.B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics*, 114, 2, 497-532.

Krueger, A. B. (2003). Economic Considerations and Class Size. *Economic Journal*, 113, 485, 34-63.

Lambert, S. (2004). *Teacher's Pay and Conditions: an Assessment of Recent Trends in Africa*. Mimeo. Paris : Paris-Jourdan : LEA-INRA

Lancaster, K.J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74, 132-157.

Lassibille, G., & Tan, J. P. (2003). Student learning in public and private primary schools in Madagascar. *Economic Development and Cultural Change*, 51, 3, 699-717.

Laveault, D., & Grégoire, J. (1997). *Introduction aux théories des tests en sciences humaines*. Bruxelles : De Boeck Université.

Leamer, E.E. (1983). Let's take The Con out of Econometrics. *The American Economic Review*, 73, 1, 31-43.

Leclercq, F. (2005). *The Relationship between Educational Expenditures and Outcomes*. Working Paper, DIAL, DT/2005/05. [en ligne] http://www.dial.prd.fr/dial_publications/PDF/Doc_travail/2005-05.pdf (page consultée le 3 novembre 2007).

Leibenstein, H. (1966). Allocative efficiency vs. "X-Efficiency". *The American Economic Review*, 56, 3, 392-415.

Lemelin, C. (1998). *L'économiste et l'éducation*. Sainte-Foy : Presses de l'Université du Québec.

Levin, H.M. (1997). Raising School Productivity : an X-Efficiency Approach. *Economics of Education Review*, 16, 3, 303-311.

Long, J.S and Ervin L.H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician*, 54, 217-224.

Lucas, R.E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22, 3-42.

Maurin, E. (2007). *La nouvelle question scolaire : Les bénéfices de la démocratisation*. Paris : Seuil

McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., & Hamilton, L.S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica : Rand Corporation.

McKinnon, J. G., & White, H. (1985). Some Heteroscedasticity Consistent Covariance Matrix Estimators With Improved Finite Sample Properties. *Journal of Econometrics*, 29, 305-325.

McMeekin, R.W. (2003). *Incentives to Improve Education. A New Perspective*. Cheltenham [Royaume-Uni] : Edward Elgar Publishing.

Michaelowa, K. (2001a). Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations. *World Development*, 29, 10, 1699-1716.

Michaelowa, K. (2001b). Scolarisation et acquis des élèves : les indicateurs de résultats dans l'analyse des politiques de l'enseignement en Afrique francophone. *Politiques d'éducation et de formation : Analyses et comparaisons internationales*, 1, 3, 77-94.

Michaelowa, K. (2002). *Teacher Job Satisfaction, Student Achievement, and the Cost of Primary Education in Francophone Sub-Saharan Africa*. Discussion Paper n°188, Hamburg : Hamburg Institute of International Economics.

Michaelowa, K., & Wechtler, A. (2006). *The cost-effectiveness of inputs in primary education: Insights from the literature and recent student surveys for sub-Saharan Africa*. Working document. Paris : ADEA.

Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66, 4, 281-302

Mingat, A. (1984). Les acquisitions scolaires au CP : l'origine des différences ? *Revue française de pédagogie*, 69, 49-62.

Mingat, A. (1987). Sur la dynamique des acquisitions à l'école élémentaire. *Revue française de pédagogie*, 79, 5- 14.

Mingat, A. (1991). Expliquer la variété des acquisitions au cours préparatoire : les rôles de l'enfant, la famille et l'école. *Revue française de pédagogie*, 95, 47-63.

Mingat, A. (2001). *Teachers Salary Issues in African Countries*. Mimeo. Equipe Analyse et Politiques pour le Développement Humain, Banque Mondiale, Région Afrique. Washington, D.C. : World Bank

Mingat, A. (2004). *La rémunération/le statut des enseignants dans la perspective de l'atteinte des objectifs du millénaire dans les pays d'Afrique subsaharienne francophone en 2015*. Mimeo. Equipe Analyse et Politiques pour le Développement Humain, Banque Mondiale, Région Afrique. Washington, D.C : World Bank

Mingat, A., & Suchaut, B. (2000). *Les systèmes éducatifs africains : une analyse économique comparative*. Bruxelles : De Boeck Université.

Moulton, B.R. (1986). Random Group Effects and the Precision of Regression Estimates. *Journal of Econometrics*, 32, 3, 385-397.

Murname, R.J., Willett, J.B., Somers, M.A., & Uribe, C. (2005). Learning Why More Learning Takes Place in Some Classrooms Than Others. *German Economic Review*, 6, 3, 309-330.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis*, 26, 3, 237-257.

Paxson, C., & Schady, N. (2007). Cognitive Development among Young Children in Ecuador. The Roles of Wealth, Health, and Parenting. *Journal of Human Resources*, 42, 1, 49-84

PASEC (2000). *Guide pour l'évaluation des facteurs de performance à l'école primaire : manuel pratique d'évaluation*. Dakar : CONFEMEN

PASEC (2003). *Le Programme de formation initiale des maîtres et la double vacation en Guinée*. Dakar : CONFEMEN.

PASEC (2004a). *Les enseignants contractuels et la qualité de l'enseignement de base I au Niger : Quel bilan ?* Dakar : CONFEMEN.

PASEC (2004b). *Impact du statut enseignant sur les acquisitions dans le premier cycle l'enseignement fondamental public au Mali.* Dakar : CONFEMEN.

PASEC (2006a). *La qualité de l'éducation en Mauritanie. Quelles ressources pour quels résultats ?* Dakar : CONFEMEN.

PASEC (2006b). *La formation des enseignants contractuels : étude thématique – Guinée.* Dakar : CONFEMEN.

Pôle de Dakar (2005). *Education pour Tous en Afrique : repères pour l'action.* Dakar : UNESCO-BREDA.

Pritchett, L. (2001). Where has All the Education Gone ? *World Bank Economic Review*, 15, 3, 367–391.

Pritchett, L., & Filmer, D. (1999). What Education Production Functions Really Show: A Positive Theory of Education Expenditures. *Economics of Education Review*, 18, 2, 223-39.

Psacharopoulos, G. (1994). Returns to Investment in Education: a Global Update. *World Development*, 22, 9, 1325-1343.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teacher, Schools and Academic Achievement. *Econometrica*, 73, 2, 417-445.

Schultz, T. W. (1961). Investment in Human Capital. *The American Economic Review*, 51, 1, 1-17

Simon, H.A. (1991). Organizations and Markets. *Journal of Economic Perspectives*, 5, 2, 25-44.

Tan, J.-P., Lane, J., & Coustere, P. (1997). Putting Inputs to Work in Elementary Schools: What Can Be Done in the Philippines ? *Economic Development and Cultural Change*, 45, 4, 857-879.

Tiebout, C.M. (1956). A pure theory of public expenditure. *Journal of Political Economy*, 64, 416-424.

Todd, P. E., & Wolpin K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *Economic Journal*, 113, 485, 3-33.

UNESCO-BREDA (2007). *Rapport Dakar + 7. EPT : L'urgence de politiques sectorielles intégrées*. Dakar : UNESCO-BREDA.

UNESCO Institute for Statistics (2006). *Teachers and Educational Quality : Monitoring Global Needs for 2015*. Montreal : UNESCO Institute for Statistics.

White, H. (1980). Nonlinear Regression on Cross-Section Data. *Econometrica, Econometric Society*, 48, 3, 721-46

Wooldridge, J.M. (2006). *Introductory Econometrics. A Modern Approach*. Third Edition. Crawfordsville [Etats-Unis] : Thomson South-Western.

Table des matières

Introduction générale	9
Chapitre 1 : L'analyse des déterminants des acquisitions scolaires : des débats théoriques aux analyses empiriques	15
1.1 La quête désespérée de la fonction de production éducative	17
1.1.1 Le concept de fonction de production dans la théorie microéconomique	17
1.1.2 Spécificité et mesure du produit de l'éducation	18
1.1.3 La remise en cause de la fonction de production éducative	24
1.1.4 Débats théoriques et enjeux empiriques	31
1.2 La modélisation statistique des acquis scolaires : débats méthodologiques	33
1.2.1 Formalisation du processus d'acquisitions scolaires	33
1.2.2 Les problèmes d'estimation des modèles explicatifs des acquis scolaires : quid des modèles à valeur ajoutée ?	35
1.2.3 La sensibilité des estimations des modèles de gains annuels : le cas des données du PASEC	38
1.2.4 L'estimation alternative des modèles à valeur ajoutée : progrès ou régression ? ..	48
1.2.5 L'utilisation des modèles hiérarchiques ou multiniveaux	52
1.2.6 L'estimation des modèles explicatifs des acquis scolaires : un monde de biais	57
1.2.7 Quelques perspectives pour l'analyse des déterminants des acquisitions scolaires	58
Chapitre 2 : La problématique enseignante à l'école primaire en Afrique francophone : entre enjeux de scolarisation et qualité de l'éducation	60
2.1 Extension de la scolarisation et évolution du corps enseignant	62
2.1.1 Les attentes en matière de scolarisation	62
2.1.2 Les implications de l'accroissement des effectifs d'élèves sur le recrutement d'enseignants	63
2.1.3 La composition du corps enseignant : tendance actuelle et évolution	66
2.1.4 La rémunération des enseignants	69
2.1.5 L'impact sur la scolarisation du recrutement d'enseignants non fonctionnaires	74
2.1.6 Au-delà de l'extension de la scolarisation, la question de la qualité de l'éducation ...	75
2.2 Caractéristiques des enseignants et acquisitions scolaires : des convergences dans les résultats des recherches	76
2.2.1 Caractéristiques des enseignants et acquisitions scolaires dans les pays développés .	77
2.2.2 Caractéristiques des enseignants et acquisitions scolaires dans les pays pauvres...	84
2.3 Les caractéristiques des enseignants au cœur des enjeux de la scolarisation primaire universelle	95

Chapitre 3 : statut des enseignants et acquisitions à l'école primaire au Mali : faux problème et vraies questions97

3.1 Protocole d'enquête et données disponibles.....	98
3.1.1 Le protocole d'enquête	99
3.1.2 Les données	102
3.2 Statut et acquisitions scolaires : quelles évidences ?.....	108
3.2.1 Les résultats en 2 ^{ème} année : Les enseignants contractuels à leur avantage	110
3.2.2 Les résultats en 5 ^{ème} année : un écart moins net en faveur des enseignants contractuels.....	116
3.2.3 Les enseignants contractuels et la qualité des apprentissages	120
3.3 Où se situent les marges de manœuvre pour l'amélioration de la qualité des apprentissages dans l'enseignement fondamental malien ?	121
3.4 Statut des enseignants et qualité des apprentissages : une question essentiellement politique.....	125

Chapitre 4 : l'amélioration de la qualité des apprentissages à l'école primaire en Afrique : les enjeux de la gestion des systèmes éducatifs..... 127

4.1 Le poids des disparités entre écoles et entre classes dans le processus d'apprentissage	129
4.1.1 La variabilité des résultats imputables au niveau classe ou école	129
4.1.2 Les disparités entre classes : « Effet-maître », « effet-classe » ou ?	134
4.1.3 Les questions soulevées par l'effet-classe	144
4.2 De l'effet-classe à l'effet-école	145
4.3 Disparités entre écoles et gestion de la qualité des apprentissages	151
4.3.1 Les sources potentielles des disparités entre classes ou écoles	151
4.3.2 Des indicateurs pour la gestion locale de la qualité des apprentissages.....	154
4.4 La gestion locale comme moyen de réduire les disparités entre écoles	162

Conclusion générale 164

Liste des annexes

ANNEXE 1.2.1 : Méthodes et instruments du Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC)	172
ANNEXE 1.2.2 : Estimation des MCO avec données hiérarchiques	177
ANNEXE 1.2.3 : Comparaison des modèles explicatifs des scores de 5 ^{ème} année en français et mathématiques selon la prise en compte du score initial	187
ANNEXE 1.2.4 : Comparaison des scores moyens de début et de fin d'année selon certaines caractéristiques de l'élève et du contexte	200
ANNEXE 1.2.5 : Création de la variable basée sur l'évolution du classement de l'élève aux tests standardisés sous STATA	205
ANNEXE 1.2.6 : Endogénéité du score de début d'année dans le modèle à valeur ajoutée	207
ANNEXE 1.2.7 : Comparaison des modèles MCO avec les modèles hiérarchiques	208
ANNEXE 1.2.8 : Les modèles hiérarchiques avec et sans score initial	213
ANNEXE 2.1.1 : Principaux paramètres du cadre indicatif pour l'initiative de mise en œuvre accélérée de l'Education pour tous	214
ANNEXE 2.1.2 : Proportion d'élèves scolarisés dans le privé à l'école primaire dans 12 pays d'Afrique francophone	215
ANNEXE 2.2.1 : Les résultats des variables maître dans les modèles avec et sans score initial dans l'étude de Michaelowa et Wechtler (2006)	216
ANNEXE 3.1.1 : Echantillon théorique de l'évaluation PASEC Mali	217
ANNEXE 3.1.2 : Traitement des valeurs extrêmes	219
ANNEXE 3.1.3 : Comparaison des statistiques descriptives selon le statut de l'enseignant	223
ANNEXE 3.1.4 : Construction des variables pour mesurer le niveau de vie de la famille de l'élève	230
ANNEXE 3.2.1 : Modèles explicatifs des scores de 2 ^{ème} année	231
ANNEXE 3.2.2 : Modèles explicatifs des scores de 5 ^{ème} année	237
ANNEXE 3.2.3 : Tentative d'instrumentation de la variable contractuelle dans les modèles explicatifs des scores de 2 ^{ème} et 5 ^{ème} années	241
ANNEXE 3.3.1 : Les variances inter-élèves et inter-classes dans les modèles multiniveau	249

ANNEXE 4.1.1 : Mesure de la variance au niveau classe et école	251
ANNEXE 4.1.2 : La mesure de l'effet classe	256
ANNEXE 4.1.3 : Test de la relation entre l'effet-classe et le nombre d'élèves par classe retenus dans les enquêtes	261
ANNEXE 4.1.4 : La mise en évidence d'un effet école : Le cas du Mali	263

Liste des tableaux

Tableau 1.1 :	Fiabilité des tests de mathématiques de 5 ^{ème} année du PASEC	22
Tableau 1.2.1 :	Estimation de modèles avec et sans score initial en 5 ^{ème} année d'école primaire dans trois pays du PASEC	41
Tableau 1.2.2 :	Différences entre les coefficients des deux modèles (valeurs absolues de MSI-MVA).....	42
Tableau 1.2.3 :	Tableau récapitulatif de la variabilité des coefficients	44
Tableau 1.2.4 :	Comparaisons des scores moyens selon l'origine sociale des élèves à Madagascar (Cf. annexe 1.2.4).....	45
Tableau 1.2.5 :	Comparaisons des scores moyens selon l'origine sociale des élèves au Tchad (Cf. annexe 1.2.4)	45
Tableau 1.2.6 :	Comparaisons des scores moyens selon la formation académique de l'enseignant à Madagascar (Cf. annexe 1.2.4)	46
Tableau 1.2.7 :	Comparaisons des scores moyens selon l'organisation de la classe en double flux à Madagascar (Cf. annexe 1.2.4)	47
Tableau 1.2.8 :	Caractéristiques des variables de rang des élèves selon les pays	49
Tableau 1.2.9 :	Comparaison des résultats des modèles à valeur ajoutée et des modèles de rang	51
Tableau 1.2.10 :	Coefficient de corrélation intra-classe dans les données PASEC	53
Tableau 1.2.11 :	Comparaison des MCO avec le modèle multiniveau en 5 ^{ème} année d'école primaire dans trois pays du PASEC	56
Tableau 2.1.1 :	Accroissement du nombre d'élèves entre 1987 et 2001 et entre 2001 et 2015	62
Tableau 2.1.2 :	Evolution passée et future du nombre d'enseignants du secteur public pour atteindre la scolarisation primaire universelle	65
Tableau 2.1.3 :	Distribution des enseignants du primaire selon le statut	67
Tableau 2.1.4 :	Recrutement du personnel fonctionnaire et contractuel de 1998 à 2002 au Mali	68
Tableau 2.1.5 :	Taux de dépendance en 2015 selon trois hypothèses pour le salaire des enseignants recrutés pour assurer l'achèvement universel du primaire	72

Tableau 2.1.6 : Niveau de rémunération des enseignants du primaire selon le statut.....	73
Tableau 2.1.7 : Estimation du gain de scolarisation imputable au recrutement d'enseignants non-fonctionnaires.....	74
Tableau 2.2.1 : Pouvoir explicatif du modèle de Tan et al. (1997) selon les différentes spécifications.....	85
Tableau 2.2.2 : Résultats relatifs aux variables maître des modèles de Michaelowa et Wechtler (2006).....	88
Tableau 2.2.3 : Les enseignants de l'échantillon PASEC Guinée.....	91
Tableau 2.2.4 : Répartition selon le statut des enseignants des échantillons PASEC Mali et Niger.....	92
Tableau 2.2.5 : Efficacité pédagogique des enseignants non-fonctionnaires en comparaison aux fonctionnaires au Mali et au Niger.....	92
Tableau 3.1.1 : Caractéristiques des élèves, des enseignants et des écoles de l'échantillon de 2 ^{ème} année.....	103
Tableau 3.1.2 : Caractéristiques des élèves, des enseignants et des écoles de l'échantillon de 5 ^{ème} année.....	107
Tableau 3.2.1 : Modèles explicatifs des scores de fin de 2 ^{ème} année au Mali.....	111
Tableau 3.2.2 : Modèles explicatifs des scores de fin de 5 ^{ème} année au Mali.....	117
Tableau 3.3.1 : Les variances inter-classes et inter-élèves.....	122
Tableau 3.3.2 : Effets fixes « classe » en 2 ^{ème} et 5 ^{ème} années au Mali.....	124
Tableau 4.1.1 : Part de variance du score de mathématiques de 5 ^{ème} année du PASEC expliquée par les variables muettes « classe ».....	130
Tableau 4.1.2 : Part de variance du score de mathématiques de 6 ^{ème} année du SACMEQ expliquée par les variables muettes « classes » et « écoles ».....	131
Tableau 4.1.3 : Part de variance du score de mathématiques de 4 ^{ème} année du LLECE 1997 expliquée par les variables muettes « écoles ».....	132
Tableau 4.1.4 : Part de variance du score de mathématiques de 4 ^{ème} année de TIMSS 2003 expliquée par les variables muettes « classes » et « écoles ».....	133
Tableau 4.1.5 : Part de variance expliquée par les variables muettes « classes » dans les modèles à valeur ajoutée sur les données du PASEC.....	136
Tableau 4.1.6 : Comparaison des scores moyens minimaux et maximaux de français aux tests PASEC de 5 ^{ème} année de l'école primaire.....	137

Tableau 4.1.7 : Part de variance expliquée par l'ensemble des variables explicatives à l'exception du score initial dans les modèles à valeur ajoutée sur les données du PASEC	140
Tableau 4.1.8 : Les résultats des études sur les effets-classe aux Etats-Unis et en France ..	143
Tableau 4.2.1 : Incidence des classes sous performantes sur la qualité des apprentissages au sein des systèmes éducatifs de quatre pays francophones	146
Tableau 4.2.2 : La relation entre le score moyen de 5 ^{ème} année et les acquis en 2 ^{ème} année	149
Tableau 4.3.1 : Les taux de connaissance de base en 5 ^{ème} année de l'enseignement primaire dans cinq pays d'Afrique francophone	157

Liste des illustrations

Graphique 1.1 :	Score moyen ajusté de français et mathématiques et coût unitaire dans chaque classe de 5 ^{ème} année de l'enseignement fondamental de l'échantillon PASEC en Mauritanie.....	28
Graphique 2.1.1 :	Evolution du salaire moyen des enseignants de l'école primaire par région et sous-région du monde (PIB/tête) (pays ayant un PIB par tête inférieur à 2000 \$US en 1993).....	70
Graphique 3.3.1 :	Relation entre score moyen initial et score moyen final au niveau des classes en 5 ^{ème} année de l'école fondamentale au Mali	123
Graphique 4.1.1 :	Comparaison des parts de variance expliquée par les variables caractérisant les enseignants et les variables muettes « classe » en 2 ^{ème} année.....	138
Graphique 4.1.2 :	Comparaison des parts de variance expliquée par les variables caractérisant les enseignants et les variables muettes « classe » en 5 ^{ème} année.....	139
Schéma 4.1 :	Les interactions entre Offre scolaire, demande d'éducation et qualité de l'enseignement	152

Développement des sigles et acronymes utilisés

BRAC	Bangladesh Rural Advancement Committee
BREDA	Bureau régional de l'Unesco pour l'éducation en Afrique
CAP	Cellule d'Animation Pédagogique
CEDEAO	Communauté économique des États de l'Afrique de l'Ouest
CEMAC	Communauté Economique et Monétaire de l'Afrique Centrale
CONFEMEN	Conférence des Ministres de l'Éducation des pays ayant le français en partage
DEPP	Direction de l'Évaluation, de la Prospective et de la Performance
ENI	Ecole Normale d'Instituteurs
FIME	Formation Initiale des Maîtres de l'Elémentaire
FIMG	Formation Initiale des Maîtres de Guinée
IEA	Association Internationale pour l'évaluation du rendement scolaire
ISO	International Organization for Standardization
LLECE	Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación
MCG	Moindres Carrés Généralisés
MCO	Moindres Carrés Ordinaires
MLA	Monitoring Learning Achievement
MLN	Modèle multiniveau
MSI	Modèle sans Score Initial
MVA	Modèle à Valeur Ajoutée
OCDE	Organisation de Coopération et de Développement Économiques
PASEC	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN
PIB	Produit Intérieur Brut

PIRLS	Programme international de recherche en lecture scolaire
RESEN	Rapports d'État des Systèmes Éducatifs Nationaux
SACMEQ	Southern Africa Consortium for Monitoring Educational Quality
SAT	Scholastic Aptitude Test
STAR	Tennessee Student/Teacher Achievement Ratio experiment
TAAS	Texas Assessment of Academic Skills
TCB	Taux de Connaissance de Base
TIMSS	Trends in International Mathematics and Science Study
UNESCO	Organisation des Nations unies pour l'Éducation, la Science et la Culture

**The Education Production Function reviewed
in the framework of Education For All in Sub-Saharan Africa:
From theoretical and methodological bounds to the contribution to educational policies**

Abstract

The Coleman report (1966) was the first of numerous researches exploring quality of learning determinants analysis. With the Education Production Function (EPF), economists have greatly contributed to the extension of this research domain. Actually, empirical works are leading the field with important methodological debates which have fuelled controversies (Hanushek, 2003; Krueger, 2003). Thus, methodological issues have particularly rallied researchers' interests on empirical estimations, while academic researches on the renewal of the theoretical framework have remained scarce despite obvious bounds. Without contesting the validity of miscellaneous estimation problems and the necessity of other kind of researches as experimental approach, it seems that their consequences are not insurmountable if using a cautious interpretation of the EPF results.

The challenges of Education for All (EFA) in Africa and particularly primary universal education stress out the teacher issue. The need of a massive recruitment of teachers till 2015 raises tricky issues as financial sustainability, linked to socio-political problems, but also the profile of these new teachers. On this last point, a selected researches survey shows a relative convergence of results. According to the analyses, teachers' characteristics, such as training, length of service, status, etc. have limited effect on pupils' learning achievement. Teachers' status is a controversial subject in francophone Africa where recruitment of contractual teachers has increased during the last years. The results of empirical studies, as PASEC studies in primary education, do not confirm their responsibility in the weakening education quality. At the opposite, our analysis on Mali shows, in this particular case, a positive effect. Nevertheless, according to this analysis, status and other teachers' characteristics do not play a key role in quality of learning whereas disparities between classes and schools explain a large part of the differences noted in learning achievements between pupils. The literature supposes that it is due to the teacher effect (Bressoux, 2000). This assumption raises certain reserves. Analyses conducted in 11 African countries demonstrate clear higher values attributed to this effect than in developed countries. Measurement issues suggest cautiousness in the interpretation of the results. However, they show that disparities between classes and schools are at the heart of the issue credibly related to the quality of learning in numerous African countries. It underlines the crucial importance to manage the quality of learning with a specific attention paid to the disparities between schools in current management. Relevant indicators can be mobilized quite easily but their use for a result-based management still remains a prerequisite.

Keywords: Economics of Education, Allocative Efficiency, Cost–Benefit Analysis, Sub-Saharan Africa, Millennium Development Goals

**La fonction de production éducative revisitée
dans le cadre de l'Education Pour Tous en Afrique subsaharienne :
Des limites théoriques et méthodologiques aux apports à la politique éducative**

Résumé

Le rapport Coleman (1966) est le premier d'une longue liste de recherches sur l'analyse des déterminants de la qualité des apprentissages. Avec la fonction de production éducative (FPE), les économistes ont largement contribué au développement de ce domaine de recherche. A vrai dire, ce sont surtout les travaux empiriques qui ont la primeur avec de nombreux débats méthodologiques qui ont alimenté certaines controverses (Hanushek, 2003 ; Krueger, 2003). Ainsi, ce sont surtout les questions méthodologiques sur les estimations empiriques qui ont mobilisé les chercheurs, tandis que les travaux sur le renouvellement du cadre théorique sont restés rares malgré des limites manifestes. Sans contester la validité d'un certain nombre de problèmes d'estimation et de la nécessité de compléter les travaux avec l'approche quasi expérimentale, il apparaît que leurs conséquences ne sont pas insurmontables dans le cadre d'une interprétation prudente des résultats de la FPE.

Les enjeux de l'éducation pour tous (EPT) en Afrique et particulièrement de la scolarisation primaire universelle mettent l'accent sur la question enseignante. La nécessité d'un recrutement massif d'enseignants d'ici 2015 soulève des questions épineuses sur le plan de la soutenabilité financière, avec les enjeux sociopolitiques qui y sont liés, mais aussi sur le profil de ces nouveaux enseignants. De ce point de vue, la mobilisation d'un certain nombre de travaux montre une relative convergence des résultats. Il ressort une influence modeste des caractéristiques de l'enseignant comme sa formation, son ancienneté ou encore son statut. Ce dernier fait l'objet de débats houleux en Afrique francophone où le recrutement d'enseignants contractuels s'est intensifié au cours des dernières années. Les résultats des travaux empiriques, notamment ceux menés par le PASEC à l'école primaire, ne confirment pas les accusations de dégradation de la qualité de l'éducation dont font l'objet ces nouveaux enseignants. L'analyse réalisée sur le Mali montre même que dans ce cas particulier l'impact est positif. Cependant, il ressort également que le statut ainsi que les autres caractéristiques des enseignants, sans être négligeables, ne sont pas pour autant des déterminants majeurs de la qualité des apprentissages. En revanche, les disparités entre classes expliquent une part importante des différences d'acquisitions scolaires entre les élèves. La littérature sur la question attribue cela à un effet-maître (Bressoux, 2000) qui soulève cependant des réserves. Les analyses menées sur 11 pays africains montrent des valeurs nettement plus élevées de cet effet que dans les pays développés. Les problèmes de mesure suggèrent la prudence dans l'interprétation des résultats. Ils permettent cependant de montrer que les disparités entre classes sont au cœur de la problématique de la qualité des apprentissages et ceci dans un grand nombre de pays africains. Cela souligne l'importance que les systèmes éducatifs doivent accorder à la question de la gestion de la qualité des apprentissages en attachant une attention particulière sur les disparités entre établissements dans la gestion courante. Des indicateurs utiles peuvent être assez facilement mobilisés. Il reste à ce qu'ils soient utilisés dans le cadre d'une gestion qui accordera une priorité aux résultats.

Mots clés : Economie de l'éducation, Efficacité de la gestion, Analyse coût-efficace, Afrique subsaharienne, Objectifs du millénaire.

JEL Classification System: D61, H75, I21