



HAL
open science

Qualité de l'éducation, taille des classes et mixité sociale : Un réexamen à partir des méthodes à variables instrumentales et semi-paramétriques sur données multiniveaux - Cas du Maroc -

Mohammed Bijou

► **To cite this version:**

Mohammed Bijou. Qualité de l'éducation, taille des classes et mixité sociale : Un réexamen à partir des méthodes à variables instrumentales et semi-paramétriques sur données multiniveaux - Cas du Maroc -. Sociologie. Université de Toulon, 2021. Français. NNT : 2021TOUL2004 . tel-04531250

HAL Id: tel-04531250

<https://theses.hal.science/tel-04531250v1>

Submitted on 3 Apr 2024

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.

Qualité de l'éducation, taille des classes et mixité sociale

Un réexamen à partir des méthodes à variables instrumentales et semi-paramétriques sur données multinationales

- Cas du Maroc -

Thèse de Doctorat en Sciences Economiques

Mohammed BIJOU

Directeur de Thèse : **Valérie BERENGER**
Co-Directeur de Thèse : **Saïd HANCHANE**

Jury :

Eric VERDIER <i>Président de jury</i>	Directeur de Recherche au CNRS Laboratoire d'Economie et de Sociologie du Travail-LEST, Aix-Marseille Université et CNRS
Valérie BERENGER <i>Directeur de thèse</i>	Professeur, Laboratoire d'Economie Appliquée au Développement-LEAD, Université du Sud Toulon-Var
Saïd HANCHANE <i>Co-directeur de thèse</i>	Professeur, Faculté de Gouvernance Science Economiques & Sociales -FGSES, Université Mohamed VI Polytechnique-UM6P
Najat EL MEKKAOU-DE FREITAS <i>Rapporteur</i>	Professeur, Université Paris Dauphine-PSL
Jean-François GIRET <i>Rapporteur</i>	Professeur et Directeur de l'institut de recherche en éducation-IREDU, Université de Bourgogne

L'université de Toulon n'entend donner aucune approbation ou improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Cette thèse de doctorat n'a pu devenir réalité sans le soutien et la forte implication de nombreuses personnes qui m'ont épaulé tout au long de ces années de recherche. A ce titre, je souhaite aujourd'hui leur adresser mes plus vifs remerciements.

En premier lieu, je souhaite exprimer ma plus profonde reconnaissance et toute ma sympathie à Madame le Professeur Valérie Berenger qui a accepté de diriger cette thèse. Son encadrement est avant tout un honneur. Ses remarques et orientations tant sur le plan analytique que méthodologique ont été indispensables et déterminantes pour améliorer la qualité du travail accompli.

Ma sincère gratitude va également à mon codirecteur de thèse, Monsieur le Professeur Said Hanchane. Son enthousiasme scientifique a suscité mon intérêt pour la voie doctorale après le diplôme de Master de recherche en micro-économétrie de l'éducation. Son soutien scientifique, méthodologique et parfois psychologique m'ont été très précieux.

Je suis très sensible à l'honneur que m'ont fait Professeurs EL Mekkaoui–De Freitas Najat et Jean-Francois Giret d'avoir accepté d'évaluer mon travail en tant que rapporteurs. Je voudrais également exprimer ma reconnaissance au Professeur Eric Verdier pour sa présence en tant que président du jury.

Que soient aussi chaleureusement remercié tout le personnel de l'école doctorale, les membres du LÉAD, mes amis et collègues de l'Université de Toulon. Mes remerciements vont également à Benaabdelaali Wail pour m'avoir accompagné et encouragé pendant toute la durée de ce travail.

Je voudrais également remercier toutes les personnes attentives à l'aboutissement de cette thèse. Ils sont très nombreux pour que je puisse les nommer mais qu'ils soient tous assurés de ma sincère gratitude.

Ces remerciements seraient incomplets sans mentionner mes collègues de tous les jours de la faculté de gouvernance, d'économie et sociale de rabat en particulier et de l'université UM6P en général, avec qui j'ai partagé plus de cinq années de ma vie.

Et évidemment...mes chers parents, ma femme, mes sœurs, ma famille, vous qui m'avez accordé votre affectueuse bienveillance, précieux soutien et à qui j'ai dû manquer, cette thèse est un peu la vôtre.

A ma famille

Table des matières

Introduction générale	7
Chapitre 1 : Les déterminants du rendement scolaire chez les élèves de l'enseignement primaire et collégial au Maroc : une nouvelle approche économétrique des modèles multiniveaux	19
Introduction.....	20
1. Littérature empirique sur les déterminants des performances scolaires	22
1.1 Caractéristiques intrinsèques de l'élève et acquis scolaires	22
1.2 Origine sociale, environnement familial et acquis scolaires de l'élève	24
1.3 L'effet établissement et acquis scolaires	26
1.4 Effets de pairs.....	28
1.5 Les travaux empiriques analysant la performance scolaire des élèves marocains	28
2. Description du dispositif des deux programmes internationales TIMSS-PIRLS 2011 : Maroc et analyse en benchmark.....	29
2.1 Présentation de l'enquête	30
2.1.1. Echantillonnage	30
2.1.2. Profil des élèves	31
2.1.3. Profil des parents d'élèves	32
2.1.4. Profil des enseignants	34
2.1.5. Caractéristiques des établissements scolaires	37
2.2. Performance scolaire des élèves marocains : Une analyse descriptive en benchmarking	39
2.2.1 Performances et classement des élèves marocains.....	40
2.3. Le niveau des acquis selon le niveau de compétence.....	44
2.3.1. Niveau des acquis des élèves marocains selon leurs caractéristiques individuelles, familiales et scolaires	45
2.3.1.1. Niveau des acquis scolaires selon les caractéristiques individuelles et les antécédents scolaires	45
2.3.1.2. Niveau des acquis selon les variables du contexte familial	47
2.3.1.3. Niveau des acquis selon les caractéristiques des enseignants	49
2.3.2. Niveau des acquis selon l'environnement de l'établissement.....	52
3. Données exploitées et méthodologie d'estimation	55
3.1. Description des variables utilisées et traitement des données	55
3.2. Evaluation des acquis scolaires par le recours aux modèles multiniveaux	57
3.2.1. Justification du modèle multiniveau.....	57
3.2.2. Présentation du modèle multiniveau.....	59
4. Les résultats de la modélisation	65
4.1. Le rôle des caractéristiques individuelles et familiales de l'élève dans l'explication des acquis scolaires	65
4.2. Les effets de l'établissement scolaire sur les acquis des élèves	72
Conclusion	74
Chapitre 2 : Taille des classes et acquis scolaires des élèves marocains, une nouvelle approche semi paramétrique des modèles multiniveaux	77
Introduction.....	78
1. Revue de la littérature	80
1.1. Etudes économétriques.....	80
1.2. Etudes expérimentales.....	81
1.3. Méthode des variables instrumentales.....	85
2. Données utilisées et faits stylisés	92
3. Statistiques descriptives sur la taille des classes, les caractéristiques des élèves de la 4e année primaire et 2e année collégiale et des inégalités des acquis scolaires	93
3.1. Disparités des performances scolaires en lecture des élèves de la 4e année du primaire selon la taille des classes.....	95

3.2.	Disparités des performances scolaires en mathématiques et en sciences des élèves de la 2 ^e année collégiale selon la taille des classes	99
4.	Méthodologie d'estimation	105
4.1.	Estimation marginale dans les modèles non-paramétriques sous une spécification hiérarchique.....	109
4.1.1.	Un modèle local linéaire sous une spécification hiérarchique.....	109
4.1.2.	Un modèle semi-paramétrique : Estimation du modèle à effets aléatoires	114
4.1.3.	Estimation de l'erreur quadratique moyenne avec bootstrap	115
4.1.4.	Sélection de bande passante locale optimale « Bandwidth »	118
5.	Résultats de la modélisation.....	120
5.1.	Taille optimale de la classe en lecture	120
5.2.	Taille optimale de la classe en mathématiques	124
5.3.	Taille optimale de la classe en sciences	127
	Conclusion	132
	Annexes du chapitre 2	134
	Chapitre 3 : Les effets de la composition sociale et économique de l'école sur la performance scolaire des élèves marocains	149
	Introduction.....	150
1.	Revue de la littérature.....	152
1.1.	Littérature théorique et apport méthodologique.....	152
1.2.	Littérature empirique.....	158
2.	Méthodologie d'estimation	162
3.	Données utilisées et faits stylisés	168
3.1.	Les écarts d'acquis scolaire en lecture à l'école primaire selon la composition sociale et économique de la classe	170
3.2.	Les écarts d'acquis scolaire en mathématique et en science selon la composition sociale et économique de la classe.....	171
4.	Résultats de la modélisation.....	172
4.1.	Résultats pour les performances en lecture	173
4.2.	Résultats pour les mathématiques	177
4.3.	Résultats pour les sciences	181
4.4.	Discussion et recommandations	185
	Conclusion	187
	Annexes du chapitre 3	189
	Conclusion générale.....	204
	Bibliographie	210
	Index des tableaux	222
	Résumé	224
	Abstract	226

Introduction générale

Le lien entre éducation, croissance et développement représente l'une des questions les plus controversée par les économistes.

Dès le début des deux dernières décennies précédentes est soulevée une série de problématiques qui a inspiré l'orientation de cette thèse pour s'intéresser à la qualité de l'éducation et de ses multiples facteurs d'amélioration et/ou de dégradation, ainsi qu'à sa répartition pour mieux comprendre, in fine, sa contribution au développement (Curgan, 2000, Dessus, 2000, Hanchane 2000a, 2000b, Hanchane et Verdier, 2005, etc).

En effet, le débat autour du rôle de l'éducation dans la croissance et le développement ne peut échapper à la controverse qui a toujours opposé les tenants de la théorie du capital humain à ceux de la théorie du filtre et/ou du signal. C'est surtout dans le premier cas que la scolarité générale de la population se répercutera sur le revenu national (Throw, 1975). D'où l'intérêt d'établir une corrélation significative entre l'éducation et le succès économique des nations, mesuré par le P.I.B. ou le taux de croissance.

Dans la foulée des théories de la croissance endogène et des nouvelles interrogations sur l'idée d'émergence des pays, les travaux se sont multipliés ces dernières années. Selon Hanchane et Verdier (2005), ceux-ci se subdivisent en trois familles : l'approche néoclassique à la Solow, celle de la théorie de la croissance endogène de Lucas (1998) et l'approche évolutionniste de Romer (1990), R. Nelson et Phelps (1997).

Face aux enjeux théoriques de ces travaux et leur implications décisives en matière de choix de politiques publiques, les investigations empiriques se sont multipliées grâce à une meilleure production des données et au développement spectaculaire des méthodes économétriques robustes aux problèmes de spécification (biais de sélection, d'endogénéité, d'erreurs de mesure, d'hétéroscédasticité, d'hétérogénéité des effets, etc.).

Sur le plan des études empiriques, les premiers travaux ont montré une relation positive entre des mesures agrégées de la production et du capital humain (Mankiw, Romer et Weil, 1992).

Cependant, les chercheurs se sont rendu compte que ce résultat dépendait d'un grand nombre d'hypothèses restrictives, dont celle de la ressemblance des pays relativement à leur proximité par rapport à leur équilibre stationnaire et à leur capacité de profiter du progrès technique.

Les travaux qui ont suivi ont attaché une attention toute particulière à la définition du stock de capital humain et plus particulièrement à son caractère endogène et aux caractéristiques non

observées des pays. Le recours à des méthodes économétriques plus robustes indique qu'il est très difficile de faire apparaître la relation mentionnée ci-dessus (Pritchett, 1996, Hanchane et Lemelin, 2000). Le coefficient des variables de capital humain cesse souvent d'être significatif, devenant même à l'occasion négatif. De tels résultats ne vont pas sans soulever une nouvelle interrogation : comment expliquer que ce qui est bon pour les élèves et étudiants ne le soit pas pour la société ?

Ces résultats ont donné lieu à un tournant décisif laissant paraître 'une troisième vague de travaux. Plusieurs pistes en faveur du rôle de l'éducation dans la croissance et le développement ont été explorées : elles s'appuient sur une mesure plus fine du capital humain, allant au-delà du calcul du nombre d'années d'études, une meilleure prise en compte de l'hétérogénéité de l'éducation (niveau d'enseignement, qualité des acquis, dépenses etc.), et de l'hétérogénéité de ses effets, notamment des effets de complémentarité entre l'éducation et l'environnement économique et social. Ainsi les effets du capital humain seraient variables et d'autant plus forts que certaines conditions - économiques, politiques et sociales - seraient réunies.

Cependant, beaucoup de travaux soulignent qu'il ne suffit pas de calculer le nombre moyen d'années d'études des travailleurs pour établir l'importance de l'éducation et du capital humain dans la croissance économique (Benabdelaali, Hanchane, Ziroili, 2020). C'est pourtant, à peu de chose près, ce qui a été réalisé dans un certain nombre d'études. Il se pourrait, au contraire, que les possibilités de substitution entre les divers types de travailleurs et entre le travail et les autres facteurs ne soient pas parfaites. De la même façon, les rendements de l'éducation pourraient varier selon le niveau d'études et les spécialités de formation : le capital humain serait par essence hétérogène (Willis, 1986) contrairement à l'hypothèse initiale de son homogénéité faite par Becker (1964).

Mais c'est surtout autour de la prise en compte de la qualité de l'éducation dans les pays en développement que les travaux s'orientent de plus en plus.

Nous sommes bien évidemment conscients dans cette thèse du fait que la question de la qualité pose d'importants problèmes, ne serait-ce que parce qu'il est difficile de la définir, encore plus de la mesurer. D'anciens travaux ont montré qu'il n'existe pas de relation étroite entre les ressources scolaires les plus facilement mesurables - dépenses par élève, scolarité et expérience des enseignants, taux d'encadrement- et les résultats des élèves (Hanushek, 1986). De plus, les principaux spécialistes sont rarement unanimes quant aux objectifs, même les plus immédiats,

du système scolaire : faut-il augmenter les acquis moyens ou réduire les écarts d'acquis entre les élèves ?

Il ressort de la littérature sur la contribution de l'éducation à la croissance économique que la généralisation de l'accès à l'éducation ne suffit pas et que sa contribution dépend d'autres facteurs particulièrement de sa qualité.

Notons que de nombreux travaux abordant le lien entre éducation et développement (Antoine, 2009) ont négligé l'hétérogénéité de l'effet du capital humain selon les pays et à différentes périodes. La prise en considération de cette forme d'hétérogénéité revient à faire l'hypothèse tout à fait raisonnable que certaines conditions économiques et sociales sont plus favorables à la rentabilité du capital humain. La non-identification d'un effet positif du capital humain sur la croissance et le développement peut s'expliquer par la non prise en compte de la qualité des systèmes éducatifs des pays.

Les travaux empruntant cette piste (Benaabdelaali et al 2017) nous aident à mieux comprendre pourquoi l'investissement en capital humain n'a pas été récompensé par une croissance plus rapide. A titre d'exemple, nous pouvons évoquer le fait que la scolarisation massive dans les pays en développement s'est souvent faite au détriment de la qualité de l'éducation reçue ou encore que la distribution inégalitaire des services de l'éducation a eu des effets néfastes en termes d'efficacité des dépenses publiques. Un cercle vicieux est entretenu dans la mesure où la scolarisation massive, dans de mauvaises conditions, donne lieu à un faible rendement privé de l'éducation et tire vers le bas la demande de l'éducation sans oublier le fait que la concentration inégalitaire de l'éducation réduit les financements publics.

De même, l'environnement institutionnel de certains pays en développement serait tel que des activités rentables individuellement soient contre-productives socialement ; notamment la « recherche de rentes » constitue le principal débouché des diplômés. Le secteur public a absorbé dans bon nombre de ces pays une part importante des diplômés alors que sa productivité est souvent faible. Il est même possible que l'augmentation du nombre de jeunes diplômés accroisse la pression sur les gouvernements pour qu'ils créent ou maintiennent des emplois publics indépendamment de leur contribution productive.

Ce qui vient d'être décrit dans ce qui précède reflète remarquablement la situation marocaine qui est prise comme exemple d'application dans cette thèse.

En effet, nombreux sont maintenant les travaux qui montrent que le stock de compétences disponibles aujourd'hui au Maroc ne répond pas aux attentes du marché de l'emploi et, affecte durablement le développement économique du Maroc : le niveau de compétences ne contribue que modestement à la croissance et freine les possibilités d'une augmentation significative de la productivité de la main d'œuvre (Benaabdelaali, Hanchane et Ziroili, 2020).

En effet, au Maroc, comme dans l'ensemble des pays en développement, l'orientation de l'effort public vers une éducation de masse a été, certes, socialement utile mais elle a été mise en œuvre aux dépens de la qualité de l'enseignement. La crise des systèmes éducatifs de ces pays apparaît clairement par leur faible contribution à la croissance et le développement.

Conscient des enjeux que nous venons d'exposer, nombreux sont les Etats qui insistent aujourd'hui sur un Droit d'accès à une éducation de qualité. Le Maroc ne fait pas exception dans la mesure où l'article 31 de sa nouvelle constitution de 2011 stipule « la responsabilité de l'Etat, des établissements publics et des collectivités territoriales, qui doivent assurer et veiller au droit d'accès des citoyens à une éducation moderne, accessible et de qualité ».

A ces dispositions, s'ajoutent celles relatives à la création d'instances de promotion du développement humain et durable et de la démocratie participative, parmi lesquelles figurent le Conseil Supérieur de l'Education, de la Formation et de la Recherche Scientifique (CSEFRS). Ce conseil a d'ailleurs mis à la disposition du gouvernement et des citoyens la « vision 2030 » pour la réforme du système éducatif après les nombreux échecs qu'a connus l'application de la charte d'éducation et de formation et le programme d'urgence (CSEFRS, 2014).

L'échec des réformes du système éducatif marocain qui ont démarré dès le début des années 2000 avec la charte d'éducation et de formation, suivie du programme d'urgence (2008-2012) s'est manifesté par le faible accès à l'éducation et les taux d'abandons, notamment dans le Collège, les zones rurales et parmi les populations vulnérables et pauvres (un nombre moyen d'année de scolarisation de 6 ans et un indice d'inégalité qui avoisine 0,55). Certes, le Maroc a entrepris des efforts considérables pour généraliser l'accès au primaire (taux brut de scolarisation de 113,5% en 2019) avec une nette amélioration de la transition vers le collège (taux de transition du primaire au collège de 81,6% en 2019). Cependant, en raison des forts taux d'abandon en milieu de cycle, l'espérance de vie scolaire et les taux d'achèvement demeurent problématiques (le taux d'achèvement sans redoublement en primaire est de 51,7% en 2019) et ne contribue donc que très faiblement à l'amélioration de l'indicateur le plus robuste

de l'accumulation du capital humain dans une économie à savoir « la durée moyenne de scolarisation » et, à la réduction des inégalités au sens de Gini.

En effet, la durée moyenne des années de scolarisation de la population âgée de 15 ans et plus n'est que de 5 ans (parmi les plus faibles au monde). Elle est par conséquent inférieure à celle observée dans les pays dits émergents (7 années au moins) et celle des pays développés (11 ans au moins). Ce niveau faible d'éducation est accompagné par un indice de Gini, jugé parmi les plus élevés au monde ; il est de 0,56 au Maroc alors qu'il est inférieur à 0,30 pour les pays émergents et développés (Hanchane et al ,2020 ; Barro et al ,2016).

Par ailleurs sur le plan des acquis des élèves, les évaluations TIMSS¹, en 2015 et PIRLS ² en 2016 placent toujours le Maroc comme dernier ou avant dernier de la classe parmi tous les pays qui ont participé à ces dispositifs. La participation du Maroc à la dernière vague de PISA³(2018) ne fait que confirmer ces classements du pays.

Les dimensions qualitatives de l'éducation ont été suffisamment bien étudiées au Maroc et tous les travaux recommandent aux politiques publiques d'accélérer encore plus la généralisation et l'accès à tous les cycles tout en améliorant les taux d'achèvement. Cependant cette généralisation risque de se traduire à terme dans les faits par une démocratisation quantitative génératrice d'inégalités de résultats si tous les élèves n'ont pas la chance d'acquérir un socle solide concentré de compétences.

Bien au-delà du rôle décisif de la qualité de l'éducation dans le développement que nous avons mentionné précédemment, le but de cette thèse consiste à s'intéresser aux facteurs qui génèrent cette qualité et expliquent son inégalité de répartition parmi les élèves scolarisés.

La définition de la qualité retenue ici est, certes, restrictive puisqu'elle est ramenée à une mesure assez frustrante qui passe par les acquis des élèves, tels qu'ils sont évalués dans le cadre des dispositifs TIMSS et PIRLS. Cependant c'est la seule mesure dont nous disposons à l'échelle internationale et celle d'un pays, avec celle de PISA dont le protocole d'évaluation est assez différent.

¹ Trends in International Mathematics and Science Study est une enquête internationale sur les acquis scolaires, coordonnée par l'International Association for the Evaluation of Educational Achievement. Elle porte sur les mathématiques et les sciences

²Le PIRLS est un programme de l'International Association for the Evaluation of Educational Achievement basée aux Pays-Bas et qui évalue l'apprentissage de la lecture dans différents pays participants.

³Le Programme international pour le suivi des acquis des élèves, souvent désigné par l'acronyme Pisa, est un ensemble d'études menées par l'Organisation de coopération et de développement économiques, et visant à mesurer les performances des systèmes éducatifs

Cette thèse propose trois chapitres donnant lieu à des évaluations micro-économétriques dans le cadre de modèles multiniveaux innovants reposant aussi bien sur le développement, par nos propres soins, d'estimateurs à variables instrumentales (chapitre 1) et la proposition et l'adaptation de nouvelles approches semi-paramétriques en phase avec la détermination de la taille de classes (chapitre 2) et de degré de mixité sociale (chapitre 3) favorisant significativement l'amélioration des acquis.

▪ **Des bases de données hiérarchiques à deux niveaux**

Les travaux de cette thèse reposent sur la reconstitution de bases de données à partir des dispositifs TIMSS et PIRLS (2011) qui fournissent des informations détaillées sur le contexte familial et scolaire dans lequel se réalise le processus d'apprentissage d'un élève. Les données collectées proviennent des questionnaires destinés aux élèves, aux enseignants, aux parents ainsi qu'aux directeurs des établissements.

Ces enquêtes permettent de mener des comparaisons internationales et d'évaluer le système éducatif marocain par rapport à ceux des autres pays participant à ces programmes internationaux d'évaluation.

Depuis les années 2000, le Maroc participe à ces enquêtes internationales. En 2011, le Maroc a encore utilisé ce dispositif afin d'évaluer le niveau des acquis des élèves de la 4^e année primaire en lecture (PIRLS) ainsi que celui des élèves de la 2^e année de collège en mathématiques et en sciences (TIMSS). L'enquête PIRLS a concerné 325000 élèves de la 4^e année primaire dans 45 pays. Le dispositif TIMSS concerne 600000 élèves de 63 pays.

La période d'élaboration de ce dispositif est particulièrement cruciale dans la mesure où elle permet d'évaluer la qualité du système éducatif après la mise en œuvre de deux réformes importantes au Maroc à savoir la charte nationale d'éducation et de la formation en 2000⁴ et le programme d'urgence en 2009⁵.

La base de sondage de l'enquête PIRLS concerne les élèves de la 4^e année primaire scolarisés en 2011. Pour l'enquête TIMSS (2011), il s'agit des élèves de la 4^e année primaire et de la 2^e année de collège.

Le protocole de l'enquête exploite la méthode d'échantillonnage stratifié pour avoir moins de biais et pour une meilleure représentativité statistique. De plus, l'échantillon stratifié adopté

⁴ En 1999, le gouvernement se lance dans un vaste programme de réforme. La Charte nationale de l'éducation est adoptée en 1999 et la période 2000-2009 est alors qualifiée « décennie de l'éducation »

⁵ Le Maroc a adopté en 2009 un Programme d'urgence pour l'éducation pour rattraper le lourd retard du pays. Le plan d'urgence couvre la période 2009-2012, et s'appuie sur les leçons apprises au cours de la dernière décennie.

s'appuie sur une méthode d'échantillonnage à deux étapes. Le premier tirage concerne les écoles participantes à l'enquête, et le second consiste à faire apparaître les classes choisies au sein de chaque établissement scolaire.

Ainsi, l'échantillon TIMSS spécifique au Maroc comporte des informations sur 286 établissements scolaires et 7841 élèves de 4^e année primaire. Pour la 2^e année collégiale, le dispositif TIMSS concerne 8986 élèves répartis sur 279 écoles. L'échantillon du dispositif PIRLS est composé de 7805 élèves en 4^e année primaire répartis sur 284 écoles.

Après l'identification des variables sélectionnées pour la modélisation, nous avons traité et corrigé les erreurs de mesure et les non-réponses. En mobilisant les techniques d'imputation multiple (cf. chapitre 1 pour plus de détail).

Ensuite, les 4 matrices de données reconstituées, à savoir les données sur les élèves, les enseignants, les parents et les directeurs des établissements ont été regroupées en une seule base de données. Cette dernière se caractérise par sa forme hiérarchique à deux niveaux (élève et établissement) nous permettant de spécifier, estimer et tester toutes nos spécifications à l'aide de nouvelles formulations de modèles micro-économétriques adaptés à la structure de ces données.

▪ Objectifs et plan de la thèse

Comme nous l'avons mentionné dans ce qui précède, cette thèse s'intéresse aux facteurs qui expliquent ou génèrent les niveaux et les inégalités des acquis des élèves dans le but d'éclairer les choix de politiques publiques visant l'amélioration de la qualité du capital humain au Maroc.

Ainsi, **le premier chapitre** s'intéresse tout d'abord au réexamen des controverses portant sur la spécification et l'estimation des fonctions de production de capital humain avant de proposer une modélisation innovante consistant à développer une approche plus générale de correction des biais d'endogénéité que celle actuellement proposée dans la littérature suite à l'article de Hanchane et Mostafa (2012). Plus particulièrement, notre proposition consiste à proposer un estimateur à variables instrumentales dans un contexte multiniveaux afin de parvenir à l'identification des effets des facteurs de production de la qualité du capital humain non biaisés et robustes à l'hétéroscédasticité. Le débat autour des facteurs générateurs de la qualité du capital humain repose essentiellement sur des résultats empiriques qui sont ensuite traduits en termes de positions « théoriques » donnant lieu à des choix de politiques publiques radicalement

différents. En effet, l'examen de la littérature empirique se rapportant à la question des acquis scolaires nous a permis d'identifier les variables potentiellement significatives dans la détermination du niveau des performances scolaires des élèves. Les développements économétriques renvoient à l'estimation de la fonction de production en recourant à des formes fonctionnelles linéaires. Ce choix limite d'une manière ou d'une autre le champ de l'analyse empirique et les recommandations en matière de politiques éducatives qui s'ensuivent.

Ce **chapitre 1** se distingue par rapport à la littérature existante (Hanchane et al « 2011 » ; Ibourk « 2013 » ; Mourji et al « 2013 ») sur le plan national par l'adoption d'une technique économétrique que nous jugeons très pertinente et utile pour une estimation robuste. Dans ce travail, nous avons appréhendé l'estimation de la fonction de production dans un cadre micro-économétrique similaire à celui d'Hausman et Taylor. Cette spécification nécessite une connaissance préalable des variables du niveau 1(élève) et du niveau 2(école) qui ne sont pas corrélées aux effets aléatoires⁶. L'estimateur de Hausman-Taylor utilisé repose sur une série d'instruments internes et stipule que les effets non observés spécifiques à l'école sont des effets aléatoires, permettant ainsi d'introduire des variables du niveau 2 dans le modèle. De plus, les relations entre les effets aléatoires et d'autres variables du modèle sont explicitées à l'aide des instruments internes. Il s'agit en fait de l'interaction des variables corrélées aux effets aléatoires au niveau de l'école, et des variables exogènes caractéristiques de l'école.⁷

Ce chapitre parvient donc à identifier certains effets de façon robuste à l'endogénéité, et permet par conséquent d'éclairer plus rigoureusement le choix des politiques publiques censées améliorer la qualité et, réduire les inégalités des acquis des élèves marocains.

Deux résultats importants se dégagent de ce chapitre, ils concernent les effets de taille des classes et le degré de mixité sociale sur la qualité du capital humain. L'importance de ces deux facteurs nécessite des analyses plus approfondies dans la mesure où leurs effets ne peuvent être identifiés de manière rigoureuse dans un cadre linéaire comme le propose la littérature.

De ce fait, **le chapitre 2** ambitionne d'aller encore plus loin. Il propose une approche semi-paramétrique en multiniveaux robuste à tous les biais, pour parvenir, in fine, à des scénarios de

⁶ $X_{ij} = [X1_{ij} : X2_{ij}]$ et $Z_i = [Z1_i : Z2_i]$, où X_{ij} est un ensemble de variables du niveau 1, Z_i est un ensemble de variables du niveau 2, $X1$ et $Z1$ sont supposés être non corrélés avec les effets aléatoires, et $X2$ et $Z2$ sont supposés être corrélés avec les effets aléatoires

⁷I Sur la base de Chevalier et al (2005)

tailles des classes, permettant de générer des acquis des élèves marocains au moins proches de ceux des pays émergents.

La taille des classes a donné lieu à de fortes controverses aussi bien entre théoriciens qu'entre responsables des politiques publiques d'éducation : affecte-t-elle réellement les acquis des élèves ? Si oui dans quel sens ? Notre hypothèse est que ce débat reste généraliste, dans la mesure où ce qui importe au fond est de pouvoir à partir de quel nombre d'élèves les acquis scolaires commencent à se détériorer. Aussi, dans ce chapitre, nous mobilisons les données TIMSS et PIRLS (2011) afin d'identifier la taille optimale de la classe dans les deux niveaux évalués à savoir la 4^e année primaire et la 2^e année collégiale.

En effet la taille des classes a suscité de nombreux débats et controverses. L'impact de la taille des classes sur la performance scolaire des élèves a donné lieu à une vaste littérature dont les résultats et conclusions sont très hétérogènes du fait d'estimations empiriques menées sur des systèmes éducatifs très différents. La difficulté centrale réside dans le fait que la répartition des élèves sur des classes de taille différente n'est pas aléatoire, et par nature une forte corrélation entre la taille de la classe et les caractéristiques des élèves se manifeste.

Les méthodes économétriques utilisées dans les travaux empiriques menés au cours des années soixante n'étaient pas encore suffisamment développées pour prendre en considération ces difficultés. Ainsi, en raison des limites des méthodes d'estimation classiques, les premiers travaux ont abouti à des impacts non-significatifs et parfois même contradictoires. Les méta-analyses de Hanushek (1986,2003) montrent un effet non-significatif de la taille des classes, tandis que Krueger (2003) révèle des effets opposés. Cependant, les travaux récents ont essayé de mieux estimer cet impact en exploitant les nouvelles méthodes économétriques.

Ce chapitre se distingue du premier chapitre et de la revue de la littérature nationale et internationale par son apport méthodologique et les conclusions et recommandations de politiques éducatives issues des résultats que nous obtenons. Ce chapitre ne répond pas d'une manière transversale à la question de la relation entre la taille de classe et performance scolaire, mais il essaye de décomposer et de creuser davantage sur la question. Alors que dans le chapitre 1, les résultats obtenus suggèrent que la réduction de la taille des classes est une politique publique pertinente pour l'amélioration de la performance généralisée des élèves, le modèle semi-paramétrique que nous mobilisons démontre en premier lieu qu'il existe des tailles optimales de classe assurant une performance scolaire, qui peut en outre différer selon le contexte ou l'élève est assigné. Le modèle développé exploite la structure hiérarchique de nos

données en prenant en considération la non-linéarité de l'effet de la taille des classes. Il s'agit d'estimer un modèle multiniveaux semi-paramétrique dont la partie paramétrique prend en compte les variables de contrôles relatives à l'élève et à l'école et une partie semi-paramétrique contrôlée à travers notre variable d'intérêt de la taille de la classe.

Ce modèle nous permet de réaliser des simulations de la taille de la classe optimale dans différents contextes en prenant en compte entre autres le milieu de l'établissement, le genre de l'enseignant, l'âge de l'enseignant, la charge horaire, le niveau d'éducation des parents, etc. Ces résultats permettent d'identifier des scénarii possibles sur lesquels pourraient s'appuyer les décideurs publics pour améliorer le niveau des acquis scolaire des élèves.

Enfin, **le chapitre 3** aborde l'analyse des effets de la mixité sociale. De même que pour la taille de la classe, ce facteur nécessite un cadre analytique novateur pour éclairer tout à la fois les débats théoriques et ceux concernant le choix des politiques publiques visant la cohésion sociale autour de valeurs communes qui passent nécessairement par une refonte des cartes scolaires et un brassage harmonieux des groupes sociaux par l'école.

Le modèle que nous cherchons à estimer se compose d'une partie paramétrique associée aux variables de l'élève et de l'école, et une partie semi-paramétrique spécifiée à travers la variable catégorielle du niveau de composition sociale et économique de l'établissement. Ce modèle fait appel à la même structure de données multiniveaux, mais cette fois en estimant un modèle semi-paramétrique conditionné au contexte de l'élève. Ce dernier est spécifique et se distingue du modèle développé dans le chapitre 2 puisqu'il spécifie la partie semi-paramétrique à travers une variable à plusieurs modalités, à savoir les niveaux de composition sociale et économique à l'école. L'estimation d'une fonction non linéaire de ce type nécessite d'adapter les techniques des modèles semi-paramétriques à variable dichotomique et latente.

Trois types de composition sociale et économique de l'école ont été exploités afin de démontrer le degré de mixité sociale le plus juste et pertinent pour une performance scolaire généralisée. Contrairement au chapitre 2⁸, la partie semi-paramétrique est captée par une variable catégorielle, celle des 3 types de compositions sociale et économique de l'école, à savoir : une

⁸La partie semi paramétrique est captée par la variable continue celle de la taille de la classe

école défavorisée, une école favorisée et école à caractère mixte dont les seuils sont établis à travers la déclaration des directeurs des établissements.⁹

A la lumière de la revue de la littérature nationale et internationale (Hanchane et al (2012) ; Mourji et al (2013) ; Duru-Bellat et al. (2004) ; Monseur et al (2008), nos résultats empiriques ne démontrent ou n'estiment pas seulement les élasticités associées à chacune des compositions sociales et économiques de l'école, mais permettent de plus de déceler le type de composition convenable et juste pour maintenir une performance scolaire généralisée.

⁹ La variable du niveau de composition sociale et économique est une déclaration des directeurs des établissements exprimés en pourcentage des élèves issus du milieu défavorisé au sein de l'école.

Chapitre 1 : Les déterminants du rendement scolaire chez les élèves de l'enseignement primaire et collégial au Maroc : une nouvelle approche économétrique des modèles multiniveaux

Introduction

Dans un contexte mondialisé de plus en plus concurrentiel, la question de la valorisation du capital humain est devenue incontournable dans les schémas de développement, et ce aussi bien pour les pays industrialisés que pour les pays dits en transition. La question de la formation et de l'accumulation du capital humain fait nécessairement appel à la notion du rendement des systèmes éducatifs.

Dans tous les pays du monde, le secteur de l'éducation nationale revêt une importance particulière. Il permet d'abord d'endiguer le phénomène de l'analphabétisme, mais aussi et surtout de développer l'éducation et la formation des générations futures en les préparant aux défis du monde du travail. La réussite de tout système éducatif se mesure à l'aune du niveau de son rendement, lequel peut être lui-même appréhendé à travers la mesure des acquis scolaires des élèves.

Les politiques publiques en matière d'éducation nationale et de réforme du système éducatif peuvent être très diversifiées et dépendent des maux dont souffre le secteur, des ressources budgétaires disponibles mais aussi et surtout des objectifs arrêtés par les pouvoirs publics. Faut-il par exemple améliorer le climat des établissements scolaires existant ou alors construire de nouvelles unités pour augmenter le taux d'accès à l'éducation à l'échelle nationale ? La gestion des établissements scolaires devrait-elle être centralisée ou, au contraire, faudrait-il accorder plus d'autonomie aux proviseurs et directeurs d'établissements scolaires au niveau local ? La réponse à ces questions permettra de cerner les contours des politiques publiques à adopter dont l'objectif ultime est d'améliorer les acquis scolaires des élèves. Mais au-delà du cadre national, et au vu de la grande similitude des problématiques au niveau régional et parfois même au niveau mondial, l'étude de l'expérience internationale peut être très intéressante à cet égard. Les comparaisons des performances des systèmes éducatifs s'appuient sur l'exploitation des enquêtes internationales TIMSS¹⁰ et PIRLS¹¹. En effet, plusieurs travaux nationaux et internationaux exploitent ces dernières pour dévoiler les inégalités d'acquis scolaire inter et intra pays.

Les résultats de ces enquêtes rendent la comparaison internationale au niveau du rendement des systèmes éducatifs, plus justes et pertinents grâce à la standardisation des méthodologies

¹⁰Trends in International Mathematics and Science Study

¹¹Progress in International Reading Literacy Study

d'approches. L'exploitation de ces enquêtes permet de comparer les niveaux des compétences des élèves sur le plan international mais aussi d'identifier les forces et les faiblesses des différents systèmes éducatifs au sein des groupes de pays étudiés. Ces comparaisons facilitent en outre la définition des objectifs à atteindre par les pouvoirs publics, comme l'atteinte du score moyen à l'échelle internationale.

Ce chapitre aura pour objectif d'analyser les déterminants des performances scolaires des élèves marocains afin d'identifier les ingrédients des politiques publiques favorables à l'amélioration des performances scolaires.

Dans cette perspective, nous allons exploiter les données de l'enquête TIMSS et PIRLS (2011) pour le Maroc. L'examen de la littérature empirique se rapportant à la question des acquis scolaires nous permettra d'identifier les variables potentiellement significatives dans la détermination du niveau des performances scolaires enregistrées. Nous utiliserons ensuite un modèle multiniveau avec correction du problème de l'endogénéité par la méthode des variables instrumentales de Hausman Taylor. Enfin, l'analyse des résultats obtenus permettra de tirer quelques enseignements par rapport aux choix d'investissement qui seraient à même d'améliorer la qualité de l'éducation au Maroc.

Ainsi, ce chapitre se décline en cinq sections. La section 1 présente une revue de la littérature existante sur les facteurs qui peuvent influencer le rendement des élèves. Nous nous focalisons sur les caractéristiques intrinsèques de l'élève, de l'environnement familial ainsi que les caractéristiques des établissements scolaires.

Dans la section 2, nous présentons les programmes internationaux TIMSS et PIRLS, ainsi que le niveau des acquis scolaires des élèves marocains en comparaison avec les systèmes éducatifs relatifs à d'autres pays. La section 3 présente les données exploitées, le cadre théorique du modèle adopté à savoir le modèle multiniveau à deux niveaux que nous justifions.

Cependant, afin de corriger le problème de l'endogénéité engendré par la corrélation existante entre les résidus du deuxième niveau et les variables explicatives du premier niveau, nous devons recourir à la méthode des variables instrumentales de Hausman Taylor.

La section 4 présente les résultats issus de l'estimation du modèle des déterminants des performances scolaires des élèves marocains. Il s'agira notamment de souligner le rôle des différents facteurs tels que ceux associés aux caractéristiques qui sont propres aux élèves (effet élève) ou ceux liés aux établissements scolaires fréquentés (effet établissement) dans

l'explication des écarts de performances scolaires des élèves. Enfin, la section 5 conclut le chapitre.

1. Littérature empirique sur les déterminants des performances scolaires

L'étude des déterminants des performances a donné lieu à une littérature empirique abondante depuis plusieurs décennies. Bien que certaines études soutiennent que les déterminants des performances scolaires des élèves peuvent varier selon le niveau de développement du pays (Duru-Bellat, Mons, et Suchaut, 2004), d'autres suggèrent que l'élève est indépendant de son contexte social et que les écoles ont peu d'influence sur sa réussite scolaire (Coleman, et al., 1966).

La revue de la littérature empirique des déterminants des performances scolaires permet d'identifier quatre catégories de déterminants. Il s'agit des déterminants liés à l'environnement et au milieu familial de l'élève, aux caractéristiques individuelles de l'élève, aux caractéristiques des établissements scolaires et enfin les déterminants liés aux caractéristiques du système éducatif.

1.1 Caractéristiques intrinsèques de l'élève et acquis scolaires

Les caractéristiques de l'élève se réfèrent en général à l'âge, le genre, la passion pour la matière, la confiance en soi, l'implication dans les activités scolaires, etc.

Genre de l'élève et acquis scolaires

Plusieurs contributions suggèrent que les résultats scolaires des élèves sont influencés par les différences de genre. Selon ces études, les filles réussiraient mieux que les garçons dans la lecture et moins bien dans les matières quantitatives qui sont considérées comme des matières à performances masculines. Ainsi, Murphy (2000) précise que les filles et les garçons développent des stratégies d'apprentissage différentes. Les travaux de Hanchane, Benbiga, et Idir (2012) ; Linn, Hyde, et Else-Quest (2010) ; Ian, Armstrong, et Rounds (2009) ; Nosek, et al. (2009) ; Halpern et al. (2007) ; Fuchs et Wößmann (2004) ; Murphy (2000) ; Lamon, Fennema, et Hyde (1990) et bien d'autres aboutissent au même constat. Dans leur méta-

analyse¹², Ian, Armstrong, et Rounds (2009) expliquent que les hommes sont passionnés par les matières scientifiques alors que les femmes préfèrent travailler en groupe et développent de ce fait un intérêt pour les activités artistiques et sociales.

L'âge d'entrée à l'école et acquis scolaires

L'âge auquel un enfant commence sa scolarité a attiré l'attention des décideurs, des parents et des chercheurs dans le domaine de l'éducation. Cependant, l'effet de l'âge de la scolarisation sur les acquis scolaires reste mitigé. L'étude menée par Moussa (2012) montre que l'âge d'entrée à l'école n'a pas d'effet à long terme sur le niveau de scolarité et le revenu à l'âge adulte. De même, Black, Devereux, et Salvan (2008) constatent que le fait de commencer l'école à un jeune âge a un effet positif à court terme sur les revenus, mais cet effet disparaît vers l'âge de 30 ans. En revanche, Datar (2006) constate que les étudiants qui entrent à l'école un an plus tard que leurs camarades du même niveau réussissent mieux en lecture et en mathématiques que leurs homologues plus jeunes. Par ailleurs, Cascio et Lewis (2006) concluent que l'effet de commencer l'école avec une année de retard sur les performances scolaires diffère selon les méthodologies adoptées par les auteurs.

L'estime de soi et acquis scolaires

Dans les domaines académiques, l'estime de soi serait composée, selon Chapman et Tunmer (1995), de trois éléments : l'intérêt envers le domaine, l'impression d'avoir des facilités ou des difficultés et le sentiment de compétence. Alors que certaines études soutiennent que l'estime de soi chez l'élève n'est pas un élément pertinent sur les réalisations scolaires (Emler, 2002) ; d'autres la considèrent comme partie prenante du problème de l'échec scolaire (Bastard et Landrier, 2005 ; Valentine, Dubois, et Cooper 2004; Stout, 2000 ; Stevenson ,1992 ; etc.). En revanche, l'étude de Marsh et Craven (2006) montre que l'estime de soi et la réussite scolaire sont à la fois une cause et une conséquence l'une de l'autre.

Activités scolaires, efforts individuels et performances scolaires

Les activités scolaires et l'effort individuel sont considérés comme d'importants déterminants de la performance scolaire des élèves. L'étude menée par Fan, Tai, et Maltese (2012) met en

¹²Une méta-analyse est une méthode scientifique systématique combinant les résultats d'une série d'études indépendantes sur un problème donné, selon un protocole reproductible. La méta-analyse permet une analyse plus précise des données par l'augmentation du nombre de cas étudiés et de tirer une conclusion globale.

évidence une relation positive et significative entre le temps consacré à faire les devoirs scolaires et les performances aux examens standardisés. Cette conclusion a également été confirmée par Henderson et Eren (2008) ; Postlethwaite et Wiley (1992) ; Fehrmann et al. (1986).

En revanche, Cooper et Valentine (2001) indiquent que comprendre la nature des devoirs scolaires et leur influence sur la réussite scolaire permet d'isoler les effets de variables nombreuses et complexes qui influent également sur les devoirs et leur achèvement. Ces influences comprennent : (a) la manière dont les enseignants structurent et gèrent la quantité de devoirs (b) la décision des élèves de savoir si, quand et comment compléter les devoirs demandés, (c) un milieu familial propice pour faire leurs devoirs, et (d) les activités de loisirs qui attirent l'attention de l'élève.

Compte tenu de ce dernier constat, il est à signaler que d'autres travaux de recherche s'intéressent plus spécifiquement aux déterminants ou facteurs associés aux caractéristiques de l'institution scolaire ou encore à l'effet établissement.

1.2 Origine sociale, environnement familial et acquis scolaires de l'élève

Plusieurs études empiriques ont montré qu'un faible rendement scolaire est étroitement corrélé avec l'environnement familial. Ces travaux ont généralement tendance à estimer l'environnement familial par le statut socio-économique (SSE) de l'élève qui est influencé par de nombreux facteurs tels que la classe sociale de la famille, le niveau de scolarité des parents et les ressources matérielles de la famille.

En plus de ces facteurs, Nechyba, McEwan, et Older-Aguilar (1999) précisent que les caractéristiques familiales et parentales comprennent également des éléments aussi divers que les dotations génétiques qui pourraient être transmises passivement à l'enfant à travers le processus héréditaire. Ce processus façonne les caractéristiques intrinsèques de l'enfant telles que la capacité cognitive et les traits de personnalité. Par conséquent, l'analyse de la relation entre SSE de l'élève et ses performances scolaires peut être appréhendée sous plusieurs angles.

Revenu familial et performances scolaires

A partir de l'analyse de l'effet causal du revenu sur les performances des enfants en mathématiques et en lecture, Lochner et Dahl (2012) estiment qu'une augmentation du revenu de 1000 dollars augmente les scores combinés dans les deux matières de 6%, tout en précisant que ces gains sont plus importants pour les enfants issus de familles défavorisées. De leur part, Keane et Wolpin (2001) trouvent que les transferts monétaires effectués par les parents au profit de leurs enfants améliorent les résultats scolaires. Par ailleurs, Costello et Copeland (2010) montrent que l'augmentation exogène du revenu contribue à une augmentation du niveau de scolarité d'un an à l'âge de 21 ans. En revanche, selon Chevalier et Lanot (2002), les enfants issus de milieux défavorisés ont des résultats scolaires inférieurs à ceux des autres enfants tout en signalant que le mécanisme par lequel le revenu du ménage influence les résultats de l'enfant reste incertain.

Certes, la situation économique des parents est cruciale mais elle ne peut à elle seule expliquer les écarts de résultats scolaires observés entre les enfants.

Niveau d'instruction des parents et performances scolaires

L'effet positif du niveau d'instruction des parents sur les réalisations scolaires des enfants a été mis en évidence dans plusieurs travaux. En effet, Huesmann, Boxer, et Dubow (2009) suggèrent que les effets bénéfiques du niveau de scolarité des parents sur l'enfant ne se limitent pas à la réussite scolaire tout au long de l'année scolaire, mais ont aussi des implications à long terme pour la vie adulte de l'enfant. En outre, selon Chevalier et al. (2005), le niveau d'éducation de la mère aurait des effets plus importants que celui du père sur la scolarisation des enfants. Cet effet serait d'ailleurs d'autant plus important chez les garçons que chez les filles. De même, Hoff (2003) montre que les enfants de mères à statut socio-économique élevé ont des compétences linguistiques plus avancées que les enfants du même âge issus de milieux moins favorisés. Dans la même perspective, Ganzach (2000) montre qu'il existe une relation symétrique entre le niveau d'éducation de la mère et les performances scolaires de ses enfants. De même, Fuchs et Wößmann (2004) concluent que les antécédents familiaux tels que l'instruction des parents affectent les performances des élèves.

1.3 L'effet établissement et acquis scolaires

Le rapport de Coleman (1966) qui est le premier rapport publié à souligner l'influence des facteurs scolaires sur les performances des élèves est à l'origine d'une vaste littérature analysant les facteurs scolaires en relation avec les performances scolaires à savoir : les ressources matérielles de l'école, la sécurité et le civisme à l'école, la taille des classes, la qualité des enseignants et le taux d'encadrement etc.

Les ressources scolaires

Les études examinant la relation ressources scolaires-acquis scolaires des élèves aboutissent à des résultats mitigés. Alors que certaines études avancent que des ressources additionnelles ne se traduisent pas forcément par un gain de performance pour les élèves, d'autres contributions révèlent un impact positif des caractéristiques de l'école sur le rendement des élèves (Hanushek, 2006).

En effet, Barro et Lee (1996) montrent que les ressources scolaires ont un impact significatif sur les tests de compétence des élèves. Par ailleurs, l'étude méta-analytique menée par Greenwald et Hedges (1996) révèle que les résultats des élèves sont positivement liés à un large éventail de ressources. Al Samarrai (2002) et Häkkinen, Kirjavainen, et Uusitalo (2003) aboutissent à la même conclusion. En revanche, les analyses menées à partir de données transversales montrent que l'effet des ressources scolaires varie selon le niveau de développement des pays. Selon Michaelowa, 2001 ; Fuller et Clarke 1994 et Heyneman et Loxley, 1983 cet effet serait plus important dans les pays en développement.

La taille des classes

L'effet de la taille des classes sur les performances scolaires est toujours controversé. En effet, Kingdon et Altinok (2012), Hanushek (1997 ; 2003 ; 2006) et Hoxby (2000) constatent que les effets de la taille des classes sont moindres dans les pays disposant d'enseignants performants capables de gérer les classes encombrées. Par ailleurs, Woessmann et West (2006) estiment l'effet de la taille des classes sur le rendement des élèves dans 18 pays et montrent qu'à quelques exceptions près, des classes plus petites présentent des effets bénéfiques uniquement dans les pays où les salaires des enseignants sont relativement faibles. Pour sa part, Krueger (1999 ; 2003) révèle que les élèves issus de classes à effectif réduit réussissent mieux.

Enfin, selon d'autres études, l'effet de la taille des classes sur les performances scolaires varie selon les niveaux scolaires. Ainsi, Lavy et Angrist (1999) montrent que la réduction de la taille des classes induit une augmentation significative et substantielle dans les résultats des tests pour les élèves de fin du cycle primaire.

L'éducation préscolaire pour une meilleure préparation à l'école

L'éducation préscolaire permet à l'enfant l'acquisition précoce et durable des gestes et des compétences qui préservent et protègent son capital santé. Elle favorise aussi une initiation précoce à la citoyenneté et aide l'enfant à atteindre de meilleurs résultats scolaires réduisant ainsi les risques de l'échec scolaire.

L'étude récente menée par Sandoval-Hernandez, Taniguchi, et Aghakasis (2013) montre que dans près de la moitié des 37 pays examinés, l'éducation préscolaire est positivement et significativement associée à de meilleures performances en mathématiques, quel que soit le niveau socio-économique familial de l'élève. De même, analysant l'effet à long terme d'une éducation préscolaire, Chetty et al. (2010) mettent en perspective une forte corrélation entre les résultats scolaires obtenus à l'école maternelle et le rendement éducatif observé sur le marché du travail à l'âge de 27 ans. Berlinski, Galiani, et Manacorda (2008) considèrent que l'éducation préscolaire représente un instrument efficace pour éviter l'échec scolaire dans les pays à faible revenu car les gains induits par cet enseignement s'amplifieraient à fur et à mesure que les enfants grandissent. En parallèle, Caille (2001) montre que les enfants inscrits dans le préscolaire à l'âge de deux ans ont moins de risque de redoubler comparativement aux élèves scolarisés après l'âge de trois ans. Enfin, Shultz (2009) conclut que la fréquentation d'un établissement d'enseignement préscolaire est positivement associée aux performances scolaires à l'âge de 15 ans mais que l'ampleur de cette relation varie considérablement entre les pays. Au-delà du rôle de l'enseignement préscolaire dans l'amélioration des acquis scolaires, d'autres éléments tels que la composition sociale de l'école ou encore «les effets de pairs ¹³» exercent également une influence déterminante sur la performance scolaire de l'élève.

¹³Les pairs désignent les agents avec lesquels un individu est en interaction. Dans le cas de l'éducation, les pairs peuvent être les camarades de classe à l'école primaire, les camarades de travaux dirigés à l'université, ou encore les compagnons de chambres sur le campus universitaire

1.4 Effets de pairs

Le concept des effets de pairs ou de composition fait référence aux théories de l'interaction sociale. Dans le domaine de l'éducation, ce concept signifie que l'élève évolue au sein de l'établissement avec une population qui l'influence et qui est influencée par lui (Hanchane, Benbiga, et Idir, 2012).

En effet, Flashman (2012) montre que la similitude dans la réussite scolaire est un prédicteur important et cohérent des liens d'amitié dans la mesure où les élèves très performants sont plus susceptibles d'étendre les liens avec d'autres étudiants très performants. Pour sa part, Hoxby (2000) souligne qu'un changement crédible exogène de 1 point dans les notes de lecture des pairs contribue à augmenter le score d'un élève entre 0,15 et 0,4 points, selon la spécification du modèle économétrique et la méthodologie d'estimation utilisée. De même, Hanushek et al (1999 ; 2003) montrent que les élèves semblent tirer bénéfice de la présence de camarades performants dans la mesure où une augmentation de 1 point dans la note moyenne des pairs améliore la note en mathématiques des élèves de 0,17 point. Enfin, mesurant les effets de pairs parmi les colocataires et les camarades de dortoirs dans une université, Sacerdote (2001) aboutit à la conclusion que l'existence d'effets de pairs influence largement les résultats des élèves.

Devant la multiplicité des travaux qui analysent les déterminants des performances scolaires des élèves à travers le monde entier, il s'avère essentiel, de s'intéresser aux résultats obtenus dans le cas particulier du Maroc

1.5 Les travaux empiriques analysant la performance scolaire des élèves marocains

Il existe peu d'études empiriques qui traitent la question des acquis scolaire des élèves marocains. Benbiga, Hanchane, Idir et Mostafa (2011) ont analysé les facteurs explicatifs de la qualité des apprentissages des élèves marocains de la 4^e année primaire et de la 2^e année collégiale. En utilisant les données issues de l'enquête nationale sur l'évaluation des acquis des élèves marocains, et en adoptant l'approche de Mundlak (1978) dans un contexte hiérarchique pour la correction du biais d'endogénéité, les auteurs montrent que plusieurs composantes participent à l'explication des performances des élèves marocains. Les conditions matérielles, la taille de la famille, les antécédents des élèves, l'environnement scolaire, l'encadrement

familial sont tous des facteurs qui contribuent de manière significative à l'explication des rendements scolaires des élèves marocains.

Ibourk (2012) a analysé les déterminants de la mauvaise performance des élèves marocains en utilisant la régression quantile et l'analyse ANOVA à plusieurs facteurs. En exploitant les données TIMSS et PIRLS (2007), l'auteur aboutit aux mêmes conclusions que celles issues des études réalisées pour les pays en développement. A partir des résultats obtenus, l'auteur formule certaines propositions et recommandations. Ainsi, il met en perspective l'importance d'un enseignement de qualité et d'une généralisation du préscolaire dans le milieu rural dans la lutte contre les inégalités spatiales. En outre, Les résultats obtenus recommandent d'encourager l'enseignement privé pour élargir le dispositif de l'éducation nationale tout en mettant en place un contrôle, un suivi et une évaluation de ce secteur.

Mourji et Abbaia (2013) ont analysé les déterminants du rendement scolaire en mathématiques chez les élèves de l'enseignement secondaire collégial au Maroc. En exploitant les données TIMSS (2003), et en spécifiant un modèle hiérarchique, les auteurs montrent que les performances scolaires sont expliquées essentiellement par le genre de l'élève et les attributs matériels du ménage. En revanche, les variables liées au contexte scolaire telles que le milieu d'implantation de l'école, l'âge de l'enseignant et son ancienneté n'ont pas un effet significatif sur les performances scolaires des élèves.

2. Description du dispositif des deux programmes internationales TIMSS-PIRLS 2011 : Maroc et analyse en benchmark

Les enquêtes internationales TIMSS et PIRLS (2011) fournissent des informations très pertinentes et très détaillées sur le contexte familial et scolaire dans lequel se réalise le processus d'apprentissage. Les données collectées proviennent des questionnaires destinés aux élèves, aux enseignants, aux parents ainsi qu'aux directeurs des établissements.

Ces enquêtes permettent de mener des comparaisons internationales très intéressantes et plus particulièrement d'évaluer le système éducatif marocain vis-à-vis des autres systèmes éducatifs participant à ces programmes internationaux.

En 2011, le Maroc a participé à ces enquêtes afin d'évaluer le niveau des acquis des élèves de la 4^e année primaire en lecture (PIRLS) ainsi que celui des élèves de la 2^e année collégial en mathématiques et en sciences (TIMSS).

L'enquête PIRLS a concerné 325000 élèves de la 4^e année primaire dans 45 systèmes éducatifs distribués sur l'ensemble des continents du monde. De même, l'enquête TIMSS inclut des informations collectées auprès de 600000 élèves répartis dans 63 pays afin de réaliser un diagnostic des apprentissages en mathématiques et en science.

La période d'élaboration de ce dispositif est particulièrement cruciale dans la mesure où elle permet d'évaluer la qualité du système éducatif après la mise en œuvre de deux réformes importantes au Maroc à savoir la Charte nationale ¹⁴d'éducation et de la formation en 2000 et le programme d'urgence¹⁵ en 2009.

Aussi, nous commençons par une analyse descriptive des résultats de l'enquête TIMSS et PIRLS. Il s'agit de présenter les résultats des données collectées auprès des élèves, des parents et des directeurs des établissements scolaires.

Ensuite, nous présentons les résultats des acquis scolaires des élèves marocains, tout en examinant leurs corrélations avec les variables contextuelles, à savoir le milieu de l'établissement, l'éducation des parents, l'ancienneté de l'enseignant, etc.

2.1 Présentation de l'enquête

2.1.1. Echantillonnage

La base de sondage de l'enquête PIRLS concerne les élèves de la 4^e année primaire scolarisés en 2011. Pour l'enquête TIMSS (2011), il s'agit des élèves de la 4^e année primaire et de la 2^e année collégiale.

Concernant la méthode d'échantillonnage, l'IEA ¹⁶a adopté la méthode d'échantillonnage stratifié pour réduire les biais et pour une meilleure représentativité statistique. De plus, l'échantillon stratifié adopté s'appuie sur une méthode d'échantillonnage à deux étapes. Le premier tirage concerne les écoles participantes à l'enquête, et le second consiste à faire apparaître les classes choisies au sein de chaque établissement scolaire.

¹⁴ En 1999, le gouvernement se lance dans un vaste programme de réformes. La Charte nationale de l'éducation est adoptée en 1999 et la période 2000-2009 est alors qualifiée « décennie de l'éducation »

¹⁵ Le Maroc a adopté en 2009 un Programme d'urgence pour l'éducation pour rattraper le lourd retard du pays. Le plan d'urgence couvre la période 2009-2012, et s'appuie sur les leçons apprises au cours de la dernière décennie.

¹⁶ L'IEA (The International Association for the Evaluation of Educational Achievement) est une coopérative internationale d'institutions de recherche nationales, d'agences de recherche gouvernementales, de chercheurs et d'analystes qui travaillent à évaluer, comprendre et améliorer l'éducation dans le monde entier.

Comparée aux méthodes d'échantillonnage utilisées par l'OCDE 17 et l'IEA, la méthode de stratification des enquêtes TIMSS et PIRLS permet de fournir une information plus fine sur les profils des enseignants

Concernant l'enquête TIMSS, l'échantillon final comporte des informations sur 286 établissements scolaires et 7841 élèves marocains de 4^e année du primaire et 279 écoles incluant 8986 élèves en 2^e année collégiale. S'agissant de l'enquête PIRLS, les données concernent 284 écoles et 7805 élèves en 4^e année primaire.

2.1.2. Profil des élèves

Commençons tout d'abord par examiner le profil des élèves selon le genre, l'âge et la préscolarisation des élèves.

- **Genre**

Les résultats concernant la 4^e année du primaire montrent les filles représentent 48% de l'échantillon total contre 52% de garçons. La répartition est quasiment similaire pour les élèves de la 2^e année collégiale (Tableau 1).

Tableau 1. Répartition des élèves selon le genre

	4 ^{ème} année primaire	2 ^{ème} année collégiale
Filles	48%	47%
Garçons	52%	53%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Âge**

Nous observons que l'âge moyen des élèves de la 4^e année primaire est de 10 ans, sachant que l'âge d'entrée de l'élève à ce niveau d'enseignement en 2011 doit être de 9 ans.

Ainsi, presque la totalité des élèves ont un retard scolaire de plus d'un an : 62 % des élèves de la 4^e année primaire ont un âge supérieur à 9 ans, et 68% des élèves de la 2^e année collégiale sont âgés de plus de 13 ans. Ces retards scolaires tiennent essentiellement au redoublement et à la rentrée tardive des élèves en première année du cycle primaire. En revanche, les élèves en avance par rapport à l'âge légal ne représentent que 1 % de la distribution totale pour les deux niveaux d'enseignements évalués (Tableau 2).

¹⁷Organisation de Coopération et de Développement Économiques

Tableau 2. Répartition des élèves selon l'âge

4 ^{ème} année primaire		2 ^{ème} année collégiale	
Moins de 9 ans	1%	Moins de 13 ans	1%
9 ans (l'âge légal)	37%	13 ans (l'âge légal)	31%
Plus de 9 ans	62%	Plus de 13 ans	68%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Précolaire**

Concernant la préscolarisation des élèves, 78% des élèves des deux niveaux d'enseignement ont bénéficié de la préscolarisation. Ce résultat doit être interprété en tenant compte de l'âge d'entrée à la première année d'enseignement scolaire. En effet, la première année du primaire a concerné les élèves actuellement en 2^e année collégiale en 2003 et ceux en 4^e année primaire. Par ailleurs, après la mise en place de la Charte nationale d'éducation et de formation en 2000, cette période a été une période de généralisation de l'enseignement pour l'ensemble des cycles d'éducation, et du préscolaire en particulier (Tableau 3).

Tableau 3. Répartition des élèves de la 4^{ème} année primaire selon le nombre d'années passées dans le préscolaire

	TIMSS	PIRLS
N'a pas fait du préscolaire	23%	22%
A fait le préscolaire	77%	78%
Un an ou moins	17%	17%
Plus d'un an et moins de 3 ans	39%	39%
3 ans ou plus	21%	22%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

2.1.3. Profil des parents d'élèves

Il s'agit de décrire le contexte familial des élèves enquêtés selon les différentes dimensions que sont le niveau d'instruction des parents, la situation professionnelle des parents ainsi que d'autres caractéristiques liées à la famille.

- **Langue parlée à la maison**

Concernant la langue parlée à la maison, nous observons que 83% des élèves de la 4^e année primaire pratiquent l'arabe à la maison, 22% d'entre eux parlent l'amazigh et le reste le français, l'espagnol et l'anglais.

Tableau 4. Répartition des élèves de la 4^{ème} année primaire selon la langue parlée à la maison¹⁸

Langues parlées à la maison	% des élèves
L'arabe	83%
L'amazigh	22%
Le français	19%
L'espagnol	4%
L'anglais	5%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

Pour les élèves de la deuxième année collégiale, 63% des élèves parlent l'arabe à la maison, 29% le parlent quelques fois, et seulement 8% n'ont jamais pratiqué la langue arabe.

- **Niveau d'instruction des parents ou tuteurs des élèves**

Les résultats du tableau 5 révèlent que 62% des élèves de la 4^e année du primaire ont une mère analphabète et seulement 6% ont une mère diplômée du supérieur. Ces chiffres sont de 66% et 11% respectivement pour les élèves de la 2^e année collégiale.

Tableau 5. Répartition des élèves selon le plus haut niveau d'instruction des parents ou tuteurs

	4 ^{ème} année primaire		2 ^{ème} année collégiale	
	Niveau d'éducation du père	Niveau d'éducation de la mère	Niveau d'éducation du père	Niveau d'éducation de la mère
Primaire, préscolaire ou analphabète	46%	62%	52%	66%
Collégial	27%	21%	11%	12%
Qualifiant	17%	11%	15%	11%
Supérieur	10%	6%	22%	11%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Situation professionnelle des pères ou tuteurs des élèves**

Concernant la situation professionnelle des parents, les résultats du tableau 6 indiquent que 64% des élèves ont des pères qui occupent un travail rémunéré, 29% d'entre eux ont des pères occupant un emploi à temps plein, et 35% déclarent avoir des pères avec un emploi à temps partiel. Enfin, seuls 10 % des élèves ayant des pères occupant un emploi non rémunéré alors que ce pourcentage devient 44% lorsqu'il s'agit des mères.

¹⁸Les résultats statistiques ne représentent pas 100% en total, car la variable est en choix multiples.

Tableau 6. Répartition des élèves de la 4^{ème} année primaire selon la situation d'emploi des parents ou tuteurs

Type d'emploi occupé	Père/Tuteur	Mère/Tutrice
Emploi rémunéré à temps plein	29%	18%
Emploi rémunéré à temps partiel	35%	15%
Emploi non rémunéré	10%	44%
Autre	26%	22%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Conditions de vie familiales**

La disponibilité de livres, d'une chambre individuelle, et d'une connexion internet représentent des indicateurs de la situation et des conditions de vie familiale de l'élève. Selon le tableau 7, 61% des élèves de la 4^{ème} année primaire disposent de plus de 10 livres chez eux, 19% ont entre 11 et 25 livres et 20% déclarent avoir plus de 25 livres.

Plus d'un tiers disposent d'une chambre individuelle, 42% détiennent un bureau, et près de 30% bénéficient d'une connexion internet. Le tableau 7 présente également la répartition des élèves de la 2^{ème} année collégiale. Les résultats indiquent des conditions manifestement plus favorables pour les élèves de la 2^{ème} année collégiale par rapport à ceux de la 4^{ème} année primaire.

Tableau 7. Répartition des élèves selon leurs conditions de vie familiales

	4 ^{ème} année primaire	2 ^{ème} année collégiale
Livres à la maison		
0-10 livres	61%	32%
11-25 livres	19%	44%
Plus de 25 livres	20%	24%
Chambre individuelle	35%	35%
Ordinateur	35%	48%
Bureau	42%	55%
Connexion internet	31%	40%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

2.1.4. Profil des enseignants

L'analyse du profil des enseignants révèle des informations intéressantes. L'enseignant joue un rôle déterminant dans le processus d'apprentissage et le transfert des savoirs. De fait, certains aspects de l'enseignant, à savoir : le genre, l'ancienneté, le niveau d'instruction et formation continue peuvent avoir un impact sur le processus d'acquisition des compétences de l'élève.

- **Genre de l'enseignant**

Tableau 8. Répartition des enseignants selon le genre

	4 ^{ème} année primaire			2 ^{ème} année collégiale		
	Mathématiques	Sciences	Lecture	Mathématiques	Physique-Chimie	Sciences de la vie et de la terre
Femmes	50%	46%	48%	29%	19%	56%
Hommes	50%	54%	52%	71%	81%	44%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

Selon le tableau 8, les élèves de la 4^e année du primaire sont répartis de manière équitable entre enseignants hommes et femmes quelque soient les matières enseignées.

En revanche, pour la 2^e année collégiale, le tableau 8 indique que les mathématiques et la physique-chimie sont des matières plutôt enseignées par des hommes (reps. 71% et 81%) que par des femmes. En revanche, ce biais de genre s'inverse pour l'enseignement des sciences de la vie et de la terre avec 56% d'enseignantes femmes.

- **Ancienneté**

L'ancienneté de l'enseignant est considérée aussi comme un facteur pouvant expliquer les écarts de performance entre les différents élèves. Les statistiques observées pour les deux niveaux montrent que le nombre moyen d'années d'expérience des enseignants varie entre 19 et 22 ans. En 4^e année primaire, le nombre moyen d'ancienneté est de 20 ans pour les enseignants de mathématiques, et de 21 ans pour les enseignants de matières scientifiques (physique-chimie et sciences de la vie et de la terre. Pour la 2^e année collégiale, ces résultats avoisinent 22 ans pour les enseignants de mathématiques et 19 ans pour les enseignants des matières scientifiques (Tableau 9). Globalement, les statistiques ne montrent pas de différences entre les niveaux d'étude.

Tableau 9. Répartition des enseignants selon le nombre d'année d'expérience

	4 ^{ème} année primaire			2 ^{ème} année collégiale		
	Mathématiques	Sciences/ Lecture	Mathématiques	Physique-Chimie	Sciences de la vie et de la terre	
Moins de 5 ans	8%	5%	15%	11%	10%	
[5,10[8%	7%	5%	7%	11%	
[10,20[33%	33%	11%	29%	27%	
20 ans et plus	51%	55%	69%	53%	52%	

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Niveau d'instruction**

Concernant le niveau d'instruction des enseignants, selon les résultats du tableau 10, la majorité des enseignants ont au moins un diplôme du baccalauréat. Les enseignants ayant obtenu un diplôme de licence représentent une part importante indépendamment des matières enseignées et des niveaux d'enseignement. En revanche, les diplômés du niveau Master et/ou Doctorat ne représentent qu'une minorité.

Tableau 10. Répartition des enseignants selon le niveau d'instruction

	4 ^{ème} année primaire			2 ^{ème} année collégiale		
	Mathématiques	Sciences	Lecture	Mathématiques	Physique-Chimie	Sciences de la vie et de la terre
Collégial au plus	11%	9%	9%	1%	0%	0%
Baccalauréat	55%	56%	51%	79%	58%	55%
Licence	33%	35%	40%	19%	38%	41%
Master/Doctorat	1%	0,24%	0%	1%	4%	4%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Formation continue**

Par rapport à la formation continue des enseignants, les résultats (Tableau 11) indiquent que plus des deux tiers des enseignants de la 4^e année primaire n'ont jamais bénéficié ces deux dernières années de sessions de formation continue. En revanche, 81% des enseignants de la 2^e année collégiale ont participé à des formations continues.

Tableau 11. Répartition des enseignants selon la participation à la formation continue durant les deux dernières années

	4 ^{ème} année primaire			2 ^{ème} année collégiale		
	Math	Sciences	Lecture	Math	Physique-Chimie	Sciences de vie et de de la terre
N'ayant jamais participé	73%	76%	71%	19%	15%	14%
Ayant participé dans les domaines de:	27%	24%	29%	81%	85%	86%
Contenu	14%	7%	-	38%	47%	52%
Pédagogie	18%	9%	-	52%	65%	63%
Curricula	16%	9%	-	41%	52%	61%
TIC	8%	6%	-	60%	59%	42%
Evaluation	16%	5%	-	32%	40%	53%
Traitement des besoins spécifiques des élèves	17%	13%	-	19%	16%	16%

Source : Calcul des auteurs à partir des enquêtes TIMSS et PIRLS 2011

2.1.5. Caractéristiques des établissements scolaires

Les écoles marocaines présentent des disparités importantes au niveau des inputs (ressources matérielles et humaines par exemple), mais également au niveau des outputs mesurés par les acquis scolaires des élèves (score des élèves). Les statistiques suivantes permettent de rendre compte des disparités relatives aux caractéristiques des établissements scolaires

- **Milieu de l'établissement**

Nous observons que plus de 50% des écoles se situent en milieu urbain. Les écoles collégiales sont plutôt localisées dans les villes de taille grande ou moyenne avec un pourcentage de 26%, contre seulement 11% pour les écoles primaires. Les écoles situées dans les petites villes ou en zone rurale se sont de l'ordre de 28% et 21% respectivement.

Tableau 12. Milieu de l'établissement

	4 ^{ème} année primaire	2 ^{ème} année collégiale
Zone urbaine densément peuplées	31%	41%
Suburbain	19%	10%
Ville de taille moyenne ou grande ville	11%	26%
Petite ville ou le village	28%	21%
zone rurale éloignée	11%	2%

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Composition de l'établissement**

La composition sociale et économique de l'école indique le pourcentage des élèves issus des classes favorisées ou défavorisées.¹⁹

Selon le tableau 13, 75% des écoles primaires enquêtées sont composées d'élèves issus des classes défavorisées. Cette proportion s'élève à 81 % pour les écoles collégiales.

Tableau 13. Répartition des élèves selon la composition des établissements

	4 ^{ème} année primaire	2 ^{ème} année collégiale
Plus d'élèves aisés que d'élèves défavorisés	12%	6%
Pas plus d'élèves aisés que d'élèves défavorisés	13%	13%
Plus d'élèves défavorisés que d'élèves aisés	75%	81%

Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

¹⁹Les directeurs d'établissements renseignent toute l'information nécessaire sur cette caractéristique de l'école pour mesurer l'ampleur des phénomènes de la ségrégation scolaire entre les différentes écoles marocaines.

- **Ressources de l'établissement**

Concernant les ressources de l'établissement scolaire, 33% des écoles primaires disposent d'une bibliothèque, et seulement 3% d'entre eux ont un laboratoire de science. Ce dernier chiffre avoisine les 82% dans les écoles collégiales.

Tableau 14. Répartition des élèves selon les ressources de l'école

	4 ^{ème} année primaire	2 ^{ème} année collégiale
Bibliothèque	33%	-
Laboratoire de sciences	3%	82%
Nombre d'élèves par Ordinateur		
Pas d'ordinateurs	31%	13%
1 ordinateur pour 6 élèves et plus	49%	70%
1 ordinateur pour 3-5 élèves	9%	10%
1 ordinateur pour 1-2 élèves	11%	6%

Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

De même, 31% des écoles primaires ne disposent pas d'ordinateurs, contre 13% seulement pour les écoles collégiales. En ce qui concerne les écoles bénéficiant des ordinateurs, les résultats montrent un ratio élevé du nombre d'élèves par ordinateurs pour les deux niveaux d'enseignement. Ce dernier est respectivement de l'ordre de 49% et 70% pour la 4^e année primaire et la 2^e année collégiale respectivement.

- **Conditions de travail des enseignants**

Les conditions de travail des enseignants sont appréhendées à partir d'un indice composite²⁰ qui indique l'ampleur des problèmes rencontrés par les enseignants. Cet indice a été construit en se basant sur plusieurs éléments à savoir l'encombrement des classes, la charge horaire, l'état des classes et les ressources matérielles et pédagogiques. Les réponses aux items comportent pour chaque indicateur trois modalités : pas de problème, problème moindre, problème important.

Tableau 15. Répartition des élèves selon les conditions de travail des enseignants

	4 ^{ème} année primaire			2 ^{ème} année collégiale		
	Math	Sciences	Lecture	Math	Physique-Chimie	Sciences de La vie et De la terre
Presque pas de problème	6%	5%	5%	4%	7%	7%
Problèmes mineurs	16%	19%	20%	34%	25%	25%
Problèmes modérés	78%	75%	76%	62%	69%	68%

Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

²⁰ L'indice composite proposé dans l'enquête est calculé sur la base d'une analyse en composante principale

Ainsi, selon les résultats du tableau 15, entre 62 % et 78% des enseignants selon le niveau d'enseignement et les matières enseignées ont déclaré avoir rencontré des problèmes modérés dans leur vie quotidienne à l'école.

- **Climat de l'établissement**

Selon les déclarations des directeurs des établissements, 60% des écoles primaires et 49% des écoles collégiales se caractérisent par des problèmes de discipline et de sécurité. La tricherie, le vol et le vandalisme sont les trois problèmes les plus fréquents. Cela vaut également pour le niveau collégial avec des différences négligeables au niveau des pourcentages des deux niveaux d'enseignements. A titre d'illustration, 33% des écoles collégiales ²¹ont déclaré être confrontées à des problèmes de violence physique envers les enseignants (Tableau 16).

Tableau 16. Pourcentage des écoles avec des problèmes de discipline et de sécurité

	4^{ème} année primaire	2^{ème} année collégiale
Retard	20	29
Absentéisme	38	36
Turbulence	38	31
Triche	45	30
Insulte	41	32
Vandalisme	48	36
Vol	52	35
Menace ou violence verbale envers les élèves	32	20
Violence physique envers les élèves	37	30
Violence verbale envers les enseignants	36	29
Violence physique envers les enseignants	-	33

Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

2.2. Performance scolaire des élèves marocains : Une analyse descriptive en benchmarking

Cette section se donne pour objectif d'analyser les scores obtenus par les élèves marocains participant aux enquêtes internationales TIMSS et PIRLS 2011, d'en examiner l'évolution depuis 1999 et de comparer le système éducatif marocain à l'aune des autres systèmes éducatifs.

²¹Déclaration des directeurs d'établissement

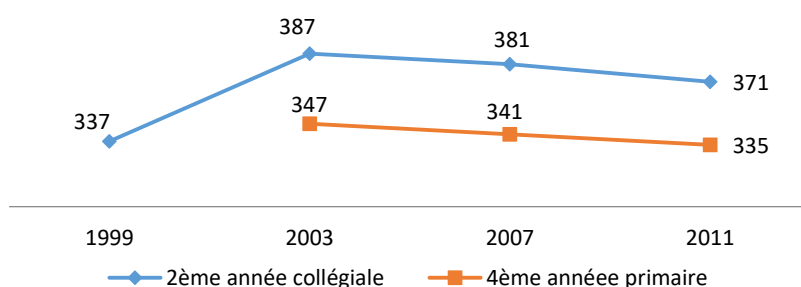
2.2.1 Performances et classement des élèves marocains

• Evolution des performances des élèves marocains

L'enquête TIMSS confirme la faiblesse du niveau des acquis des élèves marocains. Entre 2003 et 2011, le niveau des acquis scolaire des élèves de la 4^e année primaire a chuté de 12 points en mathématiques, passant d'un score de 347 à 335 (Graphique 1), tandis que le score en sciences a reculé de 40 points passant de 304 à 264 au cours de la même période (Graphique 2).

En 2^eème année collégiale, le niveau des acquis en mathématiques a connu une hausse entre 1999 et 2003 passant de 337 points à 387 (Graphique 1). Cette hausse a été suivie par une diminution de 387 points à 371 points entre 2003 et 2011. Le score en sciences qui était de l'ordre de 402 en 2007 a diminué de 26 points en 2011 (Graphique 2).

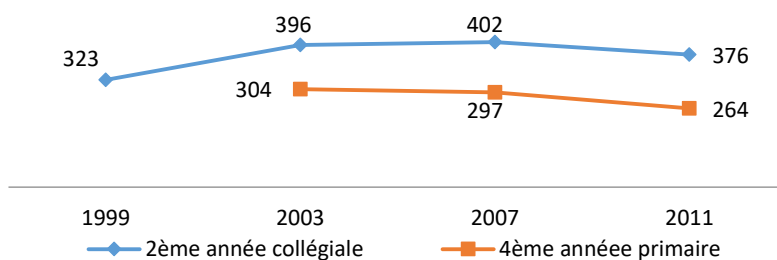
Graphique 1. Evolution des scores en mathématiques



Source : Rapports sur les résultats internationaux en mathématiques, enquêtes TIMSS

En 2^{ème} année collégiale, le niveau des acquis en mathématiques a connu une hausse entre 1999 et 2003 passant de 337 points à 387. Cette hausse a été suivie par une diminution de 387 points à 371 points entre 2003 et 2011. Le score enregistré en matière scientifique était de l'ordre de 402 en 2007, après une diminution de 26 points en 2011.

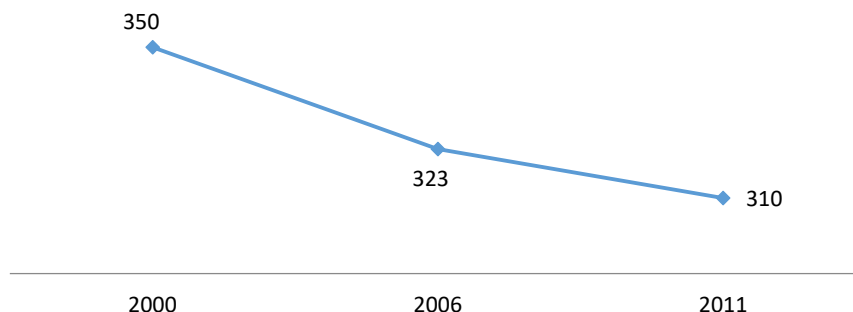
Graphique 2 . Evolution des scores en sciences



Source : Rapports sur les résultats internationaux en sciences, enquêtes TIMSS

L'enquête internationale PIRLS confirme également la faiblesse des acquis des élèves en lecture. La tendance de cet indicateur a connu une chute de 40 points passant de 350 points en 2001 à 310 en 2011 (Graphique 3).

Graphique 3. Evolution des scores des élèves de la quatrième année primaire en lecture



Source : Rapports sur les résultats internationaux en lecture, enquêtes PIRLS

- **Le classement du Maroc à l'échelle internationale**

L'analyse en benchmarking des enquêtes TIMSS et PIRLS indique que selon les scores obtenus par les élèves marocains, le Maroc se situe en avant dernière voire en dernière position dans le classement parmi l'ensemble des pays participants à cette enquête.

Concernant les élèves de la 4^e année primaire, l'analyse comparative révèle que 54% des pays participant à ces programmes internationaux ont obtenu des scores en mathématiques proches de la moyenne internationale (score de 500 points).

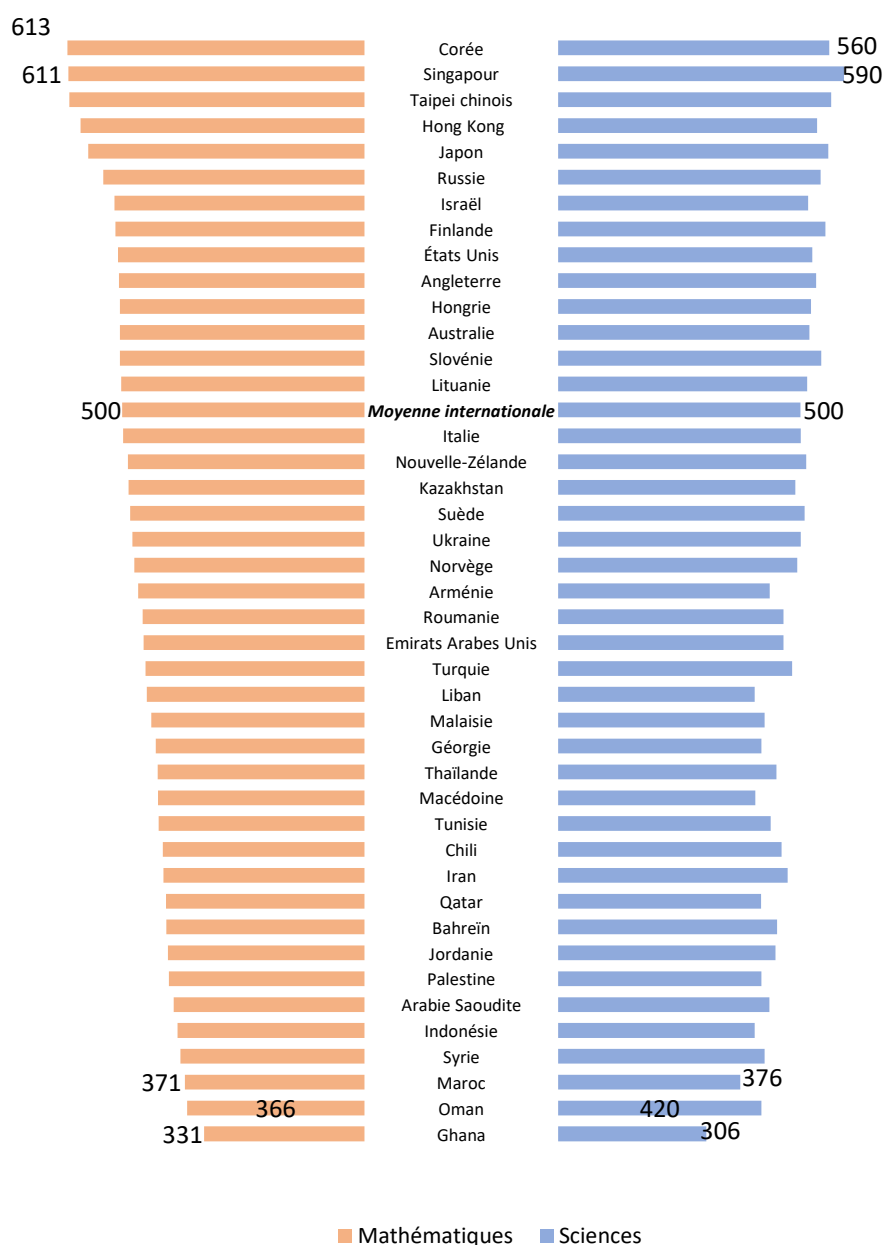
Les scores les plus élevés ont été enregistrés par les pays d'Asie de l'est (Singapour, la Corée du Sud, Hong Kong, le Taipei Chinois et le Japon) avec des scores variant de 585 à 606 points en mathématiques. Le Maroc occupe l'avant-dernière place avec des écarts importants qui s'élèvent à 271 points par rapport à Singapour et à 165 points par rapport à la moyenne internationale (Tableau 17).

De même, dans le domaine des sciences, les pays asiatiques sont les plus performants avec des scores de 587 et 583 points respectivement pour la Corée et Singapour. Il apparaît que 20 pays parmi les pays participant à l'enquête ont des scores inférieurs à la moyenne internationale parmi lesquels figurent le Maroc et le Yémen, avec les scores les plus faibles respectivement 376 et 370.

Les mêmes tendances peuvent être observées pour les élèves de la 2^e année collégiale (Graphique 4). Le Maroc, Oman, Arabie saoudite et la Syrie occupent les derniers rangs, avec

des scores plus faibles que ceux observés pour les élèves de la 4e année primaire. Les écarts de performance par rapport à la moyenne internationale sont de l'ordre de 129 points en mathématiques et de 124 points en sciences. Ainsi, les pays les plus performants sont la Corée du sud, Hong Kong et le Japon avec des scores avoisinant les 600 points.

Graphique 4. Résultats internationaux des élèves de la deuxième année collégiale



Source : Rapports sur les résultats internationaux, enquêtes TIMSS et PIRLS 2011

Tableau 17. Résultats internationaux des élèves de la quatrième année primaire

Mathématiques		Sciences		Lecture	
Singapour	606	Corée	587	Hong Kong	571
Corée	605	Singapour	583	Finlande	568
Hong Kong	602	Finlande	570	Russie	568
Taipei chinois	591	Japon	559	Singapour	567
Japon	585	Taipei chinois	552	Irlande du nord	558
Irlande du nord	562	Russie	552	États-Unis	556
Belgique (flamande)	549	États-Unis	544	Danemark	554
Finlande	545	République Tchèque	536	Taipei chinois	553
Angleterre	542	Hong Kong	535	Croatie	553
Russie	542	Hongrie	534	Angleterre	552
États-Unis	541	Suède	533	Irlande	552
Pays-Bas	540	Autriche	532	Canada	548
Danemark	537	République slovaque	532	Pays-Bas	546
Lituanie	534	Pays-Bas	531	République Tchèque	545
Portugal	532	Angleterre	529	Suède	542
Allemagne	528	Danemark	528	Portugal	541
Irlande	527	Allemagne	528	Allemagne	541
Serbie	516	Italie	524	Italie	541
Australie	516	Portugal	522	Israël	541
Hongrie	515	Slovénie	520	Hongrie	539
Slovénie	513	Irlande du nord	517	République slovaque	535
République Tchèque	511	Irlande	516	Bulgarie	532
Autriche	508	Serbie	516	Nouvelle Zélande	531
Italie	508	Australie	516	Slovénie	530
République slovaque	507	Croatie	516	Autriche	529
Suède	504	Lituanie	515	Lituanie	528
Kazakhstan	501	Belgique (flamande)	509	Australie	527
Moyenne internationale	500	Espagne	505	Pologne	526
Malte	496	Roumanie	505	France	520
Norvège	495	Pologne	505	Espagne	513
Croatie	490	Moyenne internationale	500	Norvège	507
Nouvelle Zélande	486	Nouvelle Zélande	497	Belgique (flamande)	506
Espagne	482	Kazakhstan	495	Roumanie	502
Roumanie	482	Norvège	494	Moyenne internationale	500
Pologne	481	Chili	480	Géorgie	488
Turquie	469	Thaïlande	472	Malte	477
Azerbaïdjan	463	Turquie	463	Trinité-et-Tobago	471
Chili	462	Géorgie	455	Azerbaïdjan	462
Thaïlande	458	Iran	453	Iran	457
Arménie	452	Bahreïn	449	Colombie	448
Géorgie	450	Malte	446	Emirats Arabes Unis	439
Bahreïn	436	Azerbaïdjan	438	Arabie Saoudite	430
Emirats Arabes Unis	434	Arabie Saoudite	429	Indonésie	428
Iran	431	Emirats Arabes Unis	428	Qatar	425
Qatar	413	Arménie	416	Oman	391
Arabie Saoudite	410	Qatar	394	Maroc	310
Oman	385	Oman	377		
Tunisie	359	Koweït	347		
Koweït	342	Tunisie	346		
Maroc	335	Maroc	264		
Yémen	248	Yémen	209		

Source : Rapports sur les résultats internationaux, enquêtes TIMSS et PIRLS 2011

2.3. Le niveau des acquis selon le niveau de compétence

Les enquêtes internationales TIMSS et PIRLS évaluent les acquis des élèves en fonction d'un score qui varie d'un élève à un autre, d'une école à une autre, et même d'un système éducatif à un autre système. En effet, 4 classes de modalités permettent de mieux situer l'individu de référence afin de révéler sa situation au niveau des acquis scolaires. Le premier niveau « niveau inférieur » correspond à un score global de 400 points à 475 points. Le deuxième niveau « niveau moyen » correspond à un score variant entre 475 et 550 points. Ensuite, le niveau élevé dont les scores varient entre 550 à 625 points. Et enfin, le niveau supérieur contient des scores égaux ou supérieurs à 625 points.

Selon les résultats du tableau 18, la plupart des élèves marocains des deux niveaux d'enseignements n'ont pas atteint le niveau inférieur (scores inférieurs à 400 points). Concernant les élèves de la 4^e année primaire, les pourcentages s'élèvent à 74% en mathématiques et 84 % en sciences. Ces pourcentages sont de l'ordre de 64% et 61 % respectivement pour les élèves de la 2^e année collégiale. Les élèves ayant un niveau moyen ne dépassent guère les 12% et quasiment aucun d'eux ne parvient à réaliser un niveau supérieur.

Tableau 18. Répartition des élèves marocains selon les niveaux de compétences

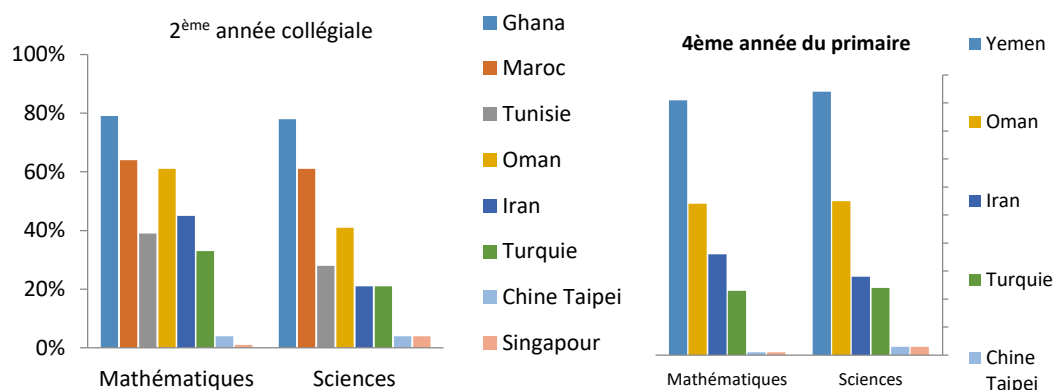
		Supérieur	Elevé	Moyen	Inférieur	N'ayant pas atteint le niveau inférieur
4 ^e année primaire	Mathématiques	0,4%	2,0%	6,9%	16,3%	74,4%
	Sciences	0,1%	1,1%	4,9%	9,7%	84,3%
2 ^e année collégiale	Mathématiques	0,2%	2,2%	9,6%	24,5%	63,5%
	Sciences	0,1%	1,9%	11,4%	25,4%	61,2%

Source : Rapports sur les résultats internationaux en mathématiques et en sciences, enquête TIMSS 2011

Selon le graphique 5, le pourcentage d'élèves n'ayant pas atteint le niveau inférieur de compétence est sensiblement plus élevé au Maroc que dans les autres pays arabes ou africains.

En effet, l'écart des acquis scolaires est très prononcé en comparant notre système éducatif avec le système d'enseignement des pays asiatiques. Les écarts avoisinent les 70 points en pourcentages en 4^e année primaire et environ 55 points en pourcentage en 2^e année collégiale.

Graphique 5 . Proportion d'élèves n'ayant pas atteint le niveau inférieur des compétences



Source : Rapports sur les résultats internationaux en mathématiques et en sciences, enquête TIMSS 2011

Les élèves de la 4^e année ne maîtrisent pas les compétences de base en lecture. En effet, 79% d'entre eux n'ayant pas atteint le niveau inférieur. Les élèves enregistrant des niveaux inférieur et moyen ne représentent que 14% et 6% respectivement. En revanche, seuls 1% d'entre eux atteignent des niveaux élevés et supérieurs d'acquis scolaires (PIRLS ,2011).

2.3.1. Niveau des acquis des élèves marocains selon leurs caractéristiques individuelles, familiales et scolaires

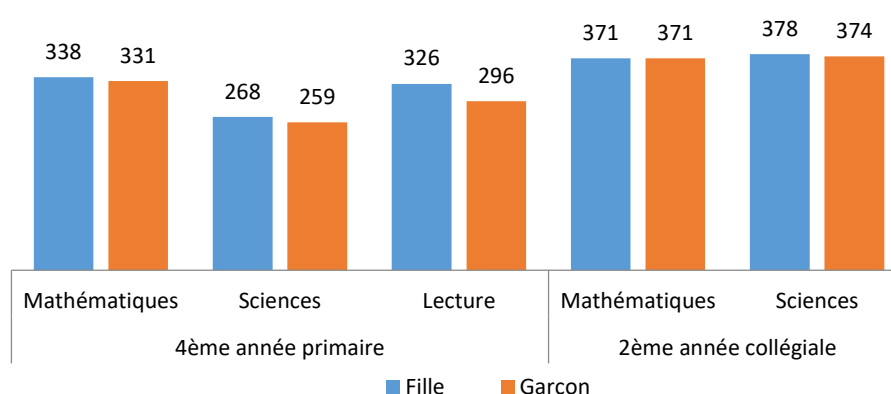
Il s'agit ici d'examiner selon une approche bi variée la relation entre les scores obtenus par les élèves et les variables contextuelles.

2.3.1.1. Niveau des acquis scolaires selon les caractéristiques individuelles et les antécédents scolaires

- **Niveau des acquis scolaire selon le genre**

En 4^e année primaire, les filles ont enregistré des scores plus élevés que les garçons de 7 à 9 points en mathématiques et en sciences respectivement. Ce biais de genre est particulièrement important pour la lecture avec un écart qui s'élève à 30 points en faveur des filles. En 2^e année collégiale, les scores obtenus en mathématiques sont similaires pour les deux groupes, alors qu'en sciences les garçons enregistrent un score plus élevé de 4 points par rapport aux filles.

Graphique 6 . Scores moyens par genre



Source : Rapports sur les résultats internationaux, enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis selon le préscolaire**

Les élèves de la 4^e année du primaire se distinguent entre eux selon qu'ils ont pu ou non bénéficier au préalable d'une éducation du préscolaire. En effet, les scores les plus élevés ont été observés chez les élèves ayant bénéficié de plus d'une année d'éducation en préscolaire comparativement à ceux qui n'ont jamais fréquenté ce niveau d'enseignement ou en ont seulement bénéficié que de quelques mois (Tableau 19).

Tableau 19. Scores moyens des élèves de la quatrième année primaire par nombre d'années passées au préscolaire

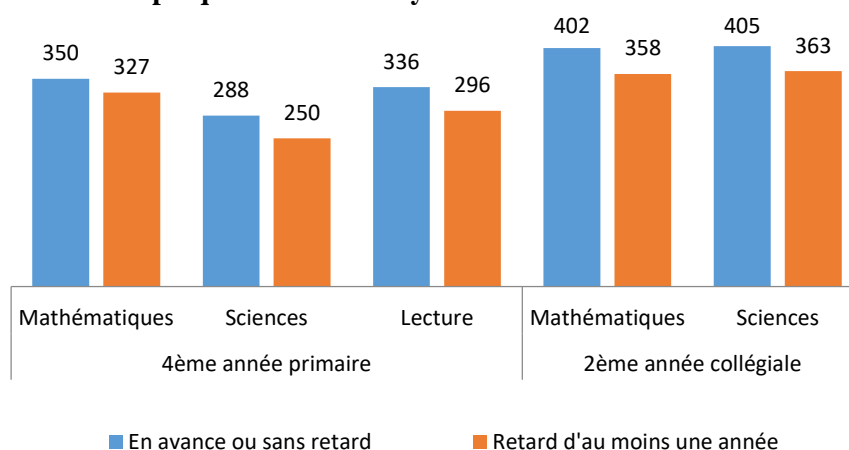
	Mathématiques	Sciences	Lecture
3 ans et plus	351	287	339
Moins de 3 ans et plus d'un an	338	269	324
Un an ou moins	324	249	298
Pas de préscolaire	333	255	293

Source : Rapports sur les résultats internationaux, enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis selon le retard scolaire**

Les chiffres révèlent que les élèves accumulant un retard scolaire d'au moins d'une année réalisent des scores inférieurs à ceux observés chez les élèves ayant l'âge légal de scolarité ou qui sont en avance par rapport à cet âge. Les écarts de performance varient de 23 à 44 points selon la discipline et le niveau d'enseignement (graphique 7). Selon le graphique 7, nous observons que les écarts sont plus importants chez les élèves de la 2^e année collégiale.

Graphique 7 . Scores moyens selon le retard scolaire



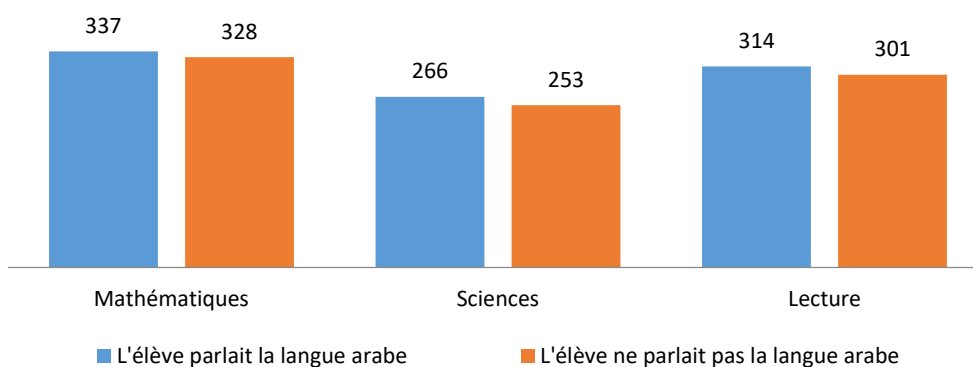
Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

2.3.1.2. Niveau des acquis selon les variables du contexte familial

- Niveau des acquis selon la langue parlée à la maison

Concernant la langue parlée à la maison, les élèves de la 4^e année primaire qui parlaient la langue du test (langue arabe) avant la première rentrée scolaire obtiennent des scores plus élevés que ceux qui ne parlent pas l'arabe. En mathématiques, l'écart est de 9 points, tandis qu'en lecture et en sciences les écarts s'élèvent à 13 points.

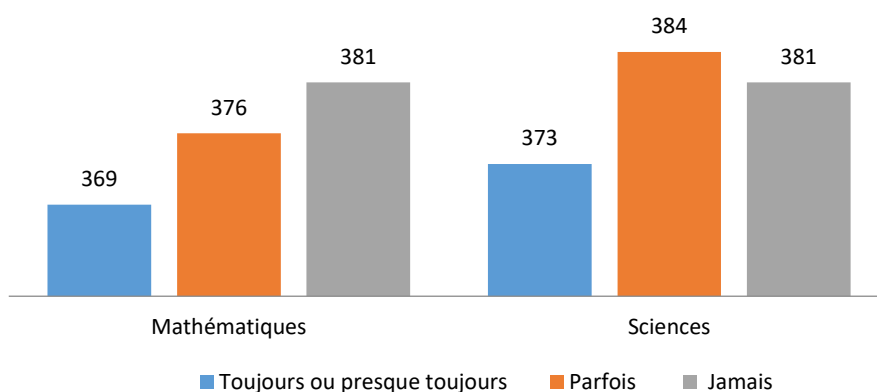
Graphique 8 . Scores moyens en quatrième année primaire selon la langue parlée à la maison



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

En revanche, il apparaît que les élèves de la 2^e année collégiale les plus performants en mathématiques et en sciences sont ceux qui ne pratiquent pas la langue arabe à la maison.

Graphique 9 . Scores moyens en deuxième année collégiale selon la fréquence avec laquelle les élèves parlent la langue arabe à la maison

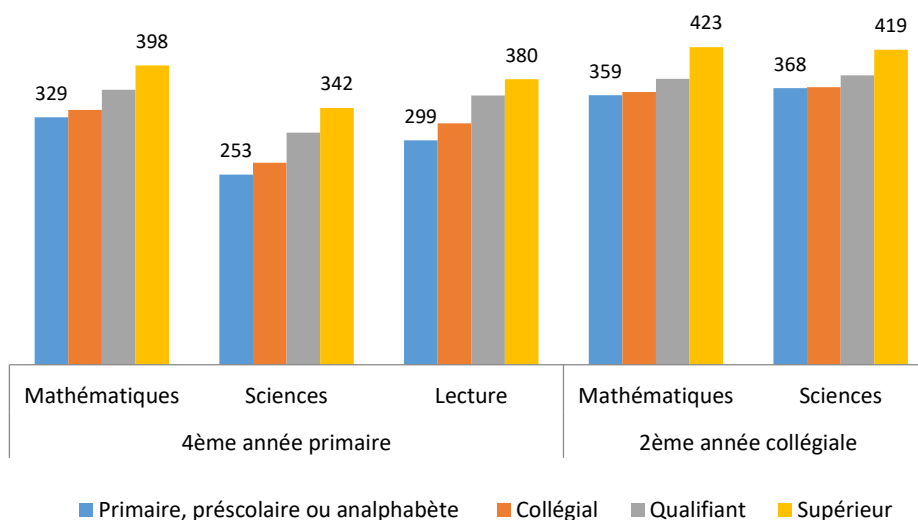


Source : Calculs de l'auteur à partir de l'enquête TIMSS 2011

• **Niveau des acquis selon le niveau socioéconomique des parents ou tuteurs**

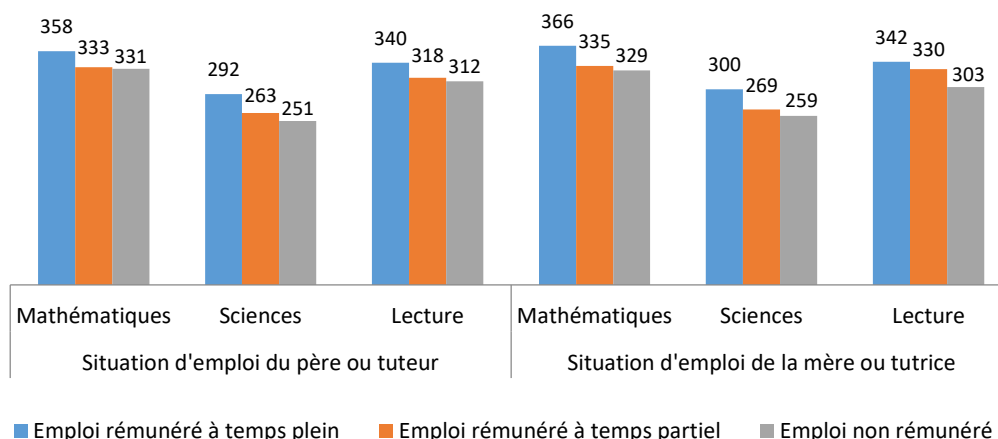
L'analyse des performances scolaires révèle une corrélation assez forte entre les scores des élèves et le niveau d'éducation des parents. En effet, les élèves dont les parents ont obtenu un diplôme de niveau supérieur obtiennent de meilleurs résultats scolaires que les autres groupes d'élèves. Les écarts constatés sont assez élevés variant de 69 points en mathématiques, 89 points en sciences et de 81 points en lecture (Graphique 10).

Graphique 10 . Scores moyens selon le niveau d'éducation des parents



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

Graphique 11 .Scores moyens en quatrième année primaire selon la situation d'emploi des parents ou tuteurs



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

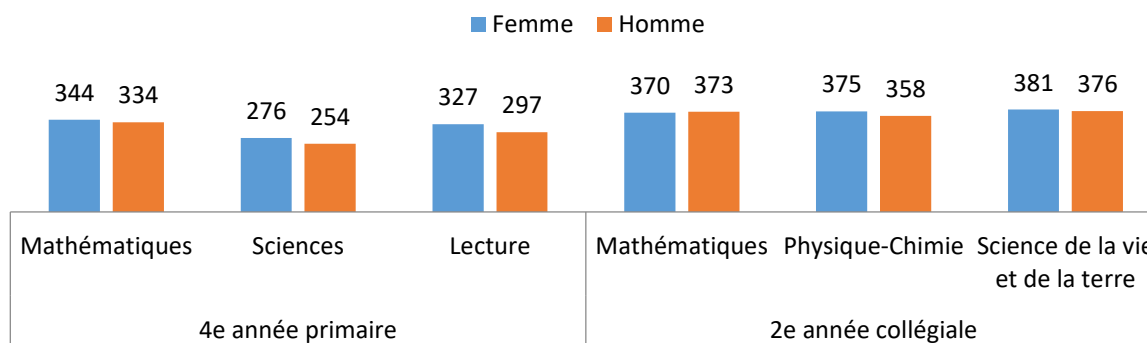
De même, l'analyse des performances scolaires selon les conditions de vie révèle que les élèves issus d'un milieu favorable réussissent mieux que ceux appartenant à des familles défavorisées en termes de ressources familiales, mais également selon la situation d'emploi des parents (Graphique 11).

2.3.1.3. Niveau des acquis selon les caractéristiques des enseignants

- Niveau des acquis selon le genre des enseignants

L'analyse des scores obtenus par les élèves selon le genre des enseignants indique que les élèves affectés à une enseignante réussissent mieux que ceux affecté un enseignant du genre masculin quels que soient le niveau d'éducation de l'enseignant et la matière concernée (Graphique 12).

Graphique 12 . Scores moyens selon le genre des enseignants

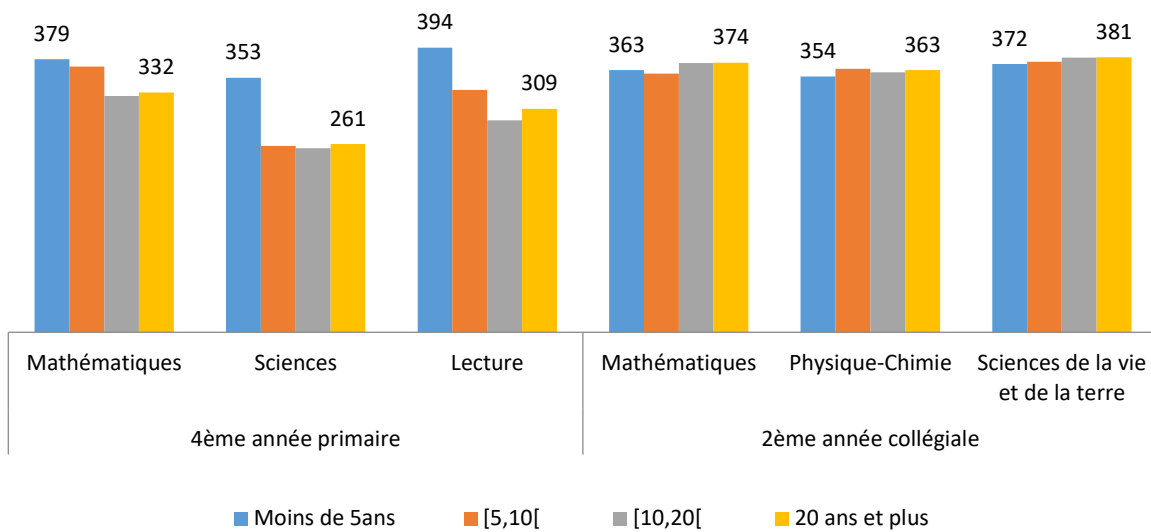


Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis selon le nombre d'années d'expérience des enseignants**

Concernant le nombre d'années d'expérience des enseignants, il apparaît que les élèves de la 4^e année primaire réussissent mieux avec des enseignants ayant moins de cinq ans d'expérience, que les élèves dont les enseignants sont expérimentés (Graphique 13). En effet, les écarts de performances des élèves varient entre 46 et 92 points. En 2^e collégiale, le constat s'inverse en faveur des élèves ayant des enseignants plus expérimentés, et dont les écarts observés s'élèvent à 11 points.

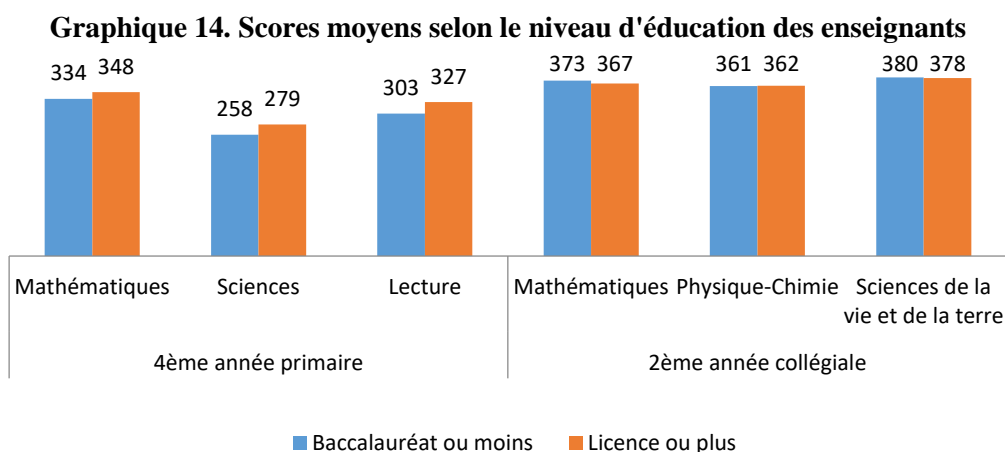
Graphique 13 . Scores moyens selon le nombre d'années d'expérience des enseignants



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis scolaires selon le niveau d'éducation des enseignants**

Les élèves de la 4^e année primaire dont les enseignants de mathématiques ont un diplôme de niveau supérieur obtiennent en moyenne un score supérieur de 14 points par rapport à ceux dont les enseignants ont un baccalauréat. En lecture et en sciences, les écarts sont de l'ordre de 24 points et de 22 points respectivement. En revanche, pour les élèves de la deuxième collégiale, le niveau d'éducation des enseignants (Graphique 14) ne semble pas avoir d'incidence sur leurs écarts de performance.



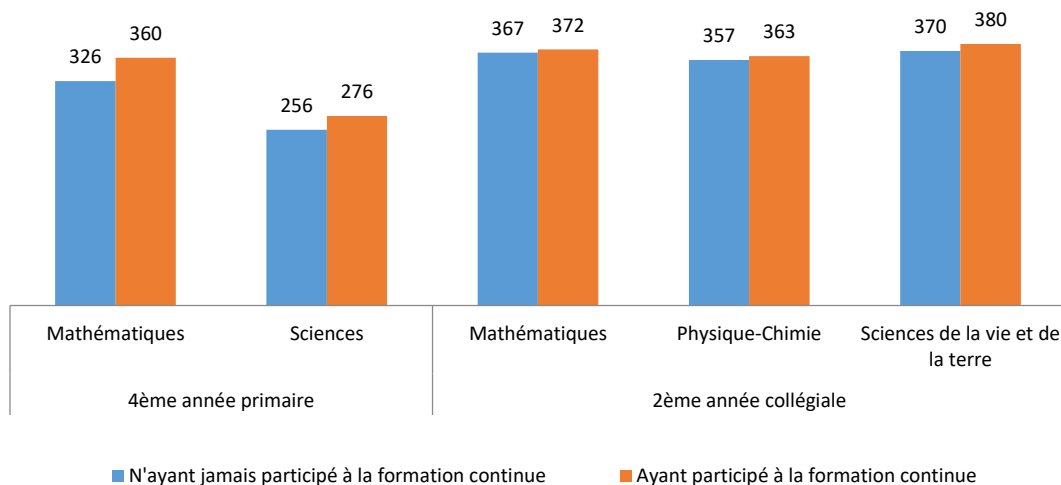
Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

• **Niveau des acquis selon la participation des enseignants aux sessions de formation professionnelle**

En 4^e année primaire, les élèves dont les enseignants déclarent avoir bénéficié d'une session de formation professionnelle ont enregistré des scores élevés variant de 20 à 34 points selon la discipline. L'écart observé en 2^e année collégiale n'est que de 12 points.

Concernant les sessions consacrées aux ateliers traitant les techniques pédagogiques de l'enseignement de la langue, les élèves dont les enseignants ont consacré entre 6 à 15 heures à ces activités ont eu des scores plus élevés de 60 points par rapport à ceux qui n'ont pas participé aux activités. L'écart diminue à 32 points pour les élèves dont les enseignants avaient consacré seulement moins de 6 heures. Le graphique 15 illustre la répartition et les écarts observés entre les deux groupes d'enseignants.

Graphique 15. Scores moyens en mathématiques et en sciences selon la participation des enseignants aux sessions de développement professionnel



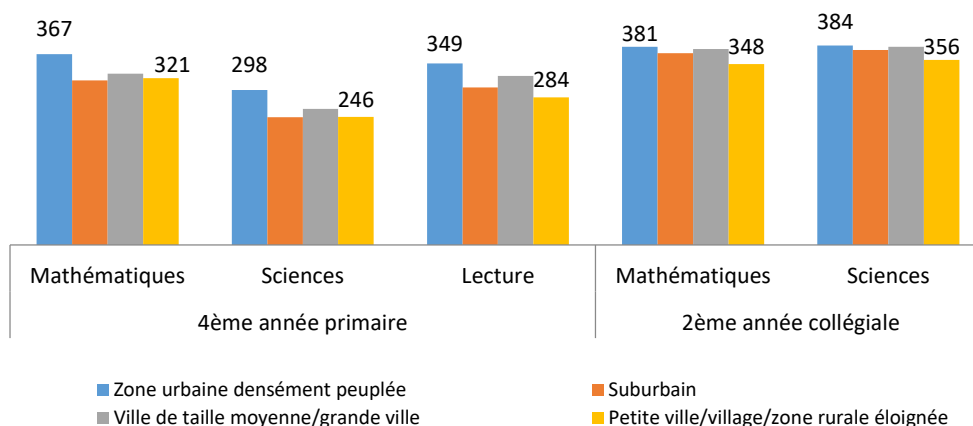
Source : Calculs de l'auteur à partir de l'enquête TIMSS 2011

2.3.2. Niveau des acquis selon l'environnement de l'établissement

- Niveau des acquis selon le milieu de l'établissement

Au niveau des acquis des élèves selon le milieu de l'établissement scolaire, les élèves issus d'une zone urbaine réussissent mieux que les élèves issus d'une zone rurale. En effet, les écarts varient de 28 à 65 points selon le niveau d'enseignement et la discipline (Graphique 16).

Graphique 16 . Scores moyens selon le milieu de l'établissement

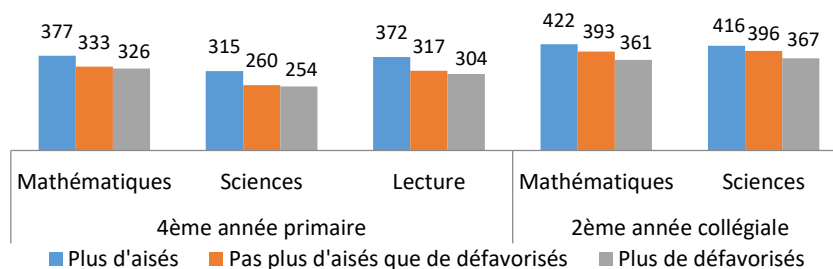


Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- Niveau des acquis selon la composition de l'établissement

Selon la composition sociale et économique de l'établissement, les élèves issus d'une école composée de plus de 75% d'élèves aisés obtiennent des scores très élevés par rapport aux écoles composées de plus de 50 % d'élèves défavorisés. Les écarts de performance sont de l'ordre de 50 points en sciences et en mathématiques et atteignent les 68 points en lecture (Graphique 17).

Graphique 17 . Scores moyens selon la composition socioéconomique de l'établissement



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis selon les ressources de l'établissement**

S'agissant de l'incidence des ressources de l'établissement sur les performances scolaires, les élèves de la 4^e année primaire issus d'une école disposant d'une bibliothèque et d'un laboratoire scientifique ont enregistré des scores plus élevés que les élèves n'ayant pas accès à ces ressources. Les écarts sont s'élèvent à 42 points en mathématiques ,47 points en lecture et 53 points en sciences.

Par ailleurs, le ratio élève /ordinateur révèle des écarts entre les élèves en termes de performance scolaire. Les meilleurs scores ont été enregistrés par les élèves issus d'une école disposant d'un ratio qui ne dépasse pas deux élèves par ordinateur. Par rapport aux élèves n'ayant pas d'ordinateur dans leurs établissements, les écarts varient entre 20 à 45 points.

Tableau 20. Scores moyens selon les ressources de l'établissement

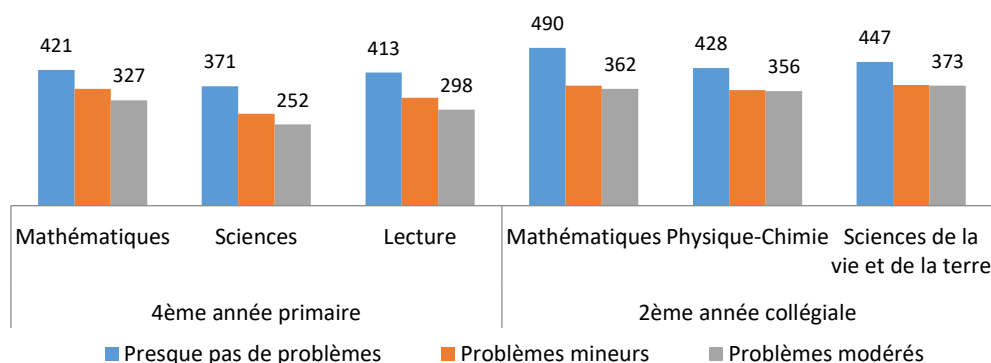
	4 ^{ème} année primaire			2 ^{ème} année collégiale	
	Mathématiques	Sciences	Lecture	Mathématiques	Sciences
Bibliothèque					
Oui	363	299	344	-	-
Non	321	247	297	-	-
Laboratoire de sciences					
Oui	-	324	-	-	377
Non	-	261	-	-	373
Nombre d'élèves par Ordinateur					
Pas d'ordinateurs	323	248	297	364	372
1 ordinateur pour 6 élèves et plus	333	264	315	368	373
1 ordinateur pour 3-5 élèves	345	271	335	394	393
1 ordinateur pour 1-2 élèves	361	293	317	405	404

Source : Calcul de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

- **Niveau des acquis selon les conditions de travail des enseignants**

Les élèves dont les enseignants ont déclaré des conditions de travail pénibles ont obtenu des scores plus faibles que ceux dont les enseignants disposent de conditions de travail favorables. L'absence de problèmes liés à l'encombrement des classes, à la charge horaire, aux ressources pédagogiques et à l'état des bâtiments engendre des scores plus élevés allant de 72 points à 128 points de gains.

Graphique 18 . Scores moyens selon les conditions de travail des enseignants

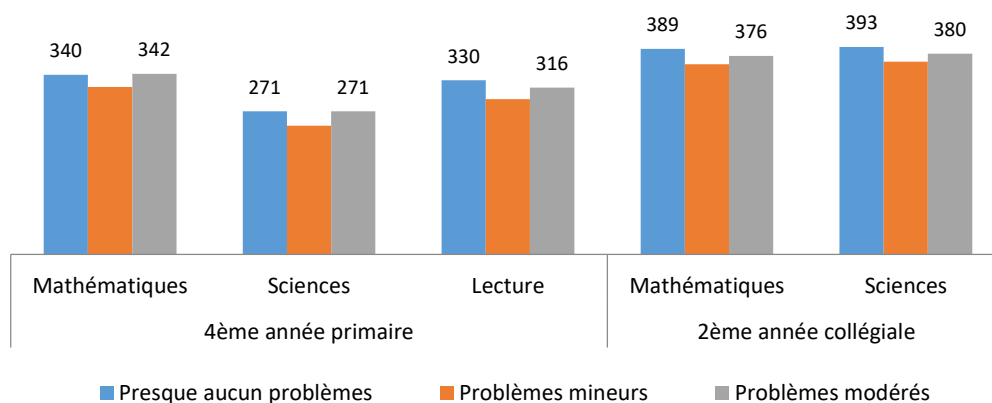


Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

• **Niveau des acquis selon le climat de l'établissement**

En exploitant les informations sur le niveau de discipline et de sécurité au sein de l'école, on constate que les élèves issus des établissements n'ayant presque aucun problème réussissent mieux que ceux fréquentant des écoles où ces problèmes sont présents. Le graphique 19 présente les différences observées selon la discipline, le niveau d'enseignement et selon les problèmes de discipline et de sécurité rencontrés dans les établissements scolaires

Graphique 19 .scores moyens selon les problèmes de discipline et de sécurité rencontrés dans les établissements scolaires



Source : Calculs de l'auteur à partir des enquêtes TIMSS et PIRLS 2011

3. Données exploitées et méthodologie d'estimation

Dans cette section, nous allons chercher à identifier les déterminants des performances scolaires des élèves marocains. Comme le suggère la revue de littérature, il semble évident que les acquis scolaires d'un élève soient influencés par de nombreux facteurs qui sont inter-liés et parfois même inobservables. A cet effet, nous utiliserons les données à caractère hiérarchique pour identifier les facteurs déterminants de la performance scolaire. Aussi, avant de présenter les résultats de nos différentes estimations, il est nécessaire de présenter l'approche méthodologique retenue ainsi que les données utilisées.

3.1. Description des variables utilisées et traitement des données

Dans ce travail, nous exploitons les données issues des enquêtes TIMSS et PIRLS (2011). Les évaluations menées au Maroc en 2011 ont porté sur une cohorte d'élèves dans deux niveaux scolaires et dans différentes matières comme l'indique le tableau 21.

Tableau 21. Description préliminaire de la base de données utilisée

Matières du test	Base de données source	Niveau	Nombre d'écoles	Effectifs des élèves
Lecture	PIRLS 2011	4 ^e année primaire	278	7183
Mathématiques Sciences de la vie et de la terre Biologie Physique Chimie	TIMSS 2011	2 ^e année collégiale	279	8986

Source : TIMSS et PIRLS 2011

Sur la base de l'examen de la revue de littérature, plusieurs variables ont été identifiées comme des déterminants potentiels des acquis scolaires des étudiants marocains. La liste des variables retenues dans notre modélisation est présentée dans les tableaux 22 et 23.

Tableau 22. Liste des variables du premier niveau « Elève »

Variables	Caractéristiques de l'élève	
	Modalités	Définition
Age	L'âge de l'élève	Age de l'élève
Genre	Fille	1, si l'élève est une fille, 0 si non.
	Garçon	
Langue parlée à la maison (arabe)	Toujours	1, si l'élève parle toujours en arabe, 0 si non.
	Souvent	1, si l'élève parle souvent en arabe, 0 si non.
	Jamais	1, si l'élève ne parle jamais l'arabe, 0 si non.
Devoirs scolaires (TIMSS)	Elevé [3h ou plus)	1, si l'élève passe plus de 3h à faire ses devoirs, 0 si non
	Moyen [45min-3h [1, si l'élève passe plus de 45min et moins de 3h à faire ses devoirs, 0 si non.
	Faible (moins de 45min)	1, si l'élève passe moins de 45min à faire ses devoirs, 0 si non.
Devoirs scolaires en lecture (PIRLS)	Plus d'une heure	1, si l'élève passe plus d'1 h à faire ses devoirs, 0 sinon.
	1 heure	1, si l'élève passe 1 heure à faire ses devoirs, 0 sinon.
	Pas de devoirs	1, si l'élève n'a pas de devoirs scolaires, 0 sinon.
Bureau	Dispose ou non d'un bureau	1, si l'élève dispose d'un bureau/ table, 0 sinon.
Ordinateur	Dispose ou non d'un ordinateur	1, si l'élève dispose d'un ordinateur, 0 sinon.
Internet	Dispose ou non d'internet	1, si l'ordinateur est connectée à internet, 0 sinon.
Livres	Possède ou non des livres propres à lui	1, si l'élève possède des livres à la maison, 0 sinon
Education des parents	Achevé le primaire ou niveau collégiale	1, si les parents ont atteint le niveau d'éducation primaire ou collège, 0 sinon
	Achevé le lycée	1, si les parents ont atteint le lycée, 0 sinon
	Niveau universitaire	1, si les parents ont atteint un niveau universitaire, 0 sinon.
Préscolaire (PIRLS)	Études préscolaires	1, si l'élève a étudié au préscolaire, 0 sinon
Passion (PIRLS)	Passion pour la lecture	Echelle (continue)
Confiance (PIRLS)	Confiant en lecture	Echelle (continue)

Source : Auteur selon les données TIMSS et PIRLS 2011

Tableau 23. Liste des variables du deuxième niveau « Ecole »

Variables	Caractéristiques de l'école	
	Modalités	Définition
Composition sociale de l'école	Majorité d'élèves favorisés	1, si l'école contient plus de 50% d'élèves issus d'un milieu favorisé, 0 sinon.
	Composition égalitaire	1, si l'école a une composition sociale égalitaire, 0 sinon.
	Majorité d'élèves défavorisés	1, si l'école contient plus de 50% d'élèves issus d'un milieu défavorisé, 0 sinon.
Milieu	Banlieue	1, si l'école est située dans une banlieue, 0 sinon.
	Urbain	1, si l'école est située dans le milieu urbain, 0 sinon.
	Rural	1, si l'école est située dans le milieu rural, 0 sinon.
Bibliothèque	Si l'école dispose oui ou non d'une bibliothèque	1, si l'école dispose d'une bibliothèque, 0 sinon.
Réussite scolaire		Echelle (continue)

Source : Auteur selon les données TIMSS et PIRLS 2011

Les variables retenues pour l'estimation ont nécessité un traitement préalable notamment l'imputation des données manquantes et le recodage des variables. Afin de pallier aux problèmes des valeurs manquantes, une imputation multiple à l'aide des équations enchaînées « Multiple Imputation by Chained Equations a été retenue. Cette méthode est une technique itérative de la méthode « Markov Chain Monte Carlo » (MCMC). Elle s'inscrit dans l'approche d'imputation de « Fully Conditional Specification » et peut être utilisée comme une approche appropriée pour imputer les larges bases de données nationales et/ou publiques incomplètes (Bouhlila et Sellaouti, 2013). Concernant le nombre requis pour les cycles de l'imputation, Rubin (1987) ; Van Buuren , Boshuizen et Knook (1999) ; Shafer et Olsen (1998) précisent qu'un nombre de $m = 3$ ou $m = 5$ procure d'excellents résultats. Ainsi, dans la présente étude, un cycle d'imputation de $m = 5$ a été retenu a été utilisé afin de corriger toutes les valeurs manquantes.

3.2. Evaluation des acquis scolaires par le recours aux modèles multiniveaux

3.2.1. Justification du modèle multiniveau

Les modèles multiniveaux sont particulièrement bien adaptés à l'analyse de données présentant des structures complexes qui incluent les différents niveaux de découpage du réel. Une des questions fondamentales que posent les sciences sociales est en effet celle des relations entre les individus et leur environnement : les acquisitions des élèves sont-elles influencées par l'école qu'ils fréquentent ?

En sciences sociales, le groupe d'appartenance des individus est souvent considéré comme un contexte très pertinent à étudier dans la mesure où s'y exerce tout un jeu d'influences qui n'est pas sans effets sur les comportements individuels (Bressoux, 2007).

Il est important de distinguer deux catégories de variables parmi celles qui identifient le groupe en tant qu'unité. Certaines variables sont des regroupements de caractéristiques individuelles : ce sont les variables agrégées. Il s'agit bien souvent de moyennes ou d'autres valeurs de tendance centrale (niveau médian par exemple) et des mesures de dispersion (l'écart-type devenant alors une mesure de l'hétérogénéité du groupe). D'autres variables ne proviennent pas de l'agrégation de données individuelles et relèvent strictement du niveau du groupe.

Les travaux sur l'analyse de contexte sont issus d'un courant de réflexion méthodologique, développé surtout dans les années 1960-70, qui a cherché comment analyser les relations entre les individus et le contexte dans lequel ils s'insèrent : en particulier, comment mettre en évidence empiriquement l'influence qu'un groupe exerce sur un individu, au-delà de ses propres caractéristiques personnelles. Parce qu'ils visaient à séparer les effets individuels des effets de groupe, les chercheurs se sont affrontés au problème de l'analyse de données qui incluent plusieurs niveaux : celui des individus et celui des groupes.

La plupart des données impliquant une problématique liée aux interactions entre les individus et leur environnement peuvent être considérées comme des sondages à plusieurs degrés. En effet, on choisit d'abord des macro-unités, puis des micro-unités au sein des macro-unités. Dans ce cas, à la différence d'un sondage aléatoire simple, les observations ne sont pas indépendantes les unes des autres. Le fait que les erreurs ne soient indépendantes génère des corrélations intra-groupes qui ne sont pas sans poser de sérieux problèmes aux techniques classiques de régression par les moindres carrés ordinaires (Bressoux, 2010).

Dans le cas de la régression par la méthode des moindres carrés ordinaires (MCO), cette dépendance est vue comme une nuisance. Or, elle doit souvent au contraire être considérée comme un phénomène intéressant à étudier, tel est précisément le cas dès lors que l'on s'intéresse aux effets de l'environnement sur les individus. Par ailleurs, la violation de l'hypothèse d'indépendance des erreurs a pour conséquence un manque de précision dans les estimations du modèle qui affecte notamment les erreurs-types des coefficients (leur variance est alors sous-estimée).

Une estimation par la méthode des moindres carrés ordinaires (MCO) repose aussi sur l'hypothèse d'homoscédasticité des erreurs : celles-ci doivent avoir une variance constante. Or, on est bien souvent incapable de dire a priori si cette hypothèse est réaliste ou non. Aussi faudrait-il pouvoir la tester et traiter l'hétéroscédasticité dans les cas où elle apparaît car, là encore, la violation de cette hypothèse conduit à sous-estimer les écarts-types des coefficients de régression.

Le modèle multiniveau remplace l'hypothèse d'homoscédasticité par une hypothèse plus faible selon laquelle la variance des résidus peut varier comme une fonction linéaire ou non linéaire des variables explicatives. Cela revient à étudier une possible hétérogénéité des relations entre les différents groupes.

En adoptant des hypothèses plus souples, le modèle multiniveau permet d'une part d'obtenir des estimations non biaisées (les erreurs-types des coefficients sont correctement estimées) et, d'autre part, il permet de tester l'homoscédasticité et l'indépendance des résidus. En effet, si l'hypothèse d'indépendance des erreurs est rejetée, cela signifie que les individus diffèrent effectivement d'un groupe à l'autre et induit donc la question du rôle du groupe lui-même dans la création de ces différences. En revanche, si l'hypothèse d'homoscédasticité est rejetée, cela signifie qu'il existe un effet d'interaction entre la variable explicative considérée et le groupe, ce qui induit de rechercher les causes de cette interaction.

Le modèle multiniveau est donc un modèle particulièrement souple, ce qui le rend très performant dans l'analyse des données ayant des structures complexes.

3.2.2. Présentation du modèle multiniveau

Les problèmes soulevés par l'application des MCO aux données à structure hiérarchique, n'ont connu de solutions satisfaisantes qu'au courant des années quatre-vingts avec l'arrivée des modèles multiniveaux ou hiérarchiques ou encore modèles à coefficients aléatoires ((Bressoux, 2007). Les auteurs précurseurs de ces modèles sont principalement AitKin et Longford (1986) ; Goldstein (1986) ; Mason, Wong et Entwisle (1983) et Raudenbush et Bryk (1988). Afin de simplifier la présentation, le modèle multiniveau peut être formalisé en distinguant dans un premier temps, les écritures en fonction des différents niveaux impliqués dans l'analyse.

Au niveau 1

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (1)$$

Où les indices i et j renvoient aux individus i (les micro-unités) et aux environnements j (les macro-unités).

Les coefficients de régression β_{0j} et β_{1j} pouvant varier d'un environnement j à l'autre.

Au niveau 2, on montre que les coefficients β_{0j} et β_{1j} sont aléatoires en introduisant les termes d'erreurs aléatoires u_{0j} et u_{1j}

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2)$$

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

En remplaçant les valeurs des coefficients β_{0j} et β_{1j} ci-dessus dans l'équation (1), on obtient l'équation :

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} + e_{ij} \quad (3)$$

Où

γ_{00} représente la moyenne générale de Y ;

γ_{10} est la pente de régression moyenne pour tous les groupes ;

u_{0j} représente une erreur aléatoire associée à chaque groupe j supposée normalement distribuée, de moyenne nulle et de variance σ_{u0}^2 .

u_{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 .

Du fait qu'il y a maintenant deux termes d'erreurs au niveau 2, on peut estimer un paramètre supplémentaire : la covariance entre les constantes et les pentes notée σ_{u01} .

L'écriture suivante montre de manière condensée la conception de la structure des paramètres aléatoires à estimer au niveau 2 soit :

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u10} & \sigma_{u1}^2 \end{pmatrix} \right] \quad (4)$$

Au niveau 1, $e_{ij} \sim N(0, \sigma_e^2)$

Avant de procéder à l'estimation du modèle multiniveau, il est toujours intéressant de commencer avec un modèle « vide ».

Le modèle « vide » est le plus simple des modèles multiniveaux possibles. Il est équivalent à une analyse de la variance avec effets aléatoires. Il ne comporte aucune variable explicative et permet simplement d'estimer comment la variance du phénomène se répartit sur les différents niveaux qui sont censés structurer les données. Le modèle « vide » se présente comme suit :

$$\begin{array}{ll} \text{Au niveau 1} & Y_{ij} = \beta_{0j} + e_{ij} \\ \text{Au niveau 2} & \beta_{0j} = \gamma_{00} + u_{0j} \end{array} \quad (5)$$

Soit en intégrant les deux équations dans une seule équation :

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \quad (6)$$

Où

- γ_{00} Représente la moyenne générale de Y ;
- u_{0j} représente une erreur aléatoire associée à chaque groupe j supposée normalement distribuée, de moyenne nulle et de variance σ_{u0}^2 .
- e_{ij} représente une erreur aléatoire associée à chaque individu i supposée normalement distribuée, de moyenne nulle et de variance σ_e^2 .

Le modèle se compose d'une partie fixe (ou structurelle) où γ_{00} représente la moyenne générale de Y et d'une partie aléatoire (ou stochastique) composée des deux termes d'erreurs u_{0j} et e_{ij} . A ce stade, il est particulièrement intéressant de rechercher ce que représente la part de variance intergroupes par rapport à la variance totale. On calcule alors le coefficient de corrélation intra classe :

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (7)$$

A travers le calcul du coefficient, le modèle vide permet de délivrer une information cruciale sur la répartition de la variance entre les différents niveaux envisagés dans l'analyse. Il permet également d'y rapporter, implicitement ou explicitement, l'évolution de la part de variance résiduelle au fil des spécifications successives caractérisées par la prise en compte des facteurs individuels et contextuels. Autrement dit, on estime trois modèles : un modèle vide, un modèle du premier niveau qui ne contient que les variables intrinsèques à l'élève et enfin le modèle complet qui contient à la fois les variables du premier niveau et du deuxième niveau « Ecole ».

Bien que le modèle multiniveau offre de multiples avantages par rapport au modèle classique des MCO, il repose sur certaines hypothèses dont la remise en cause soumet le modèle au risque d'endogénéité qui nécessite alors un traitement particulier. Ce problème provient d'une corrélation entre les résidus estimés du deuxième niveau et les variables explicatives du premier niveau. La violation de l'hypothèse d'exogénéité faible engendre forcément des paramètres et des écarts types biaisés. En effet, la violation de cette hypothèse est due à la corrélation existante entre les résidus estimés du deuxième niveau et les variables observables du modèle du premier niveau. Par conséquent, les modèles économétriques avec correction des problèmes d'endogénéité représentent plusieurs avantages par rapport aux autres méthodes économétriques qui mettent en avant une hypothèse d'exogénéité. Les modèles qui proposent une correction des biais d'endogénéité présentent plusieurs avantages, à savoir :

- Remédier au problème d'échantillonnage en tenant compte du fait que les élèves peuvent ne pas être aléatoirement distribués entre les écoles et d'identifier plusieurs coefficients associés à chaque établissement de l'échantillon.
- Une décomposition de la variance en deux types de variance : une variance inter-école et une variance intra-école.
- Examiner la relation entre les variables du premier et du deuxième niveau.

En nous appuyant sur la méthode des variables instrumentales développée par Hausman et Taylor (1981), nous montrons comment l'adapter dans le cadre du modèle multiniveau afin de corriger à la fois les paramètres estimés et la matrice de variance-covariance de notre modèle complet²².

- **La résolution du biais d'endogénéité à l'aide de la technique des variables instrumentales**

Parmi les hypothèses majeures, sur lesquelles reposent les modèles multiniveaux, il y a l'hypothèse d'absence de corrélation entre les variables explicatives de chaque niveau et les termes aléatoires de l'autre niveau (Raudenbush et Bryk, 2002). Or, dans la pratique, cette hypothèse est rarement vérifiée du fait de la présence dans la perturbation d'effets spécifiques corrélés avec les régresseurs. Or, l'absence de biais et la convergence des estimateurs usuels du modèle à erreurs composées reposent précisément sur l'hypothèse d'exogénéité stricte des régresseurs.

Selon Spencer et Fielding (2000), il n'est pas inhabituel pour un modèle multiniveau de souffrir d'un problème d'endogénéité. Reichstein (2001) estime que ce problème peut être causé par trois facteurs, à savoir : l'existence des variables omises, des erreurs de mesure et enfin la simultanéité appelée aussi la causalité inverse (X cause Y et Y cause X). Dans le cadre de notre étude, il s'agit de résoudre ce troisième et dernier facteur de simultanéité illustré par Shepherd (2010), à travers la simple représentation suivante :

$$Y = X\beta + E \tag{8}$$

$$X = Y\mu + W$$

Avec E est la partie résiduelle du modèle principale

²²Modèle complet est le modèle qui contient à la fois les variables du premier niveau et du deuxième niveau

Et W est la partie résiduelle du modèle emboité

Etant donné que les variables X et Y sont interdépendants, on aura donc :

$$\mathbf{X} = (\mathbf{X}\boldsymbol{\beta} + \mathbf{E})\boldsymbol{\mu} + \mathbf{W} \quad (9)$$

En outre, l'une des hypothèses les plus importantes de la régression avec les moindres carrés ordinaires est l'exogénéité des variables ($Cov(x, \varepsilon) = 0$), c'est-à-dire que les erreurs ne doivent pas être corrélées avec les variables dépendantes. Par conséquent, comme le montre l'exemple précédent X et E sont corrélés et violent ainsi les hypothèses d'indépendance entre la variable explicative et le terme d'erreur. La violation de ces hypothèses nécessite alors un traitement particulier.

Pour résoudre ce biais, nous allons recourir à la méthode des variables instrumentales (IV) de Hausman-Taylor (1981) dans un cadre de la modélisation multiniveau.

Bien que l'approche de Hausman et Taylor ait été initialement conçue pour le traitement de l'endogénéité dans le cadre des données de panel, cette méthode peut également être utilisée, en réalisant quelques ajustements, dans le cas des données ayant une structure hiérarchique et dans le cas des modèles multiniveaux.

Encadré 1 : Etape d'estimation d'un modèle multiniveau avec l'approche de Hausman et Taylor(1981)

Pour ce faire, on estime le modèle complet suivant :

$$Y_{ij} = c + \beta_1 X_{ij}^1 + \beta_2 X_{ij}^2 + \gamma_1 Z_j^1 + \gamma_2 Z_j^2 + V_j + e_{ij}^{23} \quad (10)$$

L'algorithme d'estimation d'un modèle à effets aléatoires corrélé avec IV (variables instrumentales).

1. $Y_{ij} - Y_j = (X_{ij}^1 - X_j^1) \beta_1 + (X_{ij}^2 - X_j^2) \beta_2 + (e_{ij} - e_j)$
2. On calcule $\bar{e}_j = \frac{\sum_{i=1}^N e_{ij}}{N}$
3. On regresse $Z_j^2 = c + a_1 Z_j^1 + a_2 X_{ij}^1 + V_j$
4. On extrait \hat{Z}_j^2 à partir de l'étape 3
5. On construit alors $Z = (Z^1, \hat{Z}_j^2)$
6. On regresse $\bar{e}_j = c + b_1 Z_j^1 + b_2 \hat{Z}_j^2 + V_j$
7. Sachant que $\sigma^{*2} = \sigma_u^2 + \frac{\sigma_\varepsilon^2}{K}$
 - i. σ^{*2} = variance totale
 - ii. σ_u^2 = variance inter – individuelle
 - iii. σ_ε^2 = variance intra – individuelle

Avec la variance totale à extraire de l'étape 6, c'est-à-dire de la variance estimée de V_j
Et la variance intra individuelle devient

$$\text{Donc } \sigma_u^2 = \sigma^{*2} - \frac{\sigma_\varepsilon^2}{K}$$

8. Estimation du modèle à effets aléatoires corrélés avec VI

a) On calcule $\hat{\theta}$

$$\hat{\theta} = \sqrt{\frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_\varepsilon^2 + K\sigma_u^2}}$$

b) On transforme toutes les variables du modèle

$$W^* = [X_{ij}^1, X_{ij}^2, Z_j^1, Z_j^2] - \hat{\theta} [X_{ij}^1, X_{ij}^2, Z_j^1, Z_j^2]$$

$$y^* = (y_{ij} - \hat{\theta} y_{ij}), (X_{ij}^1 - \hat{\theta} X_{ij}^1, \dots, \dots, Z_j^2 - \hat{\theta} Z_j^2)$$

c) les instruments

$$I = [(X_{ij}^1 - X_j^1)', (X_{ij}^2 - X_j^2)', Z_j^1', X_j^1']$$

d) On regresse w^* sur I pour obtenir \hat{w}^*

e) On regresse y^* sur \hat{w}^* pour obtenir $\hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2$

f) Test d'Hausman pour la validité des instruments (modèle à effets aléatoires vs modèle à HT)

Cette méthode sera ainsi utilisée dans l'estimation de notre modèle économétrique afin de traiter le biais d'endogénéité et par conséquent d'identifier les paramètres sans biais de notre spécification hiérarchique.

²³ La spécification d'un modèle à la Hausman et Taylor exige de distinguer entre les variables du premier niveau « X » exogènes et celles endogènes, et aussi les variables du deuxième niveau « Z » exogènes et celles endogènes.

4. Les résultats de la modélisation

4.1. Le rôle des caractéristiques individuelles et familiales de l'élève dans l'explication des acquis scolaires

Les résultats statistiques du modèle individuel sur les acquis scolaires des élèves (Tableau 24) montrent que les élèves les plus âgés progressent significativement moins bien que les élèves jeunes et ce dans l'ensemble des matières. En effet, la variable qui caractérise l'âge de l'élève en quatrième année du primaire en lecture est significative au seuil de (5%) alors qu'au niveau de la deuxième année collégiale à cette variable est significative au seuil de (1%) pour l'ensemble des matières scientifiques (Tableaux 25 et 26). Ainsi, plus on est jeune plus on a tendance à mieux assimiler de nouvelles compétences.

La variable genre apparaît significative et indique l'existence de disparités dans les performances entre les filles et les garçons. En lecture, les filles réussissent mieux que les garçons avec un seuil de significativité de l'ordre de 1%. Cependant, l'inverse se produit dans les matières scientifiques, dans la mesure où les garçons réalisent de meilleures performances que les filles. Ainsi, la variable genre est statistiquement significative au seuil de (1%) pour les mathématiques, les sciences de la vie et de la terre et pour la physique et au seuil de 5% pour la biologie alors qu'elle est non significative pour la chimie ce qui indique que les différences de scores selon le genre sont négligeables dans cette matière. Ce résultat peut être expliqué en se référant à l'étude de Ian, Armstrong et Rounds (2009) qui postule que les hommes préfèrent travailler sur des objets concrets et tendent donc à s'orienter vers des domaines quantitatifs, alors que les femmes préfèrent les relations humaines et développent de ce fait des compétences relatives aux activités artistiques et sociales. De manière générale, les résultats obtenus dans l'analyse de l'impact de l'âge et du genre sur les acquis scolaires des élèves marocains corroborent les conclusions de plusieurs études existantes dont celles de Hanchane, Benbiga, et Idiri (2012) ; Linn, Hyde, et Else-Quest (2010) ; Nosek et al. (2009) ; Fuchs et WoBmann(2004) ; Murphy(2000) ; Spencer et Steele(1999) ; Hopp, Frost, Ryan, Fennema, et Hyde (1990) ; Sherman et Fennema(1977).

**Tableau 24. Résultats du modèle multiniveaux en lecture avec résolution du biais d'endogénéité
4^e année du primaire**

Variables	Lecture		
	Coef.	P> t	
Age	-12.18	0,000	
Genre	20.67	0,000	
Préscolaire	4.58	0,061	
Ordinateur	9.36	0,000	
Bureau	4.73	0,005	
Internet	-8.87	0,000	
Livres	6.47	0,000	
Parle l'arabe à la maison			
	Souvent	11,21	0,000
	Toujours	-1.21	0,579
Niveau d'instruction des parents			
	Primaire/Niveau collégial	-6,78	0,011
	Lycée	-6.78	0,000
Devoirs scolaires			
	Une heure	24,56	0,000
	Plus d'une heure	26,37	0,000
Passion pour la lecture	5.85	0,000	
Confiance de l'élève envers la lecture	9.88	0,000	
Milieu de l'école			
	Urbain	8.40	0,054
	Banlieue	6.72	0,184
Composition sociale de l'école			
	Ecole à composition égalitaire	5.48	0,166
	Majorité d'étudiants favorisés	6,71	0,091
Effectif	0.001	0,708	
Accent de l'école sur le succès académique	2.89	0,001	
Jours d'instruction/semaine	-20.49	0,000	
Bibliothèque	5.78	0,134	
Effets fixes			
Constante	348,77	0,000	
Niveau 2 : variance inter-écoles		1996,55	
Niveau 1 : variance intra-écoles		2594,015	
Effets aléatoire			
Pourcentage de la variance inter-écoles (Niveau 2)		43,49	
Pourcentage de la variance inter-écoles (Niveau 1)		56,50	
Test d'Hausman		0,000	
Nombre d'élève		7183	
Nombre d'école		278	

Source : D'après l'auteur à partir des données PIRLS 2011

L'analyse de l'impact du préscolaire sur les performances des élèves révèle que les enfants ayant bénéficié d'une éducation préscolaire ont de meilleurs résultats en lecture que ceux qui n'ont pas reçu une telle éducation. Cette variable est statistiquement significative au seuil de (5%). Ce résultat peut s'expliquer par le fait que le programme préscolaire au Maroc intègre l'enseignement de la lecture et de l'écriture ce qui permet aux enfants ayant bénéficié de cet

enseignement d'acquérir l'apprentissage de l'alphabet à un âge précoce avant d'intégrer l'école primaire. Cela représente des avantages sans doute par rapport aux enfants qui n'ont aucune connaissance des lettres, et/ou n'ont reçu aucun enseignement préscolaire. L'effet positif de l'accès à l'éducation préscolaire est également confirmé par plusieurs études : Sandoval-Hernandez, Taniguchi, et Aghakasir (2013) ; Hanchane, Benbiga et Idir (2012) ; Schütz (2009) ; Berlinski, Galiani, et Manacorda (2008) ; Sammons, et al. (2004) ; Burkam et Lee (2002) ; Caille (2001) ; Fusaro (1997) ; Hopkins (1995).

Tableau 25. Résultats du modèle multiniveaux en matières scientifiques avec résolution du biais d'endogénéité- 2^e année collégiale

Variables	Mathématiques		S.V.T		Biologie	
	Coef.	P> t	Coef.	P> t	Coef.	P> t
Age	-15.75	0,000	-16.78	0,000	-15.99	0,000
Genre	-12.76	0,000	-14.93	0,000	-3.62	0,039
Ordinateur	8.61	0,000	4.75	0,065	6.32	0,012
Bureau	6.23	0,000	6.35	0,001	10.50	0,000
Internet	10.94	0,000	3.46	0,181	5.12	0,044
Livres	11.40	0,000	10.64	0,000	12.38	0,000
Parle l'arabe à la maison						
Souvent	0.56	0,847	2.23	0,516	4,85	0,152
Toujours	-5.63	0,047	-9.33	0,005	-0.05	0,987
Niveau d'éducation des parents						
Primaire/ collégiale	-15.89	0,000	-14.30	0,000	-11.57	0,000
Lycée	-16.01	0,000	-12.21	0,000	-12.18	0,000
Devoirs scolaires						
Plus de 3heures	10.31	0,000	-26.98	0,000	-16.11	0,000
[45 min-3h [12.93	0,000	-6.77	0,001	-2,12	0,291
Milieu de l'école						
Urbain	14.27	0,000	13.45	0,002	15.61	0,001
Banlieue	15.07	0,012	16.86	0,009	18.07	0,007
Composition sociale de l'école						
Composition égalitaire	14.56	0,000	7.08	0,087	8,23	0,043
Majorité favorisées	14.16	0,003	10.81	0,036	11.80	0,022
Effectif	-0.05	0,000	-0.04	0,001	-0.03	0,003
Accent de l'école sur le succès académique	4.16	0,000	4.04	0,000	3,63	0,000
Jours d'instruction/semaine	-19.74	0,000	-13.69	0,002	-12.98	0,005
Laboratoire	-	-	1.77	0,713	-0.39	0,938
Effets fixes						
Constante	335,2	0,000	353,46	0,000	332,83	0,000
	2					
Niveau 2 : variance inter-écoles		1732,76		1627,72		1832,76
Niveau 1 : variance intra-écoles		2124,65		2253,87		2436,34
Effets aléatoire						
Pourcentage de la variance inter-écoles (Niveau 2)		48,96		41,76		45,76
Pourcentage de la variance inter-écoles (Niveau 1)		51,04		58,24		54,24
Test d'Hausman		0,000		0,000		0,000
Nombre d'élève		8986		8986		8986
Nombre d'école		279		279		279

Source : D'après l'auteur à partir des données TIMSS 2011

Concernant l'effet des conditions socio-économiques sur les performances scolaires, elles sont représentées par la disponibilité d'un ordinateur, d'internet, d'un bureau et de livres à la maison. Le modèle à effet individuel montre que la disponibilité d'un ordinateur à la maison a un impact positif sur les performances scolaires. Cet effet semble plus prononcé dans les matières suivantes : la chimie, la lecture et les mathématiques. En revanche, cet impact est moindre dans les sciences de la vie et de la terre et la physique probablement parce que l'apprentissage dans ces deux matières nécessite plutôt une pratique dans la vie réelle. En revanche, la possession d'un ordinateur n'a pas d'effet sur la performance en lecture. Cela peut s'expliquer par le fait que l'amélioration des aptitudes en lecture des élèves de la quatrième année suppose plutôt l'initiation à la lecture de livres afin de développer leurs compétences linguistiques. En outre, au niveau de la deuxième année collégiale les élèves ayant un ordinateur connecté à internet peuvent avoir accès aux cours et aux séries d'exercices de mathématiques et de chimie contrairement à leurs camarades qui ne disposent pas de l'accès à ces outils et doivent se contenter du cours dispensé en classe.

La disponibilité d'un bureau ou d'une table pour l'élève a une incidence positive sur les acquis scolaires dans l'ensemble des matières. En effet, la disponibilité de l'une de ces commodités permet à l'élève de pouvoir disposer d'un endroit pour effectuer ses devoirs scolaires et/ou réviser son cours ce qui a pour effet d'accroître ses performances scolaires. Cette variable est statistiquement significative au seuil de 1% avec les performances scolaires des élèves en biologie et au seuil de 5% en mathématiques, sciences de la vie et de la terre et en physique.

La possession de livres propres à l'élève, autres que les manuels scolaires à son domicile a un impact positif et significatif sur les acquis scolaires des élèves dans l'ensemble des matières. La disponibilité des livres à la maison stimule ainsi la curiosité intellectuelle des élèves ce qui contribue à l'amélioration de leurs performances à l'école. En fait, bien que cette variable soit significative pour la lecture, matière pour laquelle la disponibilité d'ouvrage est pourtant considérée comme essentielle, la passion pour la matière et la confiance de l'élève en ses compétences en lecture fournissent des résultats instructifs (Tableau 24).

Tableau 26. Résultats du modèle multiniveau en matières scientifiques avec résolution du biais d'endogénéité- 2^e année collégiale (Suite)

Variables	Physique		Chimie		
	Coef.	P> t	Coef.	P> t	
Age	-20,99	0,000	-16.53	0,000	
Genre	-21.54	0,000	1.34	0,448	
Ordinateur	3,34	0,227	9.48	0,000	
Bureau	7.35	0,000	1.11	0,564	
Internet	1.54	0,579	9.78	0,000	
Livres	13.33	0,000	-2.72	0,007	
Parle l'arabe à la maison	Souvent	8.69	0,019	-3.85	0,260
	Toujours	-0,50	0,888	-10.15	0,002
Niveau d'éducation des parents	Primaire/ collégiale	-9.77	0,000	-21.30	0,000
	Lycée	-10.01	0,002	-15.68	0,000
Devoirs scolaires	Plus de 3heures	-24,53	0,000	-14.63	0,000
	[45min-3h [0.48	0,820	0,002	0,999
Milieu de l'école	Urbain	13.39	0,006	18.97	0,000
	Banlieue	17.44	0,016	20.30	0,002
Composition sociale de l'école	Composition égalitaire	8.02	0,087	13.74	0,001
	Majorité favorisées	8.78	0,109	8.90	0,076
Effectif		-,043	0,002	-,04	0,000
Insistance du directeur sur le succès académique		4.24	0,000	4,92	0,000
Jours d'instruction/semaine		-13.81	0,004	-13.30	0,003
Laboratoire		6.89	0,190	.023	0,996
Effets fixes					
Constante		334,67	0,000	365,43	0,000
Niveau 2 : variance inter-écoles		1452,87		1552,31	
Niveau 1 : variance intra-écoles		2365,85		2486,66	
Effets aléatoire					
Pourcentage de la variance inter-écoles (Niveau 2)		40,87		45,67	
Pourcentage de la variance inter-écoles (Niveau 1)		59,13		54,33	
Test d'Hausman		0,000		0,000	
Nombre d'élève		8986		8986	
Nombre d'école		279		279	

Source : D'après l'auteur à partir des données TIMSS 2011

En effet, avoir la passion de la lecture et être confiant dans ses compétences en lecture semblent être deux conditions suffisantes pour avoir de meilleures performances dans cette discipline. Ces deux variables sont positivement liées aux acquis scolaires et sont statistiquement significatives au seuil de (1%). En effet, la passion pour la lecture conduit l'élève à lire de plus en plus ce qui a pour effet d'améliorer ses compétences linguistiques et le rend donc plus

confiant. Ces résultats confortent ceux obtenus par Emler (2002), Valentine, Dubois et Cooper (2004) et Le Bastard Landrier (2005).

Par ailleurs, l'effet du plus haut niveau d'éducation atteint par les parents sur la réussite scolaire des élèves montre que les élèves dont les parents ont atteint le niveau d'enseignement primaire, ou secondaire (collège et lycée) réussissent moins bien que leurs camarades dont les parents ont un diplôme universitaire et ce dans toutes les matières. En effet, les parents qui ont un niveau supérieur d'éducation peuvent aider leurs enfants dans leurs devoirs scolaires et tendent à leur procurer un environnement favorable pour l'apprentissage. Les résultats pour cette variable sont statistiquement significatifs et s'inscrivent dans la lignée des études existantes.

Concernant la fréquence de l'utilisation de l'arabe comme langue de communication à la maison, nous observons qu'elle est négativement liée aux performances des élèves qui parlent toujours l'arabe à la maison tout en étant significative seulement pour deux matières (S.V.T et la chimie) (Tableaux 25 et 26). Cependant, ne parler que souvent en arabe à la maison impacte positivement les acquis des élèves en lecture (Tableau 24), alors que l'effet est contrasté pour les matières scientifiques dans la mesure où elle est négativement liée avec les acquis en mathématiques et en chimie et positivement avec les autres les S.V.T, la biologie et la physique (Tableaux 25 et 26).

Ainsi, les résultats obtenus confirment l'idée que les conditions socio-économiques sont des déterminants essentiels dans la réussite scolaire car elles facilitent et favorisent l'apprentissage chez l'enfant. En effet, ces analyses corroborent les résultats obtenus par les recherches menées dans ce sens, plus précisément celles de Hanchane, Idire et Benbiga (2012) ; Chevalier, Harmon, O'Sullivan et Walker (2005) ; Fuchs et Wößmann (2004) ; Baxter (2002) ; Le et Miller (2002) ; Ganzach (2000) ; Spaulding, Wolfe et Haveman (1991) ; Shah et Sewell (1968).

Enfin, le temps consacré aux devoirs scolaires a un impact positif et significatif au seuil de 1% sur les acquis scolaires en lecture et en mathématiques. En effet, les élèves deviennent de plus en plus performants en lecture au fur et à mesure que le temps consacré aux devoirs associés à la lecture augmente comparativement aux élèves dont l'enseignant ne donne pas de devoirs scolaires. A ce niveau de l'analyse les résultats confirment le proverbe « *qu'à force de forger on devient forgeron* ». Cette conclusion est également valable pour les mathématiques. Cependant, les résultats montrent que les élèves qui passent entre 45 min et 3 heures à faire les devoirs scolaires en mathématiques réussissent mieux que ceux qui y passent moins de temps

et davantage que ceux qui y passent plus de trois heures (Tableaux 25 et 26). Ceci s'explique d'une part par le fait que résoudre plus de problèmes et d'exercices mathématiques aide à exceller dans les mathématiques, et d'autre part, les élèves moyennement performants ont besoin de plus de temps et probablement d'une aide extérieure pour résoudre les exercices demandés par l'enseignant.

Cependant, le temps consacré à faire les devoirs est négativement corrélé avec les acquis scolaires dans les autres matières à savoir : la physique, les sciences de la vie et terre, la chimie et la biologie (Tableaux 25 et 26). Cela ne signifie pas que passer du temps dans les devoirs scolaires impacte négativement les résultats scolaires mais plutôt que malgré la résolution d'exercices, les élèves faibles n'arrivent pas à rehausser leur niveau comparativement à leurs camarades. L'association positive entre l'effort individuel de l'élève (devoirs scolaires) et ses acquis scolaires confirme par conséquent les conclusions des études réalisées par Reimers, Keith, Aubey et Pottebaum (1986) ; Postlethwaite et Wiley (1992) ; Muhlenbruck, Cooper, Nye, et Lindsay (1999) ; Normore et Pelletier (2007) ; Hanchane, Benbiga, et Idir (2012).

En résumé, les résultats statistiques du modèle individuel sur les acquis scolaires des élèves de la quatrième année du primaire en lecture et des élèves de la deuxième année collégiale en mathématiques, sciences de la vie et de la terre, biologie, physique et en chimie sont en général statistiquement significatifs. En outre, les caractéristiques individuelles et familiales de l'élève expliquent en partie les différences de scores entre les écoles. Les variables choisies pour caractériser ces différences contribuent à la réduction de la variance des scores d'un établissement à l'autre, d'environ 22% en lecture pour les élèves de la quatrième année du primaire (Tableau 24), de 42% en mathématiques, de 35% en SVT²⁴, de 39% en biologie (Tableau 25), de 34% en physique et de 40% en chimie pour les élèves marocains de la deuxième année collégiale (Tableau 26).

²⁴ Science de la vie et Terre

4.2. Les effets de l'établissement scolaire sur les acquis des élèves

Les résultats du modèle complet après correction du biais d'endogénéité indiquent que le milieu ou l'emplacement de l'école dans un milieu urbain ou dans une banlieue a un impact positif sur les acquis scolaires dans toutes les matières analysées. Cette dernière variable est statistiquement significative au seuil de (1%) pour la biologie, la chimie et les mathématiques et au seuil de (5%) pour la biologie, la lecture, les SVT et au seuil de (10%) pour la physique (Tableaux 25 et 26). Ainsi, les acquis des élèves dans les écoles urbaines sont meilleurs que ceux de leurs pairs des écoles rurales. Ce résultat peut être expliqué par le fait que les écoles situées dans le milieu urbain ont plus de chance d'être mieux dotées en infrastructures et en ressources pour permettre l'apprentissage dans ces matières qui sont plus pratiques que théoriques. De plus, étant donné que l'affectation des élèves aux écoles primaires publiques marocaines est déterminée par la carte scolaire, la proximité des élèves de leur domicile parental ainsi que la disponibilité des moyens de transport au sein du périmètre urbain peuvent également jouer un rôle important dans la facilité de la vie scolaire. Ceci n'est pas toujours le cas pour les élèves résidant dans le milieu rural qui généralement souffrent du problème de l'éloignement des écoles. Nos résultats s'inscrivent dans la même filiation que ceux de Hanchane, Benbiga et Idir (2012) ; Dufur et Parcel (2001) qui montrent que les élèves des écoles situées dans le milieu urbain obtiennent de meilleures performances que leurs pairs dans le milieu rural (Tableaux 24, 25 et 26).

Concernant les effets de pairs représentés par la composition sociale de l'école, l'analyse montre que cette dernière est positivement liée aux acquis scolaires dans l'ensemble des matières confirmant ainsi les apports de littérature. La variable « école à majorité d'élèves issus d'un milieu favorisé » est statistiquement significative au seuil de (5%) pour la biologie, la chimie, les mathématiques et pour les SVT (science de la vie et terre) (Tableaux 25 et 26) et au seuil de (10%) pour la lecture (Tableau 24) et la physique. Par conséquent, les effets de pairs ou de composition, influencent négativement les performances scolaires des élèves marocains. Ainsi, une école dont les élèves sont majoritairement issus d'un milieu favorisé aura une influence positive sur les rendements scolaires des élèves. Nos conclusions confirment les résultats de Hanchane, Benbiga et Idir (2012) ; Hoxby (2000) ; Hanushek et al. (1999,2003) ; Hanushek et al. (2003) et d'autres travaux.

En revanche, l'effectif des élèves inscrits dans le même niveau d'enseignement mesurant ainsi la taille des classes semble avoir un effet négatif sur les acquis des élèves. Cette variable est statistiquement significative au seuil de (1%) en mathématiques, SVT et en chimie et au seuil de (5%) en physique et biologie (Tableaux 25 et 26). En effet, un effectif d'élèves élevé peut causer l'encombrement dans les classes et rendre ainsi plus difficile pour l'enseignant la gestion de la classe. Nos résultats confirment les conclusions de Hanushek (1997, 2003 et 2006) ; Hoxby (2000) ; de Krueger (1999) ; Woessmann et West (2006) ; Lavy et Angrist (1999).

Par ailleurs, l'analyse des ressources scolaires représentées par la disponibilité d'une bibliothèque et d'un laboratoire montre qu'elles contribuent à améliorer les acquis scolaires des élèves.

En revanche, les résultats de l'analyse montrent que l'importance accordée par le directeur de l'établissement à la réussite scolaire des élèves de l'école est un déterminant important de la performance des élèves marocains. Son impact est positif et significatif de l'ordre de 1% pour l'ensemble des matières analysées. L'analyse de cette variable qui renseigne sur cinq éléments principaux dans la vie éducative de l'élève, montre que les élèves fréquentant les écoles où les objectifs pédagogiques sont bien compris par les enseignants, où les programmes scolaires sont correctement appliqués, où les attentes des enseignants en matière de rendement des élèves sont claires, où les parents sont engagés et où les élèves désirent réussir, obtiennent de meilleurs résultats dans l'ensemble des matières analysées. Ainsi, ce dernier résultat nous suggère que l'implication et l'engagement des acteurs concernés par l'éducation est un prédicateur irrévocable dans la performance des acquis des élèves marocains.

Conclusion

Ce premier chapitre a eu pour objectif d'exploiter les bases de données TIMSS et PIRLS (2011) sur l'évaluation des acquis scolaires en lecture pour les élèves de la 4^e année primaire et dans les différentes matières scientifiques pour les élèves de la deuxième année collégiale afin d'identifier les déterminants individuels et contextuels des performances scolaires des élèves. Avant de proposer une modélisation, nous avons examiné de manière descriptive les données issues des enquêtes TIMSS et PIRLS (2011) afin de présenter le niveau des acquis scolaire des élèves marocains en comparaison avec les systèmes éducatifs relatifs à d'autres pays. A l'issue de la revue de la littérature, nous avons identifié les déterminants potentiels des acquis scolaires pour lesquels nous avons également mené une analyse descriptive.

Enfin, afin d'isoler l'influence de chacun de ces facteurs qu'il agisse des caractéristiques individuelles propre à l'élève et à son environnement familial mais également des ressources et caractéristiques de l'établissement scolaire fréquenté, nous avons eu recours à une modélisation multiniveau laquelle se justifie en raison de la structure hiérarchique des données (niveau 1 : élève et niveau 2 : école) et de la pluralité des facteurs qui influencent les acquis scolaires.

Or, le recours à la méthode des variables instrumentales de Hausman et Taylor s'est avéré nécessaire pour corriger le biais d'endogénéité dû à la corrélation existante entre les variables explicatives du niveau 1(élève) et les résidus estimés du niveau 2(école). Aussi , notre apport se situe à ce niveau dans l'adaptation de la méthodologie développée par Hausman et Taylor(1981) permettant traditionnellement de traiter le problème de l'endogénéité dans le cadre de modèles en données de panel dans le cadre du modèle multiniveau. La prise en compte de l'endogénéité nous a permis ainsi d'obtenir une décomposition robuste de la variance en composantes inter et intra école.

L'étude a abouti à des résultats cohérents et significatifs et permet ainsi d'identifier un certain nombre de facteurs ayant un impact déterminant sur les performances scolaires des élèves marocains.

Les résultats obtenus révèlent que les performances des élèves marocains sont influencées aussi bien par les caractéristiques intrinsèques de l'élève, l'environnement familial que par l'école. En outre, ils confirment les principales conclusions de la revue de littérature empirique dans le domaine.

Concernant les caractéristiques de l'élève, les estimations révèlent que les facteurs de performances scolaires des élèves vont au-delà du statut socio-économique des parents, comme le suggère les travaux antérieurs. En effet, il y a lieu d'indiquer que la spécificité genre influence largement le score de l'élève. Alors que les filles sont meilleures en lecture, les garçons le sont plutôt dans les matières scientifiques. De même, l'âge influence largement l'apprentissage dans la mesure où plus on est jeune plus on a tendance à mieux assimiler de nouvelles compétences.

Par ailleurs, l'environnement familial et plus précisément le niveau d'instruction des parents joue un rôle déterminant sur les acquis des enfants car les parents influencent et transmettent en permanence des connaissances à leurs enfants. Les parents ayant un haut niveau d'étude ont tendance à être plus engagés dans l'éducation de leurs enfants et ont d'avantage connaissance du système éducatif. Ainsi plus le niveau d'éducation atteint par les parents est élevé plus les acquis des élèves sont meilleurs.

D'autres facteurs contribuent également à l'explication des performances scolaires des élèves, en particulier les effets de pairs et les caractéristiques de l'établissement scolaire.

En ce qui concerne les effets de pairs ou de composition, ces derniers influencent négativement les performances scolaires des élèves marocains. Une école dont les élèves sont majoritairement issus d'un milieu moins favorisé aura une influence négative sur les rendements scolaires des élèves.

L'analyse des caractéristiques de l'école montre que le nombre de jours d'instruction par semaine et l'effectif scolarisable au sein de l'école influencent négativement les acquis scolaires. En revanche, l'investissement du directeur dans la réussite scolaire au sein de son école semble jouer un rôle déterminant dans l'explication des facteurs de performance. Celle-ci est statistiquement significative au seuil de (1%) pour l'ensemble des matières. En effet, cette variable regroupe un ensemble de renseignements concernant l'élève, les parents, l'enseignant, les curricula et les objectifs pédagogiques. Ainsi, les élèves réalisent de meilleures performances s'ils sont inscrits dans une école caractérisée par un engagement parental et des enseignants qui assimilent au bien les objectifs pédagogiques de l'école et réussissent dans la mise en œuvre des curricula ainsi que par des élèves qui désirent réussir.

En somme, plusieurs facteurs concourent à l'explication des performances scolaires des élèves marocains. Ces facteurs sont aussi bien des éléments se rapportant à l'élève lui-même, à

l'environnement dans lequel il évolue qu'à l'école qu'il fréquente. En outre, l'engagement des parents de l'élève et de l'élève lui-même, et également du directeur dans le système éducatif peut être considéré comme l'élément clé de la performance scolaire.

A la lumière de ce qui a été mentionné dans la revue de la littérature internationale et nationale, notre étude empirique révèle un apport empirique en matière des résultats des déterminants de la performance scolaire au Maroc.

Les variables retenues pour l'analyse des déterminants des performances scolaires ne sont pas exhaustives. D'autres facteurs contribuent à l'explication des acquis scolaires notamment les caractéristiques liées aux enseignants, à la pédagogie d'instruction et bien d'autres et constituent de nouvelles pistes de recherche à explorer.

Chapitre 2 : Taille des classes et acquis scolaires des élèves marocains, une nouvelle approche semi paramétrique des modèles multiniveaux

Introduction

La question de l'effet de la taille des classes sur la performance scolaire des élèves remonte aux années 60 avec les travaux fondateurs de Coleman (1960). A partir de cette période, les travaux en économie de l'éducation ont focalisé leurs réflexions sur l'étude de l'environnement scolaire et son rôle déterminant pour la réussite scolaire des élèves. De ce fait, l'effet de la taille des classes et son degré d'impact a connu une large discussion empirique autour des méthodes d'estimations et des résultats obtenus à partir des évaluations menées essentiellement sur le système éducatif américain.

Au Maroc, alors que les effets de l'inégalité d'accès à l'éducation ont fait l'objet de plusieurs études (Benaabdelaali et al 2017), l'examen de l'impact de la taille des classes sur la performance scolaire des élèves n'a que très peu été exploré à notre connaissance, à travers l'estimation des modèles linéaires (Ibourk ,2012 ; Mourji, 2013).

En exploitant les données TIMSS et PIRLS 2011, il apparaît que plus de 60% des classes au niveau de la 4^e année primaire incluent plus de 31 élèves. Au niveau de la 2^e année collégiale, plus de 70 % des classes ont un effectif supérieur à 31 élèves.

Ce constat est révélateur d'un problème d'encombrement des classes et d'un environnement scolaire difficile auxquels les enseignants et les élèves sont confrontés au quotidien.

Ce chapitre se donne pour objectif d'une part de vérifier si la réduction de la taille de la classe joue un rôle déterminant sur la performance scolaire, et d'autre part, d'identifier la taille optimale de la taille des classes à partir de laquelle les acquis scolaires des élèves se détériorent de façon significative. Pour répondre à cet objectif, nous procéderons à une décomposition par les différentes tailles de classe (petite, moyenne et grande)²⁵, tout en comparant les degrés d'impacts pour les deux niveaux primaires et collégiaux du système éducatif marocain.

L'impact de la taille des classes sur la performance scolaire des élèves a donné lieu à une vaste littérature dont les résultats et conclusions sont très hétérogènes du fait d'estimations empiriques menées sur des systèmes éducatifs très différents. La difficulté centrale réside dans le fait que la répartition des élèves sur les différentes tailles de classe n'est pas aléatoire. En effet, il existe une forte corrélation entre la taille de la classe et les caractéristiques des élèves. Or, Les méthodes économétriques utilisées dans les travaux empiriques menés au cours des années soixante n'étaient pas encore suffisamment développées pour prendre en considération cet aspect. Ainsi, en raison des limites des méthodes d'estimation classiques, les premiers

²⁵ La définition de la taille des classes est présentée dans la partie descriptive des données

travaux ont abouti à des impacts non- significatifs et parfois même contradictoires. Les méta-analyses de Hanushek (1986,2003) montrent un effet non-significatif de la taille des classes, tandis que l'étude de Krueger (2003) révèle des effets opposés. Cependant, les travaux récents ont essayé de mieux estimer cet impact en exploitant les nouvelles méthodes économétriques.

Dans ce contexte, notre chapitre cherche à estimer l'effet associé à différentes tailles des classes pour le cas du Maroc en exploitant les données TIMSS et PIRLS 2011 relatives aux élèves de la 4^e année du primaire et la 2^e année du collège : Au regard de la littérature existante, l'apport de chapitre est d'ordre aussi bien méthodologique qu'empirique. En fait, il s'agit d'estimer le seuil optimal de la taille des classes, à partir duquel l'acquis scolaire se détériore d'une façon significative. Cet effet sera estimé en recourant à une modélisation semi-paramétrique spécifiée sous forme hiérarchique. En fait, il s'agit d'estimer un modèle à deux composantes. La première identifie les paramètres du modèle spécifié, et la deuxième estime la courbe reliant la partie non paramétrique mesurée par la taille de la classe et le niveau des acquis scolaire mesuré par le score des élèves.

Sur ce point, notre étude est la première aboutissant à une estimation des seuils de la taille maximale des classes au Maroc. D'ailleurs rares sont les études qui fournissent de telles estimations permettant de justifier la pertinence d'une politique scolaire de réduction de la taille des effectifs des classes.

Dans cette perspective, ce chapitre s'organise en 4 sections. La section 1 présente une revue de la littérature de l'impact de la taille des classes sur la réussite scolaire en attachant une attention particulière aux récentes méthodes économétriques utilisées pour remédier au caractère endogène de la taille des classes. La section 2 décrit les variables issues des données TIMSS et PIRLS 2011, que nous sélectionnons pour notre modélisation. Dans la section 3, nous justifions le choix de notre modélisation semi-paramétrique sous forme hiérarchique, afin de pouvoir rendre compte de l'effet de la taille des classes. Enfin, la section 4 présente et commente les résultats issus de cette modélisation à partir des données TIMSS et PIRLS 2011.

1. Revue de la littérature

Avant d'entamer la revue de la littérature, il faut rappeler que le problème que pose l'estimation de l'effet de la taille de la classe est son caractère endogène. Ce problème est expliqué en grande partie à l'affectation non aléatoire des élèves dans les classes, avec un degré de corrélation entre ces dernières et les caractéristiques de l'élève. Pour remédier à ce problème, plusieurs méthodes ont été proposées afin d'identifier son biais l'effet de la taille des classes sur les acquis scolaires des élèves. Les premières se réfèrent aux approches économétriques, tandis que les secondes se penchent beaucoup plus sur les études expérimentales et quasi-expérimentales

1.1. Etudes économétriques

Depuis la fin des années 90, les travaux empiriques dans le champ de l'économie de l'éducation se sont attachés à examiner l'impact de l'école sur l'amélioration des performances des élèves. Dans ce domaine, l'impact de la taille des classes apparaît comme l'une des composantes essentielles de l'environnement scolaire favorable à l'amélioration des acquis scolaires. Cependant, d'un point de vue empirique, l'estimation de l'effet de la taille des classes soulève un certain nombre de difficultés qui tiennent au fait que la répartition des élèves sur l'ensemble des tailles de classe n'est en général pas aléatoire. En effet, il existe une forte corrélation entre la taille de la classe et les caractéristiques des élèves. Il s'ensuit que les estimations obtenues peuvent être biaisées.

Cependant, la littérature récente a développé différentes méthodologies afin de contrôler les problèmes d'endogénéité associés à l'affectation non aléatoire des élèves dans des classes de tailles différentes.

De plus, différentes approches méthodologiques comme la méthode des moindres carrés ordinaires (méthode d'estimation) et les modèles linéaires (spécification du modèle) souffrent de plusieurs limites en raison de la structure des données collectées et de la période d'évaluation qui mesure la performance scolaire à un instant donné. Cette dernière limite rend notamment impossible la collecte des scores antérieurs des élèves, qui serait nécessaire pour une évaluation précise de la performance des acquis scolaires. De même, les ressources permettant une évaluation de l'efficacité ne sont observables qu'à la période considérée.

Aussi, la disponibilité et l'exploitation de données longitudinales a été au cours de ces dernières années une des avancées les plus remarquables pour les travaux économétriques menés dans ce domaine.

Hanushek (1979, 2003) a publié des études en méta-analyse de l'impact des ressources scolaires sur la performance des élèves. En général, ces analyses ont dévoilé l'absence de l'impact des ressources scolaires, et notamment l'effet de la taille des classes sur l'amélioration des acquis scolaires. Les conclusions de cette étude ont cependant été profondément remises en question par plusieurs travaux recourant à la même démarche méthodologie dont celles de Greenwald, Hedges et Laine (1996a, 1996b).

Greenwald (1996), Hedges et Laine (1996) remettent en question les conclusions de Hanushek (1979) au motif que la méthode d'intégration des résultats ne permet pas de tenir compte de l'intensité des effets.

Ainsi, ces auteurs révèlent un effet positif entre les ressources scolaires de l'école (surtout la taille des classes et la dépense par élève) et la performance scolaire. Les estimations réalisées s'appuient sur des études en méta-analyse, et les résultats montrent que le niveau de l'éducation, la formation et l'expérience des enseignants jouent un rôle important dans l'amélioration des acquis des élèves.

De même, selon Krueger (1999), l'absence d'effet de l'environnement scolaire issu de l'étude de Hanushek résulterait de l'estimation de plusieurs modèles selon des méthodes diverses des sous échantillons conduisant à un problème de représentativité de l'échantillon étudié. De fait, les effets estimés auraient été conditionnés par une diminution de la taille de l'échantillon, produisant des effets contradictoires.

1.2. Etudes expérimentales

Les études expérimentales renvoient à la nouvelle tendance permettant de mieux cerner les biais des différentes méthodologies et spécifications empiriques. Depuis 1985, la STAR « Student/Teacher Achievement Ratio » est considérée comme l'une des premières expérimentations réalisées dans ce sens. Ce dispositif repose sur une distribution aléatoire de 11600 élèves de la maternelle au CE2 dans des classes sur la base d'échantillons d'élèves de tailles différentes. Cette distribution aléatoire des élèves entre différentes classes assure une identification sans biais des effets de la taille des classes sur la performance scolaire.

L'étude expérimentale a mis en place trois catégories de classes : des classes de petite taille (13 à 17 élèves), des classes de taille moyenne (22 à 26 élèves) et des classes de taille moyenne (22 à 26 élèves) avec un enseignant assistant. Les écoles choisies dans l'expérimentation nécessitaient des tailles importantes pour s'assurer de la présence de l'une des trois catégories des classes.

L'expérimentation a été réalisée auprès de 80 écoles avec 100 classes de chaque catégorie choisie selon la méthodologie du dispositif.

Au sein de chaque école, les élèves et les enseignants ont été distribués d'une façon aléatoire entre les trois groupes des classes. L'expérimentation a focalisé le traitement en grande partie sur les élèves de la grande section de maternelle, et une minorité des élèves de la deuxième et troisième année du cycle primaire.

Le premier traitement de l'expérimentation était d'assigner les élèves aux différents groupes de classes jusqu'au troisième grade, en évaluant la performance scolaire vers la fin de chaque année scolaire. L'objectif de cette expérimentation est d'évaluer la performance scolaire de 11600 élèves avec une durée moyenne de 2 à 3 ans de traitement dans l'un des groupes de classes.

A la fin de l'expérimentation, les résultats de Finn et Achilles (1990, 1999) et Krueger (1999) révèlent des disparités très significatives pour les élèves affectés aux petites classes, alors qu'aucun écart significatif n'apparaît entre les classes de taille moyenne et celles avec un enseignant assistant. En exploitant les données de STAR, Krueger (1999) montre la présence d'une élasticité importante et significative de la taille des classes sur la réussite scolaire des élèves. Cependant, les coefficients estimés par Krueger (1999) s'avèrent trop élevés comparativement aux résultats obtenus par Angrist et Lavy (1999).

En outre, analysant l'impact de la réduction de la taille des classes selon l'ethnicité des élèves Finn, Achilles (1999) et Krueger (1999) ont mis en évidence un effet important de la taille de la classe pour les élèves issus d'une minorité raciale.

Ce dispositif expérimental présente des avantages multiples. Au-delà de l'étude des effets de très court terme, les données STAR permettent de vérifier l'hypothèse d'une réduction de la taille des classes dans le long terme et sa durabilité dans le temps.

Krueger et Whitmore (2001) ont d'ailleurs exploité les données STAR pour évaluer les impacts à long terme de l'expérimentation. Après la fin de l'expérimentation, les élèves concernés par ce programme ont regagné leurs classes normales pour le reste de leur scolarité. Les auteurs montrent qu'après quatre ans de la fin de l'expérimentation, l'impact d'appartenance à une classe de taille réduite demeure important même si les élèves appartiennent théoriquement à des classes de taille de moyenne.

Les auteurs ont également estimé le niveau de performance des élèves aux examens de fin d'études scolaires qui est une étape essentielle à leur acceptation au système d'enseignement supérieur américain. La passation de ce test constitue en général un indice de la qualité du système scolaire, et en particulier de la performance scolaire au lycée.

Les résultats montrent que le pourcentage d'élèves passant ces tests standardisés est significativement élevé parmi les élèves appartenant à des classes de petite taille lors de l'expérimentation. En termes de pourcentages, 43,7% contre seulement 40% pour les deux types de classes traitées au cours de l'expérimentation.

De plus, la correction des biais de sélection révèle des résultats supérieurs et significatifs pour les élèves assignés à des petites classes.

A partir des données « STAR », Pate-Bain et al. (1997) ont confirmé une probabilité de redoublement moins élevée chez les élèves ayant fréquenté les petites classes. Ces études donnent de meilleures estimations par rapport aux études économétriques qui testent les effets des ressources scolaires dans le long terme et qui révèlent globalement des impacts non significatifs. En fait, Ce constat révélateur réside dans l'exploitation des données longitudinales qui permettent un suivi des mêmes élèves sur plusieurs années avant et après l'expérimentation, et qui démontre par conséquent des résultats assez convaincants par rapport à l'étude économétrique basé généralement sur des échantillons représentatifs de la population.

Comme toutes les démarches d'évaluation, les études expérimentales ont été aussi critiquées, d'une part pour des raisons liées aux modalités de déroulement de l'expérimentation, et d'autre part à cause de la conscience des individus participant à l'expérimentation.

L'une des critiques adressées à l'expérimentation STAR porte sur la structure et la démarche du déroulement de l'expérimentation. Goldstein et Blatchford (1998) indiquent des problèmes de distribution et d'allocation des élèves dans les classes de tailles différentes. Hanushek (1999) et Hoxby (2000) critiquent la méthodologie du dispositif expérimental STAR, car elles consistent à assigner des élèves dans la même classe pendant toute la période d'expérimentation, autrement dit, de la première année à la quatrième année d'expérimentation. De fait, quasiment la moitié de l'échantillon est restée dans le groupe auquel elle était assignée pendant la première année. Le biais engendré dans ce type d'expérimentation est lié en grande partie à la réallocation des élèves entre les différentes tailles des classes. Toutefois, la rentrée progressive dans le programme d'expérimentation peut être une source de biais de sélection.

Le problème d'attrition est connu dans cette littérature parmi les phénomènes récurrents, puisque les élèves sont évalués sur une période de plusieurs années.

Plus que la réduction de la taille de l'échantillon qui peut biaiser l'inférence statistique, le phénomène d'attrition sous-estime les prédictions en raison des biais de sélection. A partir des données STAR, Nye, Hedges, et Konstantopoulos (1999) ont réussi à corriger les biais engendrés par les problèmes d'attrition et ont pu révéler une différence significative entre les élèves sortant des petites classes et ceux des classes de taille moyenne. Autrement dit, les élèves

appartenant à des petites classes réussissent mieux que les élèves issus de classes de taille moyenne.

D'un autre point de vue critique, les études expérimentales menées dans le champ éducatif sont très différentes de celles réalisées dans le domaine de la santé. Hanushek (1998) explique cette différence par le fait que les élèves et les enseignants sont conscients de tout le processus dans lequel se déroule l'expérimentation. Dans ce champ d'expérimentation, les élèves et les enseignants connaissent très bien les groupes dans lesquels ils sont affectés. De ce fait, les enseignants adoptent un comportement différent selon qu'ils sont affectés à des petites classes ou à des classes de taille moyenne.

Les enseignants affectés à des petites classes peuvent être influencés par cette distribution aléatoire. Les enseignants peuvent réagir en fonction de cette affectation en fournissant plus d'effort afin de tirer vers le haut la performance globale de leurs élèves.

A travers son étude expérimentale, Hoxby (2000) montre une élasticité très importante de l'effet de la taille des classes. En effet, une classe de petite taille offre à l'enseignant un environnement favorable de travail l'incitant à plus d'efficacité en matière d'apprentissage.

Ainsi, l'expérimentation modifie les incitations des enseignants des différents groupes de classes. En revanche, il est possible que les enseignants des classes de taille moyenne offrent plus d'effort afin de compenser le désavantage perçu par la taille de la classe. Selon l'analyse de court terme, l'expérimentation de dédoublement des cours préparatoires menée en France en 2002-2003 (Ministère de l'éducation, DEPP²⁶) montre les mêmes résultats positifs dégagés à partir des données STAR.

Ainsi, une redistribution des élèves des classes préparatoires (CP) d'une classe de taille moyenne (20 à 24 élèves) à une classe de petite taille (10 à 12 élèves) a augmenté les scores réalisés en fin de ce niveau d'étude (CP).

En revanche, ce résultat positif remarquable n'a pas perduré car les résultats réalisés avec des classes de petite taille ont disparu un an après l'expérimentation. Une des explications est qu'une réduction de la taille des classes doit être accompagnée d'une durabilité pendant tout le cycle du primaire.

Par ailleurs, l'analyse des caractéristiques observables des élèves choisis pour passer l'expérimentation montre un écart important entre les élèves des CP assignés à des tailles de classes réduites et celles des autres catégories des classes. Ainsi, les conditions du déroulement

²⁶Direction de l'évaluation, de la prospective et de la performance

de l'expérimentation et les critères de sélection des élèves peuvent biaiser les résultats et les analyses y afférents.

Les études expérimentales sont très rares en raison de la difficulté de mobiliser plusieurs acteurs (élèves, enseignants, etc.), et le financement nécessaire à l'élaboration d'une expérimentation dans toutes ses phases. Compte tenu de l'ampleur des impacts à estimer, les échantillons choisis doivent être de très grande taille, et toute la difficulté réside dans l'affectation aléatoire des élèves et des enseignants aux différentes catégories de classes imposées par le dispositif.

1.3. Méthode des variables instrumentales

Ainsi, selon la revue de la littérature nationale et internationale, il apparaît que la plupart des résultats empiriques proviennent des études économétriques avec ou sans correction des biais engendrés par la modélisation effectuée.

Pourtant, à partir des années quatre-vingt-dix, de nouvelles tentatives ont été menées dans le but de corriger les biais liés à l'endogénéité de la taille des classes à l'aide de la méthode des variables instrumentales. Akerhielm (1995) a utilisé deux instruments à savoir la moyenne de la taille des classes dans chaque école et la somme des effectifs des élèves à un niveau donné afin de corriger le biais de l'endogénéité.

Cette méthodologie permet d'éliminer l'effet propre estimé de l'endogénéité engendrée par les caractéristiques des élèves des différentes classes d'une école. Cependant, les instruments calculés à partir des variables observées n'ont pas permis de corriger le biais d'endogénéité engendré par l'affectation non aléatoire des élèves entre les écoles.

En exploitant un panel représentatif des élèves issus des collèges américains (24000 élèves de la deuxième année collégiale), les modèles estimés montrent un effet négatif et significatif d'une augmentation de la taille de la classe sur la performance des élèves, et ceci dans toutes les matières à savoir les mathématiques, l'anglais, l'histoire et les sciences.

En utilisant les données de 11 pays du programme d'évaluation des acquis scolaires « TIMSS », Wossman et West (2006) ont revisité l'approche des variables instrumentales afin d'estimer l'effet de la taille des classes sur les acquis scolaires des élèves en mathématiques.

Ces derniers utilisent la taille moyenne des classes de l'école comme variable instrumentale afin d'estimer l'effet intra-école tout en tenant compte de l'hétérogénéité inter-école en estimant des modèles à effets fixes.²⁷

L'impact de la taille des classes sur le score des élèves estimé par la méthode des moindres carrés pondérés est positif et négatif, significatif et non significatif selon les cas étudiés.

Toutefois, il faut souligner que les données collectées (TIMSS) par les auteurs sont moins riches que d'autres données nationales et internationales compte tenu de l'absence d'indicateurs pour une mesure historique de la performance des élèves.

A titre illustratif, les coefficients estimés pour la France, la méthode des WLS²⁸ montre un impact positif et significatif en science, mais non significatif en mathématiques. En revanche, en considérant l'estimation des effets fixes des écoles, les résultats révèlent des impacts positifs et significatifs en mathématiques.

L'estimation à l'aide de la méthode des variables instrumentales modifie les impacts estimés. Les résultats des estimations par la méthode des VI montrent un effet négatif consistant et convergent en mathématiques, tandis que les effets estimés pour les sciences restent sans effet et non significatif. L'analyse de la variance montre une association négative de -2,7% de l'écart type entre la taille de la classe et la performance scolaire des élèves.

En analysant la relation par des régressions en discontinuité, les auteurs montrent l'absence de l'effet de la taille des classes sur les performances scolaires des élèves et parfois même des résultats contradictoires avec la littérature en révélant un impact positif d'une augmentation du nombre d'élèves dans une classe.

Une des explications de ces résultats tient à la coexistence de deux impacts contradictoires. Le premier renvoie à un impact négatif d'une augmentation de la taille des classes sur le ratio élève-enseignant, et le second à l'impact positif de l'augmentation de ce ratio sur la probabilité de profiter des effets de regroupement en raison de la présence des élèves performants en termes d'acquis scolaire.

L'impact estimé de la taille des classes se modifie automatiquement après l'introduction de nouvelles variables mesurant les effets de pairs, changeant les paramètres estimés vers des signes négatifs dans certains modèles.

D'autres auteurs sont allés plus loin en proposant de nouvelles spécifications en considérant la variation inter-temporelle de la taille des classes pour une matière donnée (Heinesen ;2010).

²⁷Cette stratégie d'identification nécessite la collecte des données des deux cohortes disponibles du programme TIMSS afin d'avoir une structure longitudinale des données.

²⁸Weighted Least Squares Estimation (WLS) method

Cette évolution de la taille des classes au sein des écoles est considérée comme une variation aléatoire de cette dernière et donne lieu à une information pertinente pour la correction des biais via les scores des élèves réalisés dans les autres disciplines.

Analysant la question sur le système éducatif danois, et en appliquant la méthodologie décrite auparavant, Heinesen montre des impacts consistants, convergents et significatifs de la taille des classes sur les scores des élèves réalisés en français.

Les impacts estimés sont plus prononcés chez les élèves les moins performants et même chez les garçons.

Par ailleurs, Angrist et Lavy (1999) ont développé l'approche du seuil d'ouverture et de fermeture de classe, considérée comme l'une des stratégies d'identification les moins critiquables.

Angrist et Lavy (1999) analysent la question des écoles primaires en Israël en exploitant la règle administrative d'ouverture des classes dans les écoles publiques qui s'inspire d'ailleurs d'une règle édictée au 12^{ème} siècle par le philosophe Maimonide. La démarche décrit le raisonnement par lequel une classe est ouverte si la classe de référence dépasse 40 élèves. Cette règle engendre une diminution de la taille des classes moyenne par l'ajout d'une nouvelle classe qui contient au moins un seul élève. En fait, la réflexion de la loi de Maimonide exige l'application d'une régression en discontinuité mettant la relation entre la taille des classes et le nombre d'élèves du niveau traité.

Lorsque l'effectif global des deux classes atteint les 80 élèves, le nombre des classes reste inchangé, et la taille des classes moyenne atteint 40 élèves.

Si on poursuit le raisonnement, l'introduction d'un nouvel élève entraîne la création de la troisième classe, et par conséquent la taille des classes moyenne chute à 27 élèves.

En utilisant cette variable instrumentale, Angrist et Lavy (1999) appliquent cette spécification sur les données israéliennes afin d'identifier l'effet propre de la taille des classes sur les résultats des élèves issus des écoles primaires.

Le modèle proposé par ces auteurs repose sur la régression du score moyen des élèves d'une classe sur les variables décrivant les caractéristiques de la classe en particulier, et de l'école de façon générale. Parmi les variables explicatives de cette spécification, figurent d'une part la proportion des élèves issus des milieux défavorisés, et d'autre par la taille des classes.

Cette spécification est estimée pour chaque niveau scolaire (le 3^{ème}, 4^{ème} et le 5^{ème} grade) par la méthode classique des moindres carrés ordinaires et également en identifiant les paramètres sans biais à travers la méthode des variables instrumentales.

Sans aucune correction des biais via l'origine sociale de l'élève, les résultats de l'estimation par la méthode des moindres carrés ordinaires montrent un effet positif et significatif de la taille des classes sur la performance scolaire en mathématiques et en lecture. Ensuite, la même spécification a été modifiée par l'introduction d'une variable décrivant l'origine sociale de l'élève, et à travers laquelle les impacts estimés sont clairement plus faibles dans leur globalité. De même, le modèle révèle des impacts positifs et significatifs dans certains cas.

En effet, pour les niveaux des 4^{ème} et 5^{ème} grades, la méthode des variables instrumentales corrige entièrement les impacts estimés. Cette spécification indique un impact négatif et significatif de la taille des classes sur les scores des élèves dans les deux disciplines à savoir les mathématiques et la lecture. Cet effet significatif est rejeté pour l'estimation spécifique au 3^{ème} grade. Selon Angrist et Lavy, ce résultat s'explique par le fait que l'impact de la taille des classes procède d'une façon additionnelle. Autrement dit, si l'élève fréquente une classe de petite taille en 5^e grade, il existe une forte probabilité qu'il ait fréquenté une classe de taille comparable les années précédentes.

Hoxby (2000) propose de nouvelles stratégies d'identification en exploitant l'approche du seuil d'ouverture et de fermeture des classes aux Etats-Unis. L'auteur s'appuie sur le principe d'identification de Angrist et Lavy et en développe une version longitudinale. Il exploite des régressions en discontinuités car il apparaît clairement à partir des simples statistiques descriptives qu'il existe des changements discontinus de la taille des classes. Cette discontinuité apparaît lorsque les variations du nombre global des élèves d'un niveau d'une école engendrent un dépassement au seuil d'ouverture ou de fermeture de classe.

Les impacts estimés par les différentes stratégies d'identification montrent l'absence d'effet de la taille des classes. La précision des modèles proposés par Hoxby (2000) permet de rejeter l'existence d'effet pour les classes de taille modeste.

Le travail de recherche de Hoxby (2000) est considéré parmi les travaux les plus en désaccord avec les travaux de Angrist et Lavy (1999) et de Krueger (2000), puisque sa recherche micro économétrique propose des stratégies d'identification innovantes pour la correction des biais d'endogénéité.

A partir de ces analyses critiques fondées sur les données de STAR, Hoxby a remis en cause le dispositif des études expérimentales tout en expliquant que les acteurs participant à l'expérimentation ont au préalable toute l'information nécessaire pour modifier leurs comportements lors de toutes les phases de l'étude.

Analysant la question en Bolivie, Urquiola (2006) essaie d'évaluer l'impact de la taille des classes pour le niveau primaire des écoles rurales. L'auteur cherche de se pencher

exclusivement sur le milieu rural afin de contrôler tous mécanismes (choix des parents, affection non aléatoire des élèves, etc.) conduisant à identifier sans biais l'effet de la taille des classes.

Tout d'abord, l'auteur a fait le choix d'examiner les écoles contenant seulement une classe dont le principe d'une affection non aléatoire des élèves entre classes n'est pas pris en considération. Par ailleurs, ce type d'établissement n'offre pas aux parents la possibilité de faire des choix entre les écoles. Dans ce contexte, le biais de sélection engendré par la fréquentation des classes de petite taille est limité ce qui est vérifié par les régresseurs observables des élèves. En revanche, un biais de sélection peut être présent en raison des paramètres liés aux caractéristiques inobservables des élèves.

Pour remédier au problème d'endogénéité, l'auteur exploite le même instrument (taille maximale de 30 élèves par classe) utilisé par Angrist et Lavy. Les résultats obtenus pour les élèves du 3^{ème} grade révèlent un impact négatif de la taille des classes en mathématiques et en lecture.

Adoptant la même démarche d'identification, Jakubowski et Sakowski (2006) appliquent le principe de discontinuité pour l'étude des écoles primaire polonaises.

Les auteurs étendent les travaux empiriques en utilisant la taille moyenne des classes comme instrument pour remédier au phénomène d'une affectation non aléatoire des élèves entre classe au sein de l'école. L'objectif de l'estimation d'un modèle à effet fixe est de contrôler la variabilité entre écoles et entre zones géographiques. Cette stratégie d'identification est d'ailleurs proche de celle utilisée par Wossmann et West (2003, 2006) et Akerhielm (1995).

Cette dernière stratégie d'identification exploite le même instrument proposé par Angrist et Lavy (1999), le principe d'ouverture et de fermeture d'une classe a été appliqué à trois niveaux afin d'évaluer le cycle primaire du système éducatif polonais.

Commençant par une estimation classique par la méthode des moindres carrés ordinaires, leurs résultats montrent un effet positif de la taille des classes sur les résultats scolaires obtenus par les élèves. Par la suite, cette association positive disparaît en prenant en compte le biais d'endogénéité lié à la taille des classes, et par conséquent l'effet estimé pour toutes les classes demeure négatif et significatif. De même, les impacts estimés apparaissent plus faibles en termes d'amplitude que les résultats obtenus par Krueger (1999) et Angrist et Lavy (1999).

En étudiant le système éducatif danois, Browning et Heinesen (2007) montrent un effet négatif de la taille des classes sur le niveau maximum d'études atteint par les élèves mesurés par le nombre d'années de scolarisation. Cet examen a permis de révéler l'existence d'une règle d'ouverture ou de fermeture de classe à partir d'une classe de 24 élèves.

Ainsi, les auteurs ont estimé l'effet du ratio effectif d'élèves/nombre d'heures d'enseignement par semaine, la variable instrumentale a été déduite à partir de la politique administrative imposée par le système éducatif. L'apport de cette dernière variable mesurant l'offre scolaire tient au fait que la taille des classes comme seule variable de mesure des ressources scolaires biaise les estimations. Ce biais d'estimation est expliqué par le fait qu'une allocation additionnelle du nombre d'heures d'enseignement pour les classes de grande taille s'avère une politique de compensation par rapport aux classes de petite taille. L'effet de cette seconde stratégie d'identification est également négatif et significatif. De même, l'amplitude de ces impacts semble très importante, même si les effets ne sont pas vraiment mesurés avec précision (Valdenaire, 2011).

En adoptant la même démarche utilisée par Angrist et Lavy (1999) et Hoxby (2000), Leuven Oosterbeek et Roenning (2008) étudient les performances du système éducatif norvégien en exploitant le principe d'ouverture et de fermeture des classes. Cette dernière repose sur la règle du maximum de taille des classes.

L'objectif des modèles proposés est de mesurer l'effet de taille des classes à la fin du cycle collégial en mathématiques, anglais, et en langue norvégienne. Les résultats de l'étude ont révélé un effet non significatif de la taille des classes sur les résultats des élèves. Ce résultat est valable pour les deux méthodes d'estimations, et même après décomposition de l'échantillon en trois groupes distincts (selon la composition sociale, le sexe et les caractéristiques des professeurs).

En France, les premiers travaux sur cette question ont été menés par Meuret (2001) sur la base des dispositifs statistiques élaborés vers la fin des années 70. Ce n'est qu'à la fin des années 80 que les chercheurs ont adopté l'approche statistique économétrique pour une mesure quantitative. De même, les chercheurs et experts des sciences de l'éducation ont traité la question d'une façon générale, dans le but de répondre seulement à la problématique des déterminants de la réussite scolaire des élèves.

Dans ce contexte, le problème d'endogénéité engendré par la taille des classes a été négligé, car les modèles estimés au cours de cette période se sont limités à identifier les paramètres d'impacts via des variables observables relevant des caractéristiques des élèves. En outre, ces études ont en général été réalisées sur des échantillons qui sont généralement de taille réduite et spécifiques à une zone géographique déterminée. Les résultats de ces travaux aboutissent à des résultats très hétérogènes voire contradictoires.

Mingat (1991) montre une association négative entre la taille de la classe et la réussite scolaire en français pour les élèves des cours préparatoires. D'autres auteurs comme Suchaut (1997),

Suchaut et Le Bastard (2000) révèlent des effets non significatifs et même parfois positifs, ce qui paraît pour le moins contradictoire.

Les travaux de Meuret (2001) sur les cycles collégiaux et lycéens du système éducatif français ont introduit un doute quant à l'existence d'un impact significatif. Cependant, la plupart des résultats démontrés par cet auteur n'ont pas pris en considération les biais engendrés par l'endogénéité de la variable de la taille des classes. En fait, l'auteur montre par la suite que l'introduction des variables de contrôle dans le modèle économétrique révèle un effet négatif et significatif de la taille des classes sur la performance scolaire des élèves.

Il est par ailleurs remarquable qu'au cours du début années 2000, la plupart des études françaises se sont contentées d'analyser uniquement l'école primaire.

En faisant appel à une stratégie d'identification des seuils d'ouverture et de fermeture de classe, Gary-Bobo et Mahjoub (2013) ont proposé plusieurs instruments en se basant sur la présence de ces seuils dans les écoles du cycle collégial. Les auteurs ont exploité les données du panel secondaire 1995 réalisé par la Direction de l'Évaluation de la Performance et de la Prospective au Ministère de l'Éducation Nationale (MEN-DEPP) dans le but d'estimer l'effet de la taille de la classe sur la probabilité du redoublement à chaque niveau du cycle collégial.

Les résultats obtenus montrent un effet négatif et significatif de la taille de la classe sur la réussite scolaire des élèves de la 6^e et de la 5^e année. De même, les auteurs révèlent que pour une égalisation de la probabilité de redoublement des deux groupes d'élèves issus de milieux sociaux différents (élèves de cadres et d'ouvriers), une politique de réduction de la taille des classes à 10 élèves en 6^{ème} année serait efficace pour réduire les inégalités éducatives en matière d'acquis scolaire. En revanche, cet effet significatif n'est pas vérifié pour la 4^e et la 3^e année.

Pour conclure, Bressoux, Kramarz et Prost (2009) ont focalisé leur réflexion sur la formation des enseignants comme un facteur important pour une spécification moins biaisée mettant la relation entre la taille de la classe et la réussite des élèves du CE2. De plus, les auteurs inscrivent leur stratégie d'identification dans le cadre d'une affectation non aléatoire des élèves entre les classes.

A partir de leur méthode quasi-expérimentale, les auteurs constatent l'existence d'une corrélation entre la taille des classes et les résultats des élèves du CE2 pour l'ensemble de l'échantillon. Plus encore, leurs stratégies d'estimation reposent sur une hypothèse forte d'absence de toute corrélation entre la taille des classes et les caractéristiques non observables. Les impacts estimés sont proches des résultats révélés par la littérature internationale et en particulier des paramètres identifiés à travers le dispositif expérimental STAR. Cet effet est d'environ 3% d'un écart type. En revanche, les résultats révèlent l'absence d'un effet

significatif en mathématiques et en français. De même, l'étude a permis un contrôle de l'hétérogénéité individuelle en recourant aux modèles à effets fixes et à effets aléatoires.

Cette spécification a montré que l'effet de la taille des classes est plus important pour les élèves et les classes des écoles situées en ZEP ²⁹, et l'impact estimé est d'environ 5% d'un écart type en français et de 8% en mathématiques. Les auteurs testent leurs résultats en utilisant deux instruments à savoir la taille moyenne des classes dans l'école et la taille maximale de la classe. Les résultats obtenus par cette dernière stratégie sont presque similaires à la première stratégie d'identification. En moyenne, l'impact de la taille des classes est d'environ 2,5% d'un écart type.

Le débat autour de l'effet de la réduction de la taille des classes (RTC) sur les acquisitions des élèves a attiré beaucoup d'attention de la part des décideurs et des chercheurs dans le domaine de l'éducation. Sa mesure reste complexe et les travaux menés dans ce sens aboutissent souvent à des résultats contradictoires.

Certes, la réduction de la taille des classes est une politique éducative très coûteuse, mais une meilleure compréhension de tel effet reste nécessaire pour orienter les politiques éducatives vers des décisions sélectives en matière de cette réduction.

2. Données utilisées et faits stylisés

Notre étude empirique porte sur un échantillon représentatif de la population totale des élèves marocains de la 4^e année primaire et 2^e année collégiale. Les données collectées sont issues des enquêtes internationales TIMSS ³⁰ et PIRLS ³¹ dont les variables renseignent sur l'ensemble des caractéristiques du système éducatif marocain. La taille de l'échantillon est de 7805 élèves pour l'évaluation des acquis scolaires en lecture, tandis que l'évaluation des acquis scolaires en mathématiques porte sur un échantillon de 8986 élèves. L'enquête PIRLS inclut 289 écoles, tandis que l'enquête TIMSS repose sur un échantillon de 284 écoles. Ainsi, notre recherche porte sur l'évaluation des acquis scolaires de deux générations d'élèves marocains inscrits en 2004 et en 2008 en première année du cycle primaire.

²⁹Zone d'éducation prioritaire

³⁰Trends in International Mathematics and Science Study : Evaluation des performances en mathématique et matière scientifique pour les élèves de la 2e année collégiale.

³¹Progress in International Reading Literacy : Evaluation des performances en lecture pour les élèves de la 4e année primaire

Nous avons sélectionné les variables les plus utilisées dans la littérature afin de proposer un modèle micro économétrique sans biais d'identification. En fait, la variable continue de la taille de la classe sera utilisé pour capter le seuil de la taille optimale de la classe, tandis que les autres variables de contrôle servent à contrôler les paramètres du modèle, et aussi bien le degré de sensibilité de la courbe reliant la taille de la classe et le niveau des acquis solaire mesuré par le score des élèves. Les variables de contrôle que nous considérons renvoient aux caractéristiques de l'élève, de l'enseignant et de l'établissement fréquenté.

Les tableaux 36 et 37 en annexe du chapitre 2 présentent les variables sélectionnées et leurs descriptions détaillées, ainsi que leur codification.

3. Statistiques descriptives sur la taille des classes, les caractéristiques des élèves de la 4^e année primaire et 2^e année collégiale et des inégalités des acquis scolaires

L'objectif de cette section est d'analyser les écarts de performance des élèves de la 4^e année primaire et 2^e année collégiale, tout en focalisant notre analyse sur les différents types de taille des classes. Les bases de données « TIMSS et PIRLS » incluent plusieurs variables contextuelles qui peuvent être à l'origine d'importantes inégalités des acquis scolaires des élèves. Plus précisément, comme nous l'avons souligné, la taille des classes représente un des facteurs essentiels de l'apprentissage des élèves.

Afin de rendre compte des écarts de performance scolaire des élèves selon les différents types de taille des classes, nous choisissons de classer la taille de la classe en trois modalités en les différenciant selon le niveau d'enseignement. Le tableau 27 présente les modalités associées à chaque catégorie. Le choix de cette catégorisation se justifie en partie des pourcentages observés dans chacune de taille des classes, des scores moyens de chaque catégorie et des résultats attendus de notre modèle économétrique.

Tableau 27. Catégorie des tailles de la classe

	4^e année primaire	2^e année collégiale
Taille réduite	Inférieur à 24 élèves	Inférieur à 27 élèves
Taille moyenne	Entre 25 et 35 élèves	Entre 28 et 37 élèves
Classe chargée	Plus de 36 élèves	Plus de 38 élèves

Le tableau 28 décrit la répartition de la taille des classes des deux niveaux d'enseignement du système éducatif marocain. Nous constatons que 40% des classes de la 4^e année primaire correspondent à des classes chargées, 47 % des classes sont de taille moyenne tandis que seules 13% des classes sont de petite taille.

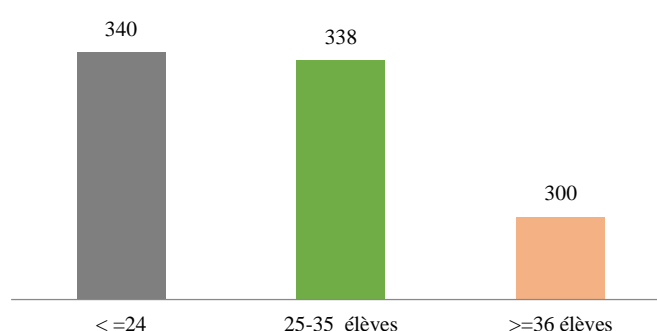
Tableau 28. La répartition de la taille des classes

Catégorie de la taille des classes	4 ^e année primaire			2 ^e année collégiale		
	Inférieur à 24 élèves	Entre 25 et 35 élèves	Plus de 36 élèves	Inférieur à 27 élèves	Entre 28 et 37 élèves	Plus de 38 élèves
%	13%	40%	47%	1%	50%	39%

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

La 2^e année collégiale présente une distribution quasiment identique avec 50% de classes chargées, 39% de classes de taille moyenne et seulement 11% de petites classes.

Graphique 20. Le score moyen en lecture selon la taille des classes



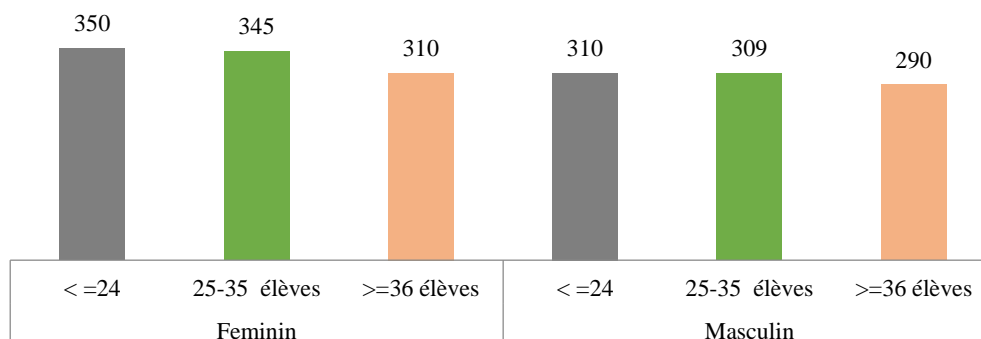
Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Comme nous l'avons mentionné dans le chapitre 1, le score moyen des élèves de la 4^e année du primaire reste faible comparativement à la moyenne internationale. Cette insuffisance est fortement corrélée avec la taille des classes car le score moyen en lecture diffère selon les types de taille des classes. En effet, on constate que les élèves associés à une taille de classe inférieure à 24 élèves réussissent mieux que les élèves assignés à des classes chargées avec un écart de scores de l'ordre de 40 points (Graphique 20). Cependant, nous observons que la différence de performance en lecture est peu tangible entre les petites et moyennes classes ; elle n'est que de deux points seulement.

L'analyse peut être approfondie en examinant l'association entre la performance scolaire et certaines variables relatives à l'élève (genre et retard scolaire), aux enseignants (genre, expérience de l'enseignant, charge horaire, etc.) ainsi qu'aux caractéristiques associées à l'école (milieu de l'établissement, disponibilité des bibliothèques, etc.) selon la taille des classes.

3.1. Disparités des performances scolaires en lecture des élèves de la 4^e année du primaire selon la taille des classes

Graphique 21. Le score moyen en lecture selon le genre et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

La prise en compte du genre de l'élève confirme la baisse de performance avec la taille de la classe. Les écarts de performance entre les petites classes et les classes chargées sont de 40 points pour les filles et de 30 points pour les garçons. En revanche, ces écarts de performances sont négligeables, voire inexistantes entre les petites classes et les classes de taille moyenne (Graphique 21).

Tableau 29. Le score moyen en lecture selon le retard scolaire et la taille des classes

Catégorie de taille des classes	2003 ³²			2002			2001			2000			1999			1998		
	TP	TM	TG ³³	TP	TM	TG	TP	TM	TG	TP	TM	TG	TP	TM	TG	TP	TM	TG
Score Moyen	400	398	340	360	355	320	360	358	330	300	290	280	290	292	260	300	290	280

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

En ce qui concerne le retard scolaire, les élèves de la 4^e année primaire ayant l'âge légal pour ce niveau d'éducation en 2011³⁴ réussissent mieux que les élèves enregistrant un retard scolaire. Cependant, le tableau 29 montre que les disparités de performance se manifestent selon la taille des classes et le retard scolaire accumulé par les élèves. Ainsi, les élèves n'ayant aucun retard scolaire progressent mieux dans les petites classes que dans les classes chargées : l'écart de score s'élève à 60 points entre ces deux groupes. Bien que les élèves ayant un retard scolaire réalisent des scores plus faibles, il s'avère que les disparités parmi les élèves selon la taille des

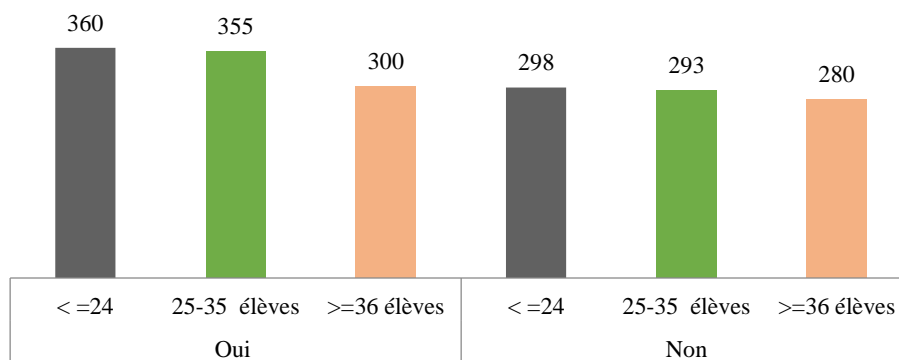
³² Date de naissance des élèves n'ayant enregistré aucun retard scolaire et qui sont scolarisés au niveau de la 4^e année primaire en 2011

³³ Taille petite, taille moyenne et taille grande

³⁴ L'année de scolarisation des élèves de la 4^e année

classes s'estompent avec le retard accumulé. En effet, l'écart de score entre les élèves de classes à effectifs réduits et de classes chargées atteint 40 points pour une année de retard et 20 points pour 3 et 5 années.

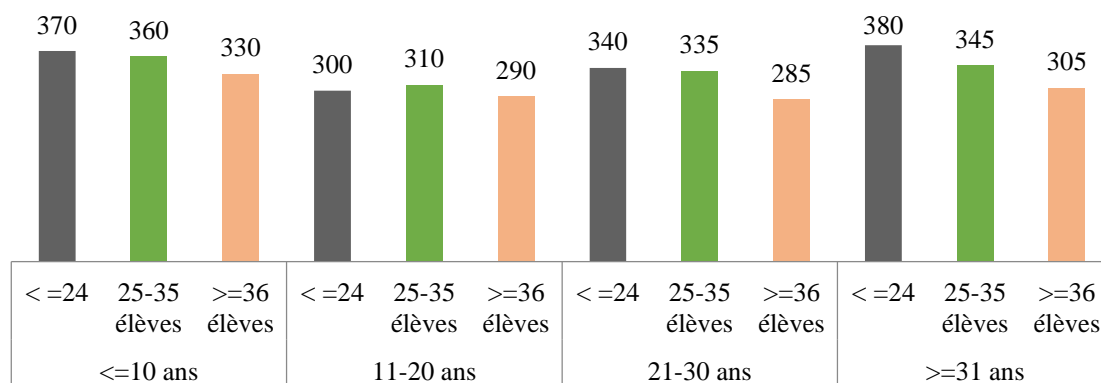
Graphique 22. Le score moyen en lecture selon la préscolarisation et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Ainsi, la performance varie selon la préscolarisation des élèves de la 4^e année du primaire mais aussi selon la taille des classes. En effet, les élèves ayant bénéficié de la préscolarisation réussissent mieux lorsqu'ils sont scolarisés dans des classes de petite taille (score de 360 points) que lorsqu'ils sont accueillis dans des classes chargées (score de 300 points). L'écart de score s'élève à 60 points pour ces deux derniers groupes tandis qu'il n'est que de 18 points entre les élèves non préscolarisés de petites classes et de classes surchargées (Graphique 22).

Graphique 23. Le score moyen en lecture selon l'expérience des enseignants et la taille des classes

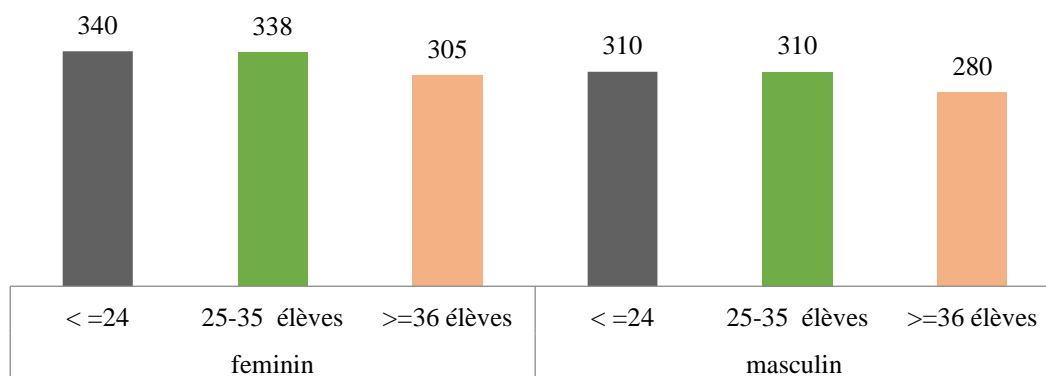


Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

S'agissant du nombre d'année d'expérience de l'enseignant et de la taille des classes, on observe une diminution de la performance scolaire en passant d'une taille de classe à une autre selon l'expérience de l'enseignant (Graphique 23). Les élèves dont l'enseignant a plus de 31 années d'expérience réussissent beaucoup mieux dans des classes de petite taille que dans des

classes chargées ; l'écart de performance s'élevant à 75 points. Ce résultat vaut également pour les élèves ayant un enseignant moins expérimenté (moins de 10 années d'expérience) dont l'écart entre les petites classes et les classes surchargées s'élève à 40 points. Ce résultat montre l'ampleur de la taille des classes dans la différenciation de la performance scolaire des élèves quelle que soit l'expérience de l'enseignant.

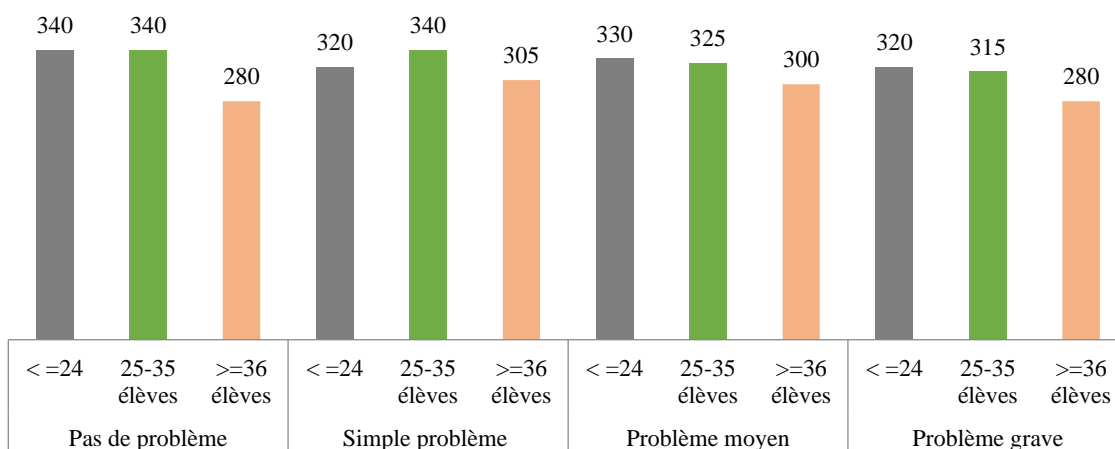
Graphique 24. Le score moyen en lecture selon le genre des enseignants et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Concernant le genre de l'enseignant, nous remarquons toujours une baisse des écarts des performances scolaires selon la taille de la classe. Cependant, la différence selon le genre de l'enseignant ne semble pas confirmer un effet de genre puisque l'écart est de fait seulement de 5 points (Graphique 24).

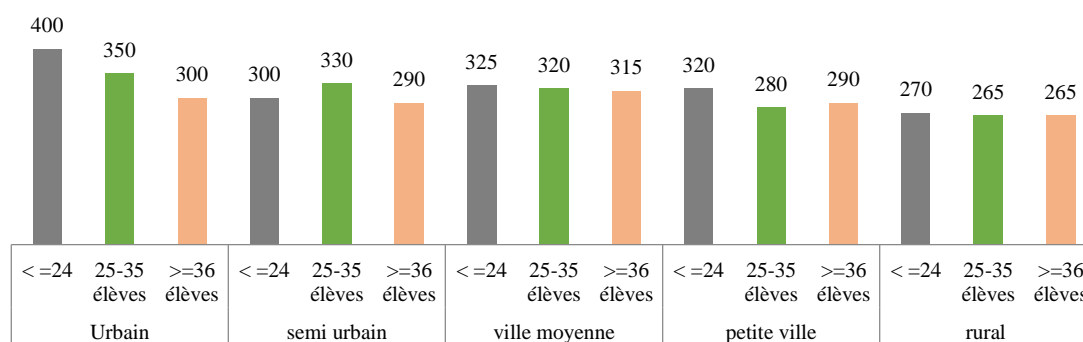
Graphique 25. Le score moyen en lecture selon le degré du problème de la charge horaire et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

S'agissant du problème de la charge horaire (Graphique25), pour les enseignants déclarant n'avoir aucun problème de charge horaire, l'écart de performance entre les élèves des petites classes et des classes de grande taille (scores de 340 contre 280) est de 60 points. Pour les élèves dont l'enseignant déclare avoir des problèmes graves de charge horaire, les écarts de performances des élèves varient dans une moindre mesure selon la taille des classes. Cet écart varie de 15 à 40 points selon la taille de la classe et selon que l'enseignant déclare respectivement la charge horaire comme étant un simple problème ou un problème grave. Bien que les disparités de performance soient plus faible entre petite et grande classes lorsque l'enseignant perçoit la charge horaire comme un problème comparativement à la situation où celle-ci n'est pas perçue comme un problème (75 points contre 40 points), il s'avère que la performance des élèves des petites classes tend à s'affaiblir avec l'ampleur de la charge horaire perçue par l'enseignant (340 points dans le premier cas contre 320 points dans le second cas)., les élèves appartenant à des classes de petite taille enregistrent le score le plus élevé (320 points), tandis que les élèves assignés à des classes chargées sont moins performants avec un score de 280 points de sorte que l'écart de performance est de l'ordre de 40 points. Enfin, si nous considérons les élèves des classes de taille moyenne, les scores s'élèvent à 340 points et pour les enseignants qui n'ont aucun problème avec la charge horaire et de 315 points pour ceux la considérant comme un problème grave.

Graphique 26. Le score moyen en lecture selon le milieu de l'établissement et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Comme nous l'avons mentionné dans le chapitre 1, les élèves issus du milieu urbain réussissent beaucoup mieux que les élèves issus du milieu rural dont le score est en moyenne seulement de 268 points. Le graphique 26 montre qu'en milieu rural, la taille des classes n'a pas d'effet tangible sur ce score. Les écarts de performance selon la taille des classes y sont négligeables voire même inexistantes. En revanche, la situation est différente en milieu urbain. En effet, en milieu urbain, les élèves réussissent beaucoup mieux dans les petites classes que dans les classes

surchargées, l'écart de performance étant de 100 points. Au sein du milieu urbain, les écarts de performance sont de l'ordre de 100 points.

En résumé, notre analyse révèle que les scores moyens peuvent masquer des écarts de performance assez importants lorsque nous considérons l'association entre la taille de la classe et les caractéristiques associées à l'élève et à son environnement. Ainsi, les élèves des classes de petite taille réussissent beaucoup mieux que les élèves des classes surchargées, et les écarts de performance s'amplifient en fonction de certains facteurs contextuels (Milieu de l'établissement, composition sociale de l'école, etc.) (Hanchane et al 2013), et qui peuvent différencier le processus des acquis scolaires des élèves en lecture.

3.2. Disparités des performances scolaires en mathématiques et en sciences des élèves de la 2^e année collégiale selon la taille des classes

Cette section analyse les disparités des performances scolaires en mathématique et en science des élèves de la 2^e année collégiale selon la taille des classes. En fait, il s'agit de décortiquer davantage les scores moyens afin de déceler toute différence susceptible de démontrer que les élèves assignés dans des classes de petite taille sont beaucoup plus performant que ceux affectés dans des classes de grande taille.

Tableau 30. Le score moyen en mathématiques et en sciences selon la taille des classes

	Mathématique			Science			Chimie			Science de la terre			Biologie			Physique					
	TP	TM	TG	TP	T	TG	TP	T	TG	TP	TM	TG	TP	T	TG	TP	T	TG			
				M			M						M			M					
Score	420	370	367	42	380	37	43	380	37	420	380	373	42	380	37	39	350	35			
Moye				0	7			0			6			0	4			0			
n																					

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Ainsi, l'écart de performance en mathématiques entre les élèves de petites classes et les élèves de classes de taille moyenne ou surchargées est de l'ordre de 50 points (tableau 30). Cela vaut également pour les matières scientifiques, puisque l'écart enregistré selon la taille des classes avoisine les 40 points. Comme nous l'avons montré dans le chapitre 1, une augmentation de la taille des classes de la 2^e année collégiale défavorise certainement les pratiques pédagogiques des enseignants, et par conséquent la performance scolaire.

Tableau 31. Le score moyen en mathématiques et en sciences selon le genre des élèves et la taille des classes

		Mathématiques	Sciences
Féminin	<=27 élèves	440	430
	28-37 élèves	370	380
	>=38 élèves	368	373
Masculin	<=27 élèves	430	420
	28-37 élèves	370	380
	>=38 élèves	369	379

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

L'analyse du score moyen en mathématiques et en sciences selon le genre de l'élève et la taille des classes indique pour les filles un écart de 70 points et 50 points en mathématiques et en sciences respectivement (tableau 31) entre les petites et les classes surchargées. Ces écarts sont proches de ceux enregistrés pour les garçons ; ils sont de 60 points en mathématiques et de 40 points en sciences.

Ainsi, les élèves appartenant à une classe de petite taille enregistrent un score de 498 points en mathématiques, tandis que les élèves accueillis dans des classes de taille moyenne et grande n'enregistrent que des scores de 440 points à 380 points respectivement. Pour le même niveau d'instruction des parents, l'écart des scores en sciences entre les différents types de taille des classes est similaire à celui observé pour l'évaluation des mathématiques.

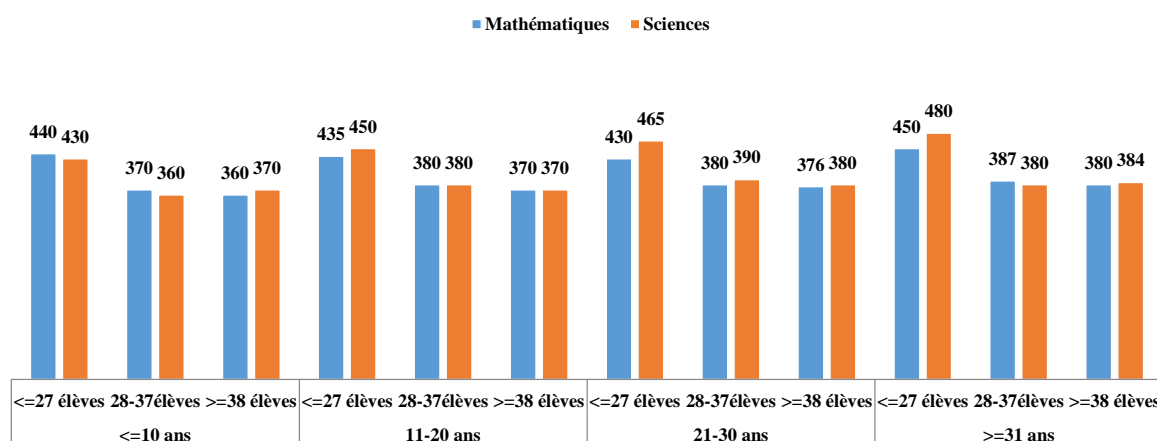
Tableau 32. Le score moyen en mathématiques et en sciences selon le niveau d'instruction des parents et la taille des classes

		Mathématiques	Sciences
Parents analphabètes	<=27 élèves	380	390
	28-37 élèves	380	390
	>=38 élèves	380	389
Parents avec un niveau d'éducation primaire	<=27 élèves	420	420
	28-37 élèves	384	380
	>=38 élèves	384	380
Parents avec un niveau d'éducation collégial ou lycée	<=27 élèves	460	440
	28-37 élèves	400	400
	>=38 élèves	398	390
Parents avec un niveau d'éducation de l'enseignement supérieur "Licence"	<=27 élèves	500	490
	28-37 élèves	440	420
	>=38 élèves	380	410
Parents avec un niveau d'éducation de l'enseignement supérieur "Master ou doctorat"	<=27 élèves	498	490
	28-37 élèves	340	345
	>=38 élèves	380	385

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Par ailleurs, les élèves ayant des parents déclarant avoir seulement un niveau primaire ou non scolarisés, réalisent des scores en mathématiques et en sciences similaires, quel que soit la taille de la classe. Pour les autres catégories de niveaux d’instruction des parents, le score moyen diminue en fonction de la taille de classe avec des écarts de performance allant de 20 à 90 points selon la matière.

Graphique 27 . Le score moyen en mathématiques et en sciences selon l’expérience des enseignants et la taille des classes

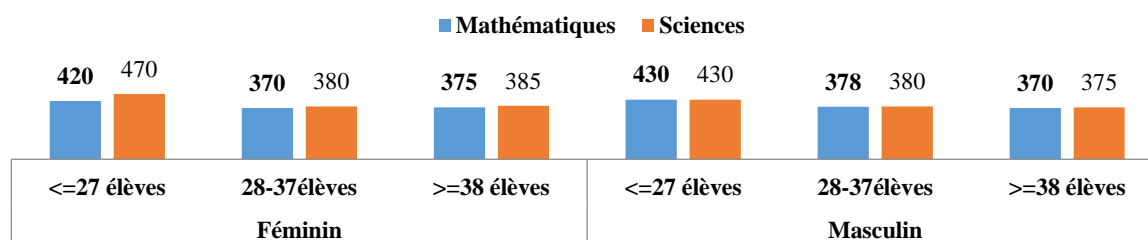


Source : Calculs de l’auteur à partir des données PIRLS et TIMSS 2011

En désagrégant les scores moyens par le nombre d’années d’expérience des enseignants, on remarque que les élèves scolarisés dans les classes de taille de classe moyenne ou chargée enregistrent des performances moins élevées que les élèves de classes de petite taille, et ceci quel que soit le nombre d’années d’expérience des enseignants (Graphique 27). A titre d’illustration, les élèves dont l’enseignant a plus de 31 années d’expérience réussissent beaucoup mieux dans les petites classes que dans les classes chargées. Les écarts de performance sont de 80 points et 100 points en mathématiques et en sciences respectivement.

Pour les élèves dont l’enseignant a moins de 10 années d’expérience, l’écart de performance enregistré entre les élèves des petites classes et des classes moyennes et surchargées avoisine les 70 points en mathématiques et en sciences.

Graphique 28 .Le score moyen en mathématiques et en sciences selon le genre des enseignants et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Les écarts de performance en mathématiques et en sciences varient selon le genre des enseignants (Graphique 28). Les écarts de performances entre les élèves de petites classes et de classes moyennes ou chargées sont de l'ordre de 45 points et 50 points en mathématiques et en sciences respectivement. Ces écarts restent comparables qu'il s'agisse d'un enseignant ou d'une enseignante.

S'agissant de la localisation de l'établissement, en milieu urbain, les élèves réussissent que les élèves issus du milieu rural (tableau 33). Cependant, en milieu urbain les disparités entre les élèves sont d'autant plus prononcées que la taille de la classe est petite plutôt que surchargée. Les écarts de performance selon la taille de la classe y sont de l'ordre de 60 points et 100 points en mathématiques et en sciences respectivement.

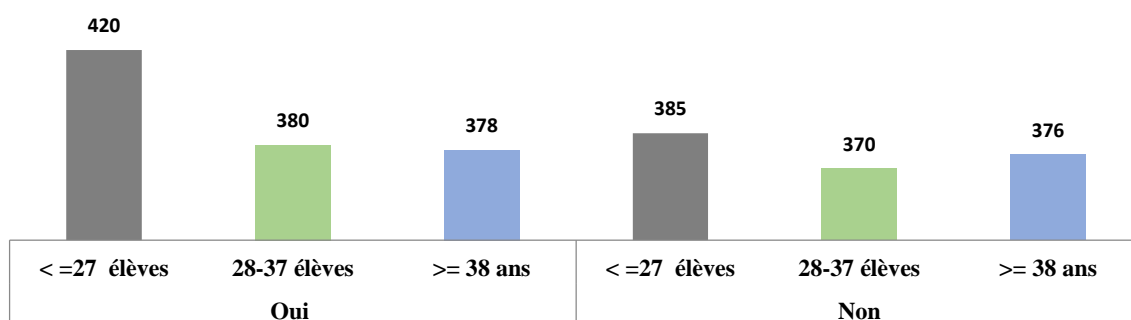
Tableau 33. Le score moyen en mathématiques et en sciences selon le milieu de l'établissement et la taille des classes

		Mathématiques	Sciences
Urbain	<=27 élèves	460	480
	28-37 élèves	380	380
	>=38 élèves	375	376
Semi-urbain	<=27 élèves	310	380
	28-37 élèves	380	385
	>=38 élèves	370	370
Ville de taille moyenne	<=27 élèves	450	430
	28-37 élèves	370	380
	>=38 élèves	365	376
Petite ville	<=27 élèves	370	380
	28-37 élèves	347	368
	>=38 élèves	340	360
Rural	<=27 élèves	310	380
	28-37 élèves	350	350
	>=38 élèves	-	-

Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

Ces disparités selon la taille de la classe sont beaucoup moins importantes en milieu rural car l'écart n'est seulement que de 30 points. En revanche, les résultats semblent indiquer qu'en mathématiques, les élèves réussissent mieux dans les classes de taille moyenne que dans les classes de petite taille avec un écart de 40 points.

Graphique 29 .Le score moyen en sciences selon la présence d'un laboratoire et la taille des classes



Source : Calculs de l'auteur à partir des données PIRLS et TIMSS 2011

La disponibilité d'un laboratoire scientifique à l'école fait ressortir des différences en termes de performance scolaire. Selon le graphique 35, nous constatons que dans les écoles disposant d'un laboratoire les élèves des petites classes enregistrent un score de 420 points contre seulement 378 pour les élèves de classes chargées. Ces écarts sont cependant moins prononcés dans les écoles ne disposant pas d'un laboratoire scientifique puisqu'ils sont de l'ordre de 10 points selon la taille des classes.

Tableau 34. Le score moyen en mathématiques et en sciences selon le degré de mixité sociale

		Mathématiques	Sciences
Ecole favorisé	<=27 élèves	430	420
	28-37élèves	370	365
	>=38 élèves	320	338
Ecole mixte	<=27 élèves	400	398
	28-37élèves	350	334
	>=38 élèves	320	310
Ecole défavorisée	<=27 élèves	379	367
	28-37élèves	355	344
	>=38 élèves	290	298

Source : Calculs de l'auteur à partir des données TIMSS 2011

Le tableau 34 montre que la performance scolaire se différencie au sein de chaque catégorie de composition sociale et économique de l'établissement selon la taille de la classe. En effet, le score en mathématique diminue en fonction de l'augmentation de la taille de classe avec des

grandeurs différenciées selon la catégorie d'école. Le score passe de 430 à 290 en mathématique et de 420 à 298 en science (tableau 34).

Tableau 35. Le score moyen en mathématiques et en sciences selon l'âge de l'enseignant

		Mathématiques	Sciences
25-29 ans	<=27 élèves	405	410
	28-37 élèves	365	355
	>=38 élèves	303	310
30-39 ans	<=27 élèves	420	412
	28-37 élèves	390	387
	>=38 élèves	320	315
40-49 ans	<=27 élèves	398	387
	28-37 élèves	365	355
	>=38 élèves	310	308
50-59 ans	<=27 élèves	410	413
	28-37 élèves	387	398
	>=38 élèves	376	387

Source : Calculs de l'auteur à partir des données TIMSS 2011

De même, le score des élèves en science et en mathématique se manifeste par des différences selon l'âge de l'enseignant et la taille de la classe dont l'enseignant est affecté. En mathématique, les scores montrent des écarts allant de 405 pour les enseignants âgés de 25-29 affectés à une classe de petite taille à 303 points pour la même catégorie d'enseignant assignée à une classe de grande taille. Ce constat est remarqué pour les différentes tailles de classes et pour les différentes classes d'âge des enseignants (tableau 35).

Enfin, l'analyse des scores des élèves selon la taille des classes et différentes variables contextuelles indique clairement que les élèves appartenant à des classes de petite taille réussissent mieux que les élèves de classes surchargées. Cette conclusion est vérifiée pour les deux niveaux d'enseignement à savoir la 4e année du primaire et la 2e année collégiale. Afin d'estimer l'effet net de la taille des classes sur la performance scolaire en lecture, en mathématiques, et en sciences et d'identifier la taille optimale de la classe à partir de laquelle le rendement scolaire se détériore d'une façon significative, nous proposons dans ce qui suit une modélisation semi-paramétrique multi-niveaux.

4. Méthodologie d'estimation

Dans cette section, nous justifions et présentons la méthodologie que nous proposons d'adopter pour estimer l'effet de la taille de la classe sur ses performances scolaires de l'élève. Il s'agit de spécifier un modèle semi-paramétrique sous une spécification hiérarchique, tout en développant la méthode d'estimation ainsi que l'apport de cette méthodologie par rapport aux modèles paramétriques.

Au cours des vingt dernières années, les modèles multiniveaux linéaires et leurs extensions sont devenus des modèles économétriques très utilisés pour analyser les données ayant une structure hiérarchique. La prise en compte de la variabilité dite « intra-groupe » permet de traiter des données longitudinales et regroupées, et par conséquent de réaliser une décomposition de la variance en composantes inter-groupe et intra-groupe.

En fait, l'inférence statistique avec des modèles à effets mixtes « multiniveaux » linéaires a été largement étudiée dans le contexte de données longitudinales, notamment par Verbeke et Molenberghs (2000), ainsi que par Diggle et al. (2013). Dans le même temps, les modèles à effets mixtes sont également adaptés à l'estimation de petites zones (Jiang et Lahiri, 2006), à l'appariement des données ou à la cartographie de la pauvreté (C. Elbers et al. 2003).

Au cours des dernières années, à la suite du travail de Wu et Zhang (2006), il y a eu un intérêt notable à étendre les modèles paramétriques à des spécifications non paramétriques et semi-paramétriques. Parallèlement aux différentes formulations non-paramétriques et semi-paramétriques notamment de Zhang et al (1998), plusieurs méthodes statistiques pour réaliser des inférences adéquates à ce type de modélisation ont été développées. Les approches fondées sur la vraisemblance ont été proposées, notamment par Lin et Zhang (1990) et Lombardía et Sperlich (2008 et 2011). Par la suite, ces auteurs ont développé des méthodes de lissage pour l'estimation des modèles non paramétriques.

Le principal problème lors de l'utilisation des méthodes de lissage réside dans le choix du paramètre approprié pour réaliser le lissage pertinent. Gu et Ma (2005) ont proposé une formulation généralisée dans le contexte du lissage spline (splinesmoothing). Plus tard, Xu et Zhu (2009) ont formulé une spécification des modèles non-paramétriques à effets mixtes basés sur le noyau (Kernel). Notre méthodologie repose sur la méthode du bootstrap itérative dans un cadre d'analyse des modèles multiniveaux non-paramétriques et semi-paramétriques. Cette dernière est très pertinente car il repose sur les techniques itératives (Baltagi 2017) et corrige

naturellement toute les problèmes engendrés par les écarts type standard très connues dans l'estimateur des moindres carrés ordinaires. En effet, le bootstrap sera utilisé pour estimer la courbe non paramétrique et les intervalles de confiances nécessaires pour tester la validité du modèle.

Nous commençons par la formulation du modèle multiniveau non-paramétrique dont l'expression est la suivante :

$$Y_{ij} = m(X_{ij}) + v_i(Z_{ij}) + \varepsilon_{ij}, \quad j = 1, \dots, n_i, \quad i = 1, \dots, q, \quad \sum_{i=1}^q n_i = n \quad (1.1)^{35} \quad (11)$$

Avec Y_{ij} les scores des élèves observés et X_{ij} ($k \times 1$), Z_{ij} ($r \times 1$) les variables explicatives liées aux facteurs intrinsèques à l'élève et à l'établissement scolaire, où souvent Z_{ij} se compose d'une constante et de certains éléments de X_{ij} . Ici, $m(\cdot)$ représente l'effet fixe, et $v(\cdot)$ représente l'effet aléatoire.

Les résidus ε_{ij} sont supposés être indépendants avec une moyenne nulle et une variance conditionnelle $\sigma^2 X_{ij}$. Les $v_i(Z_{ij})$ peuvent être considérés comme des réalisations d'un processus lisse de moyenne nulle avec une fonction de covariance $\gamma(Z_{ij1}, Z_{ij2}) = E[v_i(Z_{ij1})v_i(Z_{ij2})]$. Les variables $v_i(Z_{ij})$ et ε_{ij} sont supposées être indépendantes conditionnellement à X_{ij} .

Comme le suggère Wenceslao González Manteiga (2012), les modèles multiniveaux et les modèles des données longitudinales reposent sur des principes de spécification et des stratégies d'estimation et d'identification similaires. En effet, on suppose que les observations des différents groupes de notre échantillon sont indépendantes, tandis que les observations du même groupe sont corrélées par nature. Cette dépendance intra-groupe est souvent modélisée par $v_i(Z_{ij}) = v_i(t_{ij})$, interprétée comme les effets non observables.

Pour les données regroupées, une modélisation à plusieurs niveaux prend en considération les observations de q groupes (écoles, classes), où Y_{ij} varie d'un groupe à un autre et aussi bien au sein de chaque groupe, et avec une matrice de régresseurs X_{ij} ($j = 1, \dots, n_i, i = 1, \dots, q$).

³⁵Le modèle (1.1) est sous une spécification largement étudiée dans l'économétrie des données de panel et des données regroupées. Plus précisément, il s'agit des modèles à effets fixes, des modèles à effets aléatoires, et des modèles à effets aléatoires corrélés.

Les observations des différents groupes sont supposées être indépendantes, alors que celles appartenant au même groupe sont supposées être corrélées à divers degrés. Cette corrélation intra-groupe est alors modélisée par un terme aléatoire $v_i(Z_{ij})$, qui est souvent appelé dans la littérature « latent effects ».

Un autre cas particulier de modèle est le modèle de régression à erreur imbriquée, où $m(\cdot)$ est linéaire (ou polynomial) et v_i est simplement une fonction indicatrice pour l'appartenance au groupe i ($i = 1, \dots, q$).

Comme nous nous concentrerons sur les formes non-paramétriques de $m(\cdot)$ et $v_i(\cdot)$ basées sur des polynômes locaux, il est également utile de mentionner le modèle de régression à coefficient aléatoire dont la formulation est la suivante :

$$y_{ij} = \beta X_{ij} + b_i X_{ij} + \varepsilon_{ij} \quad (12)$$

Où β et b_i sont respectivement les effets fixes et les coefficients d'effets aléatoires. Contrairement au modèle (12), la souplesse du modèle (11) offre également de nouvelles perspectives pour les différents problèmes rencontrés lors de la validation des hypothèses de la modélisation. Autrement dit, la modélisation par des formes non-paramétriques et semi-paramétriques peut être une alternative pertinente pour éviter les problèmes de spécification du modèle proposé, et pour la validation des tests et des hypothèses imposées sur les moments d'ordre 2 à savoir la moyenne et la variance.

Notre contribution à la littérature empirique est présentée comme suit. Premièrement, nous examinerons les approches existantes pour estimer le modèle (11) par une estimation polynomiale locale. Cependant, la question principale est de savoir comment choisir les poids, c'est-à-dire comment combiner la structure des corrélations avec les poids du noyau dans les équations à estimer. De ce fait, nous considérerons principalement trois stratégies différentes pour l'estimation de ce modèle ainsi que les propriétés d'inférences statistiques appropriées à ces méthodes.

En fait, nous aurons recours à une procédure de bootstrap itérative afin d'aboutir à un cadre d'inférence adéquat à notre proposition, de corriger les problèmes d'hétéroscédasticité et pour construire des intervalles de confiance. Par la suite, nous utiliserons l'estimation par bootstrap de l'erreur quadratique moyenne afin de remédier au problème de l'estimation du modèle à effet fixe.

Cependant, dans les modèles multiniveaux non-paramétriques et semi-paramétriques, le choix des paramètres de lissage s'avère être une tâche complexe comparativement aux estimations des modèles non et semi-paramétriques classiques (modèles en coupe instantanée et modèles de séries temporelles). Les méthodes non-paramétriques et semi-paramétriques standards ne sont pas appropriées car elles ne prennent pas en compte l'hétérogénéité individuelle, puisque la structure des données s'inscrit dans le cadre de données hiérarchiques.

Ainsi, en nous inspirant de la méthode de M.J. Lombardía, S. Sperlich (2008), nous utiliserons les estimations des moyennes quadratiques des résidus pour construire les intervalles de confiance et de prévisions. Nous traiterons principalement la stratégie d'estimation de l'inférence de la forme fonctionnelle $m(\cdot)$, et aussi nous discuterons les effets mixtes ou bien les fonctions individuelles représentées par la forme $\eta_i(X, Z) = m(X) + v_i(Z)$ pour analyser les paramètres tels que $\Theta_i = \hat{m}(X_i) + \hat{v}_i(Z_i)$ avec $\hat{m}(X_i) = \sum_{j=1}^{n_i} m(X_{ij}) / n_i$ et $\hat{v}_i(Z_i) = \sum_{j=1}^{n_i} v_i(Z_{ij}) / n_i$. Ces méthodes et leurs propriétés nous fournissent des outils puissants pour traiter nos données pour l'estimation, l'inférence, la construction des intervalles de confiance ainsi que la détermination de la trajectoire de l'input (taille de la classe) par rapport à l'output (score des élèves) dans le cadre de la structure hiérarchique de nos données.

L'objectif est ainsi de proposer un modèle non-paramétrique sous une spécification hiérarchique avec un cas particulier d'un modèle semi-paramétrique où les effets aléatoires sont supposés linéaires. Dans ce contexte, nous étudions les différentes propositions d'estimation non et semi-paramétriques, en les comparant en fonction de leurs performances prédictives.

Ces stratégies d'estimation et de validation des tests d'inférences statistiques nous permettront de répondre à notre problématique du seuil optimal de la taille des classes, à partir duquel le rendement scolaire se détériore de façon significative. L'estimation des matrices des variances-covariances, l'estimation conjointe des effets fixes et des effets aléatoires sont tous des éléments indispensables pour mener à bien notre méthodologie sans biais de spécification et sans biais d'identification. Notre méthodologie commence par le développement d'un modèle local linéaire sous une spécification hiérarchique. Nous présentons ensuite les techniques et étapes à suivre pour estimer un modèle à effet aléatoire dans le cadre des modèles semi-paramétriques.

4.1. Estimation marginale dans les modèles non-paramétriques sous une spécification hiérarchique

4.1.1. Un modèle local linéaire sous une spécification hiérarchique

En supposant que les fonctions m et v_i ont $(p + 1)$ dérivées continues (si on simplifie $p = 1$) pour tout x fixé dans un domaine générique X , m peut être approché par une fonction linéaire dans un voisinage de x par le théorème connu de Taylor comme suit :

$$m(X_{ij}) \approx m(X) + (X_{ij} - X)^t \nabla m(X) \quad (13)$$

Où ∇m désigne le gradient de la fonction m .

De même, pour la fonction aléatoire v_i , nous avons :

$$v_i(z_{ij}) \approx v_i(z) + (z_{ij} - z)^t \nabla v_i(z) \quad (14)$$

Pour tout z fixe dans un domaine générique Z , il vient que :

$$X_{ij} = (1, (X_{ij} - X)^t), \beta = (m(X), \nabla m(X)^t)^t, \text{ et } Z_{ij} = (1, (z_{ij} - z)^t) \quad (15)$$

$$b_i = (v_i(z), \nabla v_i(z)^t)^t$$

Le modèle (11) peut ainsi être approché par le modèle linéaire à effets mixtes suivant :

$$y_{ij} = X_{ij}\beta + Z_{ij}b_i + \varepsilon_{ij} \quad j = 1, \dots, n_i, i = 1, \dots, q \quad (16)$$

L'hypothèse la plus cruciale est que les effets aléatoires b_i sont indépendants de X_{ij} ($j = 1, \dots, n_i, i = 1, \dots, q$) avec une moyenne nulle et $E[b_i b_i^t] = B_i$. Les résidus ε_{ij} sont mutuellement indépendants et indépendantes de b_i , avec une moyenne conditionnelle nulle $E[\varepsilon_{ij} | X_{ij}] = 0$ et des variances conditionnelles finies $\sigma^2 X_{ij}$. En empilant les observations dans les q groupes, le modèle s'exprime comme suit :

$$y_i = X_i \beta + Z_i b_i + \varepsilon_i \quad i = 1, \dots, q \quad (17)$$

Où $y_i = (y_{i1}, \dots, y_{ini})^t$, X_i est une matrice $(n_i \times 2)$ avec des lignes $X_{ij} = (1, (x_{ij} - x)^t)$

Z_i est une matrice avec des lignes $Z_{ij} = (1, (z_{ij} - z)^t)$ ($j = 1, \dots, n_i, i = 1, \dots, q$) et

$\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{ini})^t$ est un vecteur des résidus avec une matrice de covariance conditionnelle qui vaut :

$$\sum_i = \text{diag}(\sigma^2(X_{i1}), \dots, \sigma^2(X_{ini}))$$

Le modèle (17) peut aussi être formulé de la façon suivante :

$$y = X\beta + Zb + \varepsilon \quad (18)$$

Avec $y = (y_1^t, \dots, y_q^t)^t$, $X(n \times (k + 1))$ une matrice avec des lignes X_i et

$Z(n \times q(r + 1))$ une matrice avec des blocs diagonaux Z_i

Le coefficient des effets aléatoires est donné par $b = (b_1^t, \dots, b_q^t)^t$ et le vecteur des erreurs par $\varepsilon = (\varepsilon_1, \dots, \varepsilon_q)^t$.

Pour l'estimation d'un tel modèle local à effets mixtes linéaires, commençons par une définition des effets fixes β et du modèle marginal suivant :

$$y = X\beta + u \quad (19)$$

Avec $u = Zb + \varepsilon$, impliquant une corrélation intra-groupe exprimée par

$$E[uu^t | X, Z] = V = ZBZ^t + \Sigma$$

Ici B est une matrice symétrique de dimension $n(r + 1) \times n(r + 1)$ définie par les blocs diagonaux β_i .

Lin et Carroll (2000) proposent une extension formelle de l'approche paramétrique d'équations d'estimation généralisée introduisant des poids du noyau pour les données regroupées ou hiérarchisées. La méthodologie proposée par les auteurs repose sur la quasi vraisemblance pour l'inférence du paramètre à effet fixe β , donnée par

$$0 = \sum_{i=1}^q X_i^t (V_i^c)^{-1} (y_i - X_i\beta) \quad (20)$$

Avec V_i^c qui ne correspond pas nécessairement la matrice de corrélation intra-groupe $Z_i B_i Z_i^t + \sum_i$, mais plutôt à une matrice exploratoire. En effet, $V_i^c = S_i^{\frac{1}{2}} R_i(\delta) S_i^{\frac{1}{2}}$ avec $S_i^{\frac{1}{2}}$ est une matrice diagonale dépendante des éléments diagonaux de V_i , et $R_i(\delta)$ est une matrice de corrélation inversible.

La solution peut être obtenue par des moindres carrés avec des repondérations itératives, et dont l'estimateur résultant est asymptotiquement gaussien.

Si le modèle paramétrique (linéaire) est seulement considéré localement, il est nécessaire d'introduire des poids-noyaux pour tenir compte uniquement des observations dans un voisinage restreint. Lin et Carroll (2000) suggèrent deux façons différentes d'introduire les poids, à savoir

$$0 = \sum_{i=1}^q X_i^t (V_i^c)^{-1} W_{ih} (y_i - X_i \beta) \quad (21)$$

$$\text{Et } 0 = \sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} (V_i^c)^{-1} W_{ih}^{\frac{1}{2}} (y_i - X_i \beta) \quad (22)$$

Où $W_{ih} = \text{diag}(K_h(X_{i1} - X), \dots, K_h(X_{ini} - X))$, avec $K_h(\cdot) = h^{-1}K(\cdot/h)$ une fonction de noyau (multi-variée) et $h > 0$ est un paramètre bandwidth.

Il est à noter que ces deux approximations sont différentes si V_i^c n'est pas diagonal. Les auteurs recommandent d'ignorer complètement la corrélation intra-groupe puisque les estimateurs de (21) ou (22) atteignent asymptotiquement la variance minimale avec $R_i(\delta) = I$.

Li and Wang (2003) explique pourquoi asymptotiquement ceci est encore optimal et donc différent de l'approche paramétrique des estimations des équations généralisées.³⁶

Par conséquent, pour des améliorations de l'efficacité asymptotique, des méthodes alternatives de pondération sont nécessaires pour une mesure correcte de la corrélation intra-groupe. L'idée de régression généralisée (ou pondérée) des moindres carrées (GLS) repose sur une transformation de l'expression (19) de telle sorte que l'on obtient des observations non corrélées. Vilar et Francisco (2002) ont appliqué cette méthodologie à un estimateur polynomial

³⁶Asymptotiquement (c'est-à-dire quand $h \rightarrow 0$) il n'y a effectivement qu'une seule observation (i, j) pour chaque $i = 1, \dots, q$ qui contribue à estimer $m(x)$. Cette observation unique est pondérée par $K_h(X_{ij} - X) v_i^{jj}$ avec v_i^{jj} étant l'élément (j, j) de V^{-1}

local pour une spécification temporelle AR (1). En suivant l'approche de Vilar Francisco (2002), mais que nous adaptons au modèle sous une spécification hiérarchique, comme V est une matrice symétrique et définie positive, et que son inverse a une racine carrée $V^{-1/2}$ qui satisfait $V^{-1} = V^{-1/2}(V^{-1/2})^t$, il vient que :

$$\tilde{y} = V^{-1/2}y = V^{-1/2}X\beta + V^{-1/2}u = \tilde{X}\beta + \tilde{u} \quad (23)$$

est vérifié lorsque $E[\tilde{u}\tilde{u}^t|\tilde{X}] = V^{-1/2}V(V^{-1/2})^t = I$, c'est-à-dire lorsque les des résidus sont non corrélés. Par la suite, les poids du noyau sont introduits afin d'obtenir la régression pondérée des moindres carrés suivants :

$$\tilde{\tilde{y}} = W_h^{1/2}\tilde{y} = W_h^{1/2}V^{-1/2}y = W_h^{1/2}V^{-1/2}X\beta + W_h^{1/2}V^{-1/2}u = \tilde{\tilde{X}}\beta + \tilde{\tilde{u}} \quad (24)$$

Avec $W_h = \text{diag}(W_{ih})$. La régression pondérée des moindres carrés conduit à l'estimateur linéaire local marginal pour les effets fixes suivant :

$$\hat{\beta}_{M1} = \left(\sum_{i=1}^q X_i^t V_i^{-\frac{1}{2}} W_{ih} V_i^{-\frac{1}{2}} X_i \right)^{-1} \left(\sum_{i=1}^q X_i^t V_i^{-\frac{1}{2}} W_{ih} V_i^{-\frac{1}{2}} y_i \right) \quad (25)$$

Nous obtenons ainsi notre premier estimateur linéaire local marginal :

$$\hat{m}_{M1}(X) = e_1^t \hat{\beta}_{M1} = \sum_{i=1}^q w_i^{M1}(X) y_i \quad (26)$$

dont l'expression des poids est la suivante :

$$w_i^{M1}(X) = e_1^t \left(\sum_{l=1}^q X_l^t V_l^{-\frac{1}{2}} W_{lh} V_l^{-\frac{1}{2}} X_l \right)^{-1} X_i^t V_i^{-\frac{1}{2}} W_{ih} V_i^{-\frac{1}{2}} \quad (27)$$

Linton et al. (2003) ont souligné qu'une telle transformation n'améliore pas les propriétés asymptotiques du premier ordre, mais qu'elle améliore l'estimation dans le cas d'échantillons finis. Les auteurs proposent un estimateur en deux étapes, en calculant d'abord l'estimateur proposé par Lin et Carroll (2000), puis utilisent les résultats pour construire une transformation linéaire dey qui présente une matrice de covariance diagonale afin d'exécuter une régression linéaire locale sur les données transformées. De manière alternative aux moindres carrés pondérés, Chen et Jin (2005) proposent d'utiliser une fonction de pondération basée sur les

variances locales plutôt que sur les variances globales. Cela se rapproche de l'idée des estimateurs des moindres carrés généralisés et conduit à l'estimateur linéaire local marginal suivant :

$$W_h^{1/2}y = W_h^{1/2}X\beta + W_h^{1/2}u \quad (28)$$

Cet estimateur peut être transformé pour obtenir des observations non corrélées en calculant :

$$\Omega_h^{-1/2}W_h^{1/2}y = \Omega_h^{-1/2}W_h^{1/2}X\beta + \Omega_h^{-1/2}W_h^{1/2}u \quad (29)$$

Avec $\Omega_h = E \left[W_h^{\frac{1}{2}}u(W_h^{\frac{1}{2}}u)^t \right] = W_h^{\frac{1}{2}}VW_h^{\frac{1}{2}}$. où $\Omega_h^{-1/2}$ est l'inverse généralisé de Moore-Penrose de $\Omega_h^{1/2}$.

Enfin, nous introduisons les poids du noyau par l'expression suivante :

$$\tilde{y} = W_h^{\frac{1}{2}}\Omega_h^{-1/2}W_h^{\frac{1}{2}}y = W_h^{\frac{1}{2}}\Omega_h^{-1/2}W_h^{\frac{1}{2}}X\beta + W_h^{\frac{1}{2}}\Omega_h^{-1/2}W_h^{\frac{1}{2}}u = \tilde{X}\beta + \tilde{u} \quad (30)$$

Il s'ensuit que l'estimateur linéaire local marginal vaut :

$$\tilde{m}_{M2}(X) = e_1^t \tilde{\beta}_{M2} = \sum_{i=1}^q w_i^{M2}(X)y_i \quad (31)$$

$$\tilde{\beta}_{M2} = \left(\sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} \Omega_{ih}^{-\frac{1}{2}} W_{ih}^{-\frac{1}{2}} W_{ih}^{\frac{1}{2}} X_i \right)^{-1} \sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} \Omega_{ih}^{-\frac{1}{2}} W_{ih}^{-\frac{1}{2}} W_{ih}^{\frac{1}{2}} y_i \quad (32)$$

Avec $\Omega_{ih} = W_{ih}^{\frac{1}{2}} V_i W_{ih}$

Et les poids $w_i^{M2}(X) = e_1^t \left(\sum_{l=1}^q X_l^t W_{lh}^{1/2} \Omega_{lh}^{-\frac{1}{2}} W_{lh}^{-\frac{1}{2}} W_{lh}^{1/2} X_l \right)^{-1} X_i^t W_{ih}^{1/2} \Omega_{ih}^{-\frac{1}{2}} W_{ih}^{-\frac{1}{2}} W_{ih}^{1/2}$

Les propriétés asymptotiques peuvent être obtenues de la même manière que celles proposées par Vilar et Francisco (2002).

La troisième approche est une adaptation de l'approche développée par Park et al (2006). Elle est plus facile à appliquer pour notre modélisation hiérarchique. Cette approche permet

d'obtenir un troisième estimateur marginal basé sur le maximum de vraisemblance qui est le suivant :

$$\mathcal{L}_M(\beta, y) = -\frac{1}{2} \sum_{i=1}^q \left\{ [y_i - X_i \beta]^t W_{ih}^{\frac{1}{2}} V_i^{-1} W_{ih}^{\frac{1}{2}} [y_i - X_i \beta] + C_{iM} \right\} \quad (33)$$

Avec une constante $C_{iM} = \log|V_i| + 2n_i \log(2\pi)$. L'estimation est donnée par :

$$\hat{m}_{M3}(X) = e_1^t \hat{\beta}_{M3} = \sum_{i=1}^q w_i^{M3}(X) y_i \quad (34)$$

Avec $\hat{\beta}_{M3} = (\sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} V_i^{-1} W_{ih}^{\frac{1}{2}} X_i)^{-1} \sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} V_i^{-1} W_{ih}^{\frac{1}{2}} y_i$

Et le poids $w_i^{M3}(X) = (\sum_{i=1}^q X_i^t W_{ih}^{\frac{1}{2}} V_i^{-1} W_{ih}^{\frac{1}{2}} X_i)^{-1} X_i^t W_{ih}^{\frac{1}{2}} V_i^{-1} W_{ih}^{\frac{1}{2}}$

A partir d'un raisonnement semblable afin d'obtenir des estimateurs transformés et simplifiés de (30) et (33), nous considérons le modèle transformé suivant :

$$\tilde{y} = V_i^{-\frac{1}{2}} W_h^{\frac{1}{2}} y = V_i^{-\frac{1}{2}} W_h^{\frac{1}{2}} X \beta + V_i^{-\frac{1}{2}} W_h^{\frac{1}{2}} u = \tilde{X} \beta + \tilde{u} \quad (35)$$

Cependant, cette transformation ne conduit pas à un modèle non corrélé car,

$$E(\tilde{u} \tilde{u}^t | \tilde{X}) = V^{-\frac{1}{2}} W_h^{\frac{1}{2}} V W_h^{\frac{1}{2}} (V^{-\frac{1}{2}})^t \neq W_h I \quad (36)$$

Park et Wu (2006) affirment que, indépendamment des résultats asymptotiques, cet estimateur révèle des résultats beaucoup plus efficaces que l'estimateur développé par Chen et Jin (2005).

4.1.2. Un modèle semi-paramétrique : Estimation du modèle à effets aléatoires

Il s'agit ici de considérer le modèle semi-paramétrique avec des effets aléatoires linéaires qui est un cas particulier des modèles non-paramétriques. Il s'agit du modèle suivant :

$$y_{ij} = m(X_{ij}) + b_i Z_{ij} + \varepsilon_{ij}, \quad j = 1, \dots, n_i, \quad i = 1, \dots, q \quad (37)$$

Où b_i indique les effets aléatoires. Ce modèle semi-paramétrique a également été proposé par H. Wu, J.T. Zhang (2006), mais sans aucune étude théorique ou illustration empirique. Pour estimer le modèle (37), nous pouvons utiliser chacun des 3 estimateurs marginaux que nous venons de présenter. En ce qui concerne la prévision, et en particulier la détermination des intervalles de confiance, il est primordial de déterminer les effets aléatoires.

Bien que cela soit possible dans un modèle purement non-paramétrique, la tâche s'avère généralement difficile à mener d'abord pour l'interprétation, la simulation et la détermination du processus stochastique qui génère $v_i(Z_{ij})$. En effet alors que la plupart des articles de la littérature ne considèrent même pas l'analyse des prédictions pour une fonction aléatoire non-paramétrique v_i , d'autres ne discutent que de sa difficulté.³⁷. En revanche, pour un modèle d'effets mixtes semi-paramétriques, la question devient plus intéressante et innovante dans le sens où les décompositions sont multiples. Notons par $\hat{m}_M(X)$ l'un des estimateurs marginaux non-paramétriques des effets fixes. La composante des effets aléatoires peut être prédite à partir de l'expression suivante :

$$\hat{b}^{sm} = (\hat{b}_1^{sm}, \dots, \hat{b}_q^{sm})^t = B\tilde{Z}^t V^{-1}(y - \hat{m}_M) \quad (38)$$

Où \tilde{Z} est la matrice $(n \times r)$ avec les lignes (z_{i1}, \dots, z_{ini}) , $i = 1, \dots, q$, et \hat{m}_M est un vecteur. En supposant que (b, y) suit une distribution normale multivariée, ce dernier couple est le meilleur estimateur linéaire pour b . Lorsque l'estimateur marginal est dérivé des moindres carrés linéaires généralisés, l'estimateur résultant est le meilleur estimateur linéaire sans biais. Les courbes moyennes individuelles peuvent être estimées par :

$$\hat{\eta}_i(X, Z) = \hat{m}_M(X) + \hat{b}_i^{sm} Z, \text{ et la moyenne est donnée par } \hat{\Theta}_i = \sum_{j=1}^{n_i} (\hat{m}_M(X_{ij}) + \hat{b}_i^{sm} Z_{ij}) / n_i .$$

Les estimateurs et les prédicteurs dépendent des matrices de variance B et Σ .

4.1.3. Estimation de l'erreur quadratique moyenne avec bootstrap

Le critère d'efficacité habituel pour tout estimateur d'une fonction $m(x)$ est l'erreur de prédiction, habituellement mesurée par l'erreur quadratique conditionnelle (MSE) à chaque point d'estimation x :

$$MSE(\hat{m}_h(X)) = E[(\hat{m}_h(X) - m(X))^2] = (\text{bias}(\hat{m}_h(X)))^2 + \text{var}(\hat{m}_h(X)) \quad (39)$$

³⁷Dans les différents articles, nous avons constaté des différences dans les prédictions obtenues relevant de la même logique d'identification

Ou globalement par sa version intégrée :

$$\text{MISE}(\widehat{m}_h) = \int \text{MSE}(\widehat{m}_h(X)) dx$$

Dans le modèle multiniveaux non-paramétrique, nous nous intéressons à la quantification de l'erreur de l'estimateur non-paramétrique $m(x)$ et des deux types des fonctions aléatoires: les courbes individuelles η_i et les paramètres moyens $\widehat{\Theta}_i, i = 1, \dots, q$. Pour ces deux derniers paramètres, il s'agit de contrôler l'erreur quadratique moyenne de prédiction. Dans la littérature actuelle sur les modèles à effets mixtes, les méthodes de ré-échantillonnage deviennent de plus en plus utilisées pour obtenir des estimateurs contrôlés sans biais.

Les méthodes courantes sont multiples comme le bootstrap paramétrique ou le bootstrap par couplage. Lombardía et Sperlich (2008) ont utilisé une combinaison de ces deux derniers types de bootstrap. Le bootstrap peut également être utilisé pour construire des intervalles de confiance locaux. Le principal inconvénient de l'inférence basée sur le ré-échantillonnage réside dans sa longue procédure de compilation. Nous verrons maintenant comment cela peut être évité. L'algorithme bootstrap est formulé pour des matrices de variance données. Pour sa mise en œuvre pratique, les matrices de variance peuvent être estimées en utilisant la méthode de la vraisemblance maximale.

Considérons maintenant le schéma de ré-échantillonnage suivant.

Étape 1. Prenons une estimation pilote $\widehat{m}_{h_0}(X)$ avec une bande passante (Bandwidth) pilote h_0 .

Étape 2. Générer $v_i^* Z_{ij}$ à partir du processus aléatoire de moyenne nulle avec une structure de covariance connue $\gamma(Z_{ij1}, Z_{ij2})$ conditionnée par les régresseurs originales $Z_{ij}, i = 1, \dots, q; j = 1, \dots, n_i$.

Étape 3. Générer $\varepsilon_{ij}^* = \sigma(X_{ij})W_{ij}, j = 1, \dots, n_i; i = 1, \dots, q$, de W_{ij} étant des variables aléatoires indépendantes et identiquement distribuées (i.i.d.) avec $E[W_{ij}] = 0$, et $E[W_{ij}^2] = 1$ indépendantes de $v_i^* Z_{ij}$.

Étape 4. Construire le modèle bootstrap $y_{ij}^* = \widehat{m}_{h_0}(X_{ij}) + v_i^* Z_{ij} + \varepsilon_{ij}^*$, à partir de l'échantillon bootstrap $\{(y_{ij}^*, X_{ij}, Z_{ij}; j = 1, \dots, n_i; i = 1, \dots, q)\}$, et calculez l'estimateur bootstrap $\widehat{m}_h^*(X)$

Une estimation de bootstrap du MSE $\hat{m}_h(X)$ est donnée par

$$\text{MSE}_*(\hat{m}_h(X)) = E_*[(\hat{m}_h^*(X) - \hat{m}_{h0}^*(X))^2] \quad (40)$$

Où E_* désigne l'espérance sur les distributions de ré-échantillonnage. La simplicité du mécanisme du bootstrap permet de calculer les prédictions exactes de (40). Plus précisément, puisque tout estimateur marginal proposé peut être écrit sous la forme d'une combinaison linéaire des réponses indépendamment de l'expression $\hat{m}_h(X) = \sum_{i=1}^q w_i X y_i$, nous avons :

$$\begin{aligned} \text{MSE}_*(\hat{m}_h(X)) &= \{B_*(h; x)\}^2 + V_*(h; x) \\ B_*(h; x) &= \sum_{i=1}^q w_i(X) \hat{m}_{i,h0} - \hat{m}_{h0}(X) \\ V_*(h; x) &= \sum_{i=1}^q w_i(X) V_i w_i^t(X) \end{aligned} \quad (41)$$

En utilisant $\text{Var}_*(y_i^*) = B_i + \sum i = V_i$.

Il n'est pas nécessaire d'utiliser des simulations de Monte-Carlo pour calculer $\text{MSE}_*(\hat{m}_h(X))$. Pour les erreurs moyennes de prédiction de $\hat{\eta}_i$ ou $\hat{\Theta}_i$, nous n'admettons pas une telle version simplifiée. En effet, pour chaque échantillon du bootstrap, l'erreur de prédiction contient les termes $\hat{v}_i^* - \hat{v}_i$, ce qui n'est pas facile à obtenir sans effectuer une simulation Monte-Carlo. Dans ces cas, nous devons ajouter l'étape suivante.

Étape 5. Répéter la procédure pour $l = 1, \dots, R$ et notée $\hat{m}_h^{*(l)}(X)$ les estimations bootstrap calculées pour la l -ième échantillon bootstrap.

L'approximation de Monte-Carlo du $\text{MSE}_*(\hat{\eta}_{i,h}(X, Z))$ est donnée par :

$$\text{MSE}_*(\hat{\eta}_{i,h}(X, Z)) = R^{-1} \sum_{l=1}^R (\hat{\eta}_{i,h}^{*(l)}(X, Z) - \hat{\eta}_{i,h0}(X, Z))^2 \quad (42)$$

Et de manière analogue pour $\hat{\Theta}_i$.

Dans le cas des variances inconnues, il faut modifier l'étape 1 par la formulation suivante :

Étape 1. Estimer $\gamma(\cdot, \cdot)$ (ou \widehat{B}_i) et $\widehat{\Sigma}_i$. Prenons une estimation pilote $\widehat{m}_{h_0}(X)$ impliquant ces estimations de covariance, avec une bande passante pilote h_0 .

De plus, puisque dans le modèle semi-paramétrique, la partie non-paramétrique \widehat{m} domine les expressions asymptotiques (en raison de son taux de convergence plus lent), il suffit d'étudier les propriétés asymptotiques du MSE de \widehat{m}_h et la cohérence de la méthode de sélection de bandes passante (bandwidth).

En supposant certaines conditions de régularité, l'approximation bootstrap $MSE_*(\widehat{m}_h(X))$ de (40) est cohérente à l'intérieur du support de f , étant f la densité de x . En probabilité, nous avons :

$$MSE_*(\widehat{m}_h(X)) - MSE(\widehat{m}_h(X)) \rightarrow 0 \quad (43)$$

Pour des raisons de présentation, nous nous limitons au modèle semi-paramétrique, c'est-à-dire que nous supposons maintenant une composante d'effets aléatoires linéaires. Ensuite, dans l'étape 2 il s'agit de générer des effets aléatoires b_i^* , $i = 1, \dots, q$ de $B_i^{1/2}W$, où W est un vecteur aléatoire avec moyenne nulle et matrice de covariance égale à l'identité. Pour l'étape 4, nous avons alors :

$$v_i^*(Z_{ij}) = Z_{ij}' b_i^*.$$

Les intervalles de confiance $\frac{1-\alpha}{2}\%$ pour $m(x)$ sur la base des estimations $\widehat{m}_h(X)$, V_* et la normalité asymptotique de nos estimations, sont donnés par l'intervalle :

$$\left[\widehat{m}_h(X) - \frac{z_{1-\alpha/2}}{2} V_*^{1/2}(X); \widehat{m}_h(X) + \frac{z_{1-\alpha/2}}{2} V_*^{1/2}(X) \right] \quad (44)$$

Il faut noter que dans les modèles d'effets mixtes paramétriques, les intervalles de prédiction bootstrap ne s'améliorent que marginalement par rapport aux approximations linéaires typiquement utilisées dans les modèles mixtes non et semi-paramétriques.

4.1.4. Sélection de bande passante locale optimale « Bandwidth »

La méthode du bootstrap peut être une alternative de toutes les méthodes connues dans la littérature pour trouver localement les paramètres de lissage optimal.

La prise en compte d'un terme aléatoire dans le modèle de régression modifie également la valeur optimale standard pour les modèles à effets fixes. Pour les données longitudinales ou hiérarchiques, la largeur de la bande asymptotiquement optimale (globale) a été dérivée par Lin et Carroll (2000), Wu et Zhang (2006), Park et Wu (2006) et Chen et Jin (2005).

Dans tous les cas, la largeur de bande optimale a été calculée pour le cas où les matrices de variance réelle étaient connues et ensuite ajustées à la situation commune de variances inconnues. Park et Wu (2006) ont proposé une procédure de backfitting avec des algorithmes itératifs imbriqués, avec la mise à jour de toutes les composantes du modèle, des variances et des bandes passantes sur la base d'une technique de validation croisée. Zhu et al. (2009) ont proposé une méthode de validation croisée généralisée pour estimer simultanément la bande passante et les variances.

Notre proposition réside dans l'ajustement dû à la nouvelle composante à effet aléatoire. Nous proposons une sélection de paramètres de lissage basée sur la méthode bootstrap qui recherche les largeurs de bande localement optimales

$$h_{\text{opt}}(X) = \arg \min_h \text{MSE}(\hat{m}_h(X)) \quad (45)$$

En utilisant l'estimation MSE bootstrap, le sélecteur de bande passante local pour l'estimateur marginal $\hat{m}_h(\cdot)$ est par conséquent donné par

$$h_b(X) = \arg \min_h \text{MSE}_*(\hat{m}_h(X)) \quad (46)$$

La consistance de la procédure de sélection découle de celle de l'approximation bootstrap du MSE.

Pour résumer la présentation qui précède, notre objectif est d'estimer un modèle semi paramétrique ou la partie non paramétrique est la taille de la classe, tout en contrôlant notre modèle par l'ensemble des variables intrinsèques à l'élève et des variables contextuelles (partie paramétrique). La stratégie d'estimation que nous venons de présenter va nous permettre d'estimer les paramètres du modèle, de générer des prédictions à travers nos simulations, de calculer des intervalles de confiances afin d'identifier la taille optimale de la classe à dans les différents contextes : milieu géographique, âge de l'enseignant, climat scolaire, etc.

5. Résultats de la modélisation

L'objectif de cette section est d'interpréter et d'analyser les résultats du modèle semi-paramétrique global et les modèles semi-paramétriques conditionnels au contexte d'apprentissage de l'élève. Nous commençons par le modèle semi-paramétrique expliquant le score dans les trois matières, à savoir : la lecture, les mathématiques et les sciences. Ainsi, l'objectif principal est d'identifier les tailles optimales des classes où le score des élèves décroît de façon très significative, et qui certainement explique la raison pour laquelle le score à l'échelle nationale est tiré vers le bas. Dans ce qui suit, nous présentons les résultats obtenus pour chacune des 3 matières.

5.1. Taille optimale de la classe en lecture

D'après les estimations obtenues des modèles semi-paramétriques, la taille de la classe optimale à partir de laquelle la performance scolaire en lecture décroît est de 22 à 23 élèves par classe. Cette taille optimale de classe est en outre confirmée par les modèles semi-paramétriques contextuels prenant en compte différents facteurs contextuels (localisation de l'établissement, ancienneté de l'enseignant, genre de l'enseignant, le niveau de mixité sociale) associés à l'environnement scolaire de l'élève dont il acquiert ses acquis scolaires (graphique 30).

L'estimation du modèle semi-paramétrique contextualisé par le problème d'absentéisme des enseignants révèle que la taille optimale varie entre 20 et 24 élèves par classe quel que soit le degré d'absentéisme des enseignants (Graphique 30.a).

De même, la prise en compte du genre de l'enseignant révèle que les enseignantes peuvent au maximum accueillir 20 élèves par classe alors que l'effectif maximum est de 28 élèves pour les enseignants du genre masculin (Graphique 30.b).

Au regard de l'âge de l'enseignant, il apparaît que les jeunes enseignants (25-29 ans) peuvent accueillir de 28 à 29 élèves par classe tout en maintenant une performance scolaire assez élevée. Autrement dit, la taille optimale de classes pour les jeunes enseignant (e)s est de 29 élèves. De même, l'effectif optimal diminue pour les enseignants âgés ; il est de 21 élèves pour les enseignants âgés et varie entre 21 à 26 élèves pour les plus âgés de 50-59 ans (Graphique 30.c).

Concernant la charge horaire, les enseignants ayant moins de charge horaire ont la capacité d'accueillir plus d'élèves de sorte que la taille des classes peut varier entre 20 et 33 élèves.

Cependant, les enseignants qui déclarent avoir un problème de charge horaire, ne peuvent accueillir que 22 à 23 élèves dans l'objectif d'atteindre une performance scolaire conforme aux objectifs des standards internationaux. Par ailleurs, les enseignants déclarant un grave problème de charge horaire ne peuvent accueillir que 20 élèves par classe (Graphique 30.d).

En réalisant une décomposition par le degré de mixité sociale, les enseignants assignés à des classes avec peu d'élèves issus de milieux défavorisés seraient en mesure d'accueillir 31 élèves par classe. En revanche, cette taille diminue avec le nombre d'élèves issus de milieux défavorisés. Ainsi, la taille optimale de la classe serait de 20 à 24 élèves pour une composition sociale mixte et de 20 à 22 élèves par classe pour une composition caractérisée par une part importante des élèves issus de milieux défavorisés (Graphique 30.e).

En agissant de l'expérience des enseignants, nos résultats montrent que la taille optimale de la classe qui assure une performance scolaire assez élevée est de 22 à 26 élèves quel que soit le nombre d'années d'expérience de l'enseignant (Graphique 30.f).

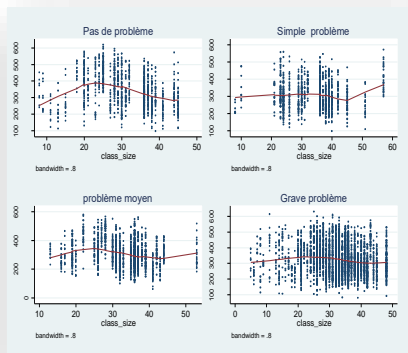
De plus, quel que soit le niveau d'instruction des parents, la taille des classes convenable pour une réussite scolaire de tous les élèves est de 20 à 25 élèves par classe. Autrement dit, la réduction de la taille des classes doit être généralisée quel que soit le niveau d'instruction des parents et leur niveau socio-économique ou leur catégorie socio professionnelle (Graphique 30.h).

D'après les résultats, un enseignant affecté dans le milieu urbain peut accueillir une classe de 21 à 23 élèves, tandis que les enseignants travaillant dans le milieu rural peuvent recevoir une taille de classe de 22 à 25 élèves (Graphique 30.i).

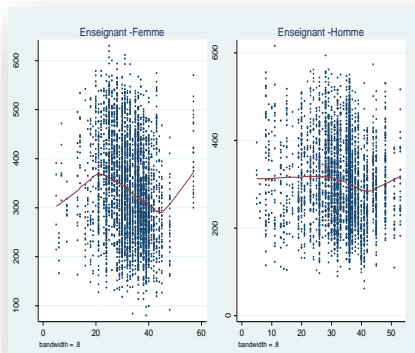
Enfin, nos résultats dévoilent que quelque le contexte ou l'élève est assigné la taille de la classe qui assure une réussite généralisée de l'ensemble des élèves en lecture varie entre 22 et 27 élèves.

Graphique 30 .Les courbes du modèle semi-paramétrique expliquant la performance scolaire en lecture

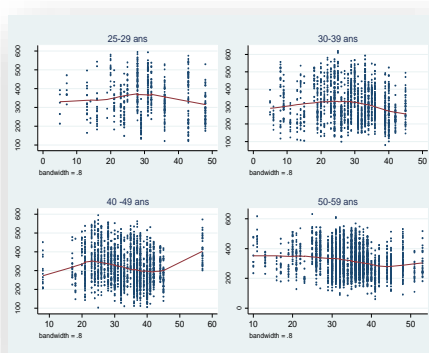
a. Absence des enseignants



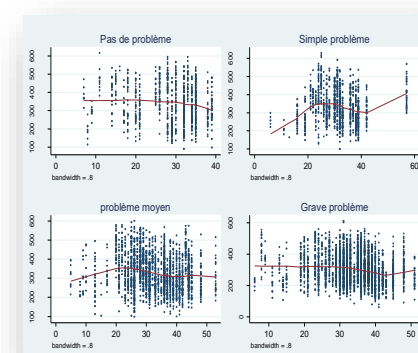
b. Genre de l'enseignant



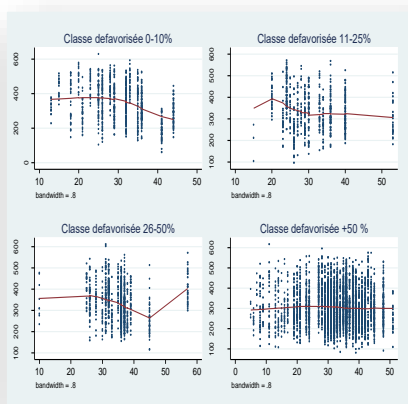
c. Age de l'enseignant



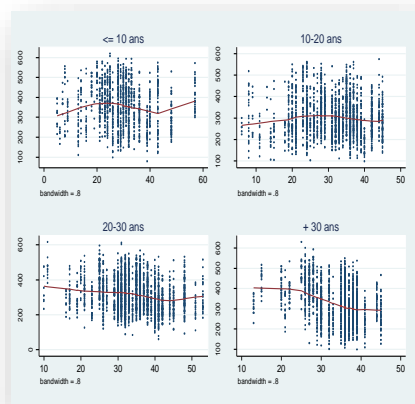
d. Charge horaire



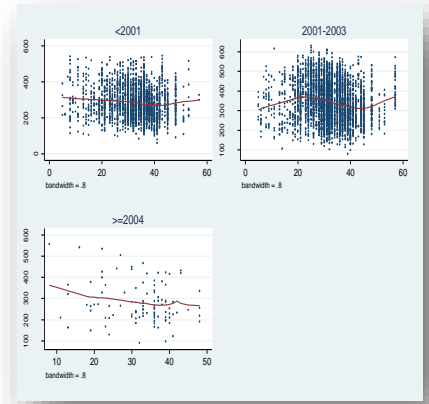
e. Degré de mixité sociale



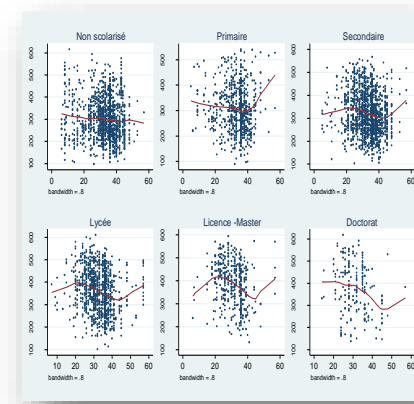
f. Expérience des enseignants



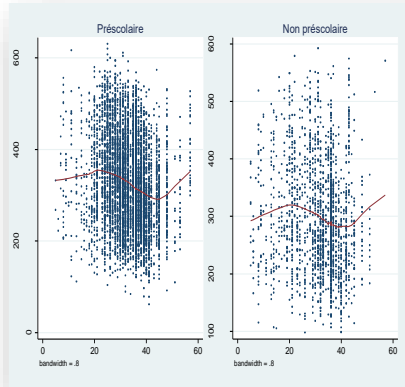
g. Age de l'élève



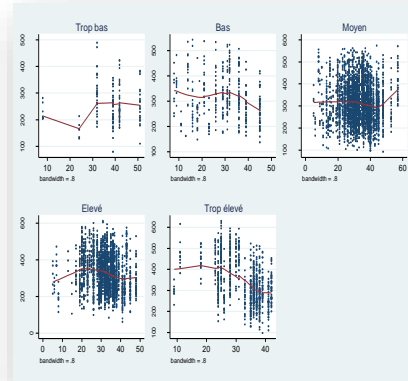
h. Niveau d'instruction des parents



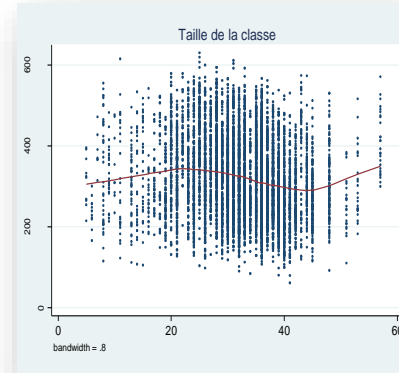
i. **Préscolarisation des élèves**



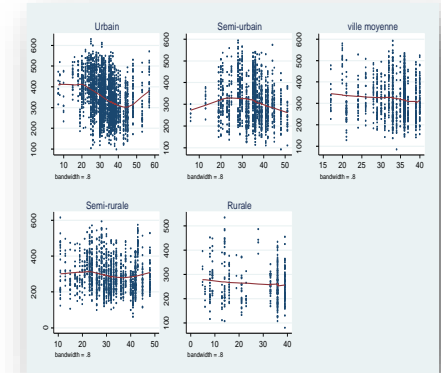
j. **Satisfaction des enseignants**



k. **Taille de la classe**



l. **Milieu de l'établissement**



Source : D'après l'auteur à partir des données PIRLS 2011

Encadré 2 : Etape d'estimation et interprétation des graphiques ci-dessus relatifs au lecture (Graphique 30)

1. Estimation du modèle MCO
2. Estimation du modèle paramétrique
3. Estimation du modèle semi paramétrique
4. Estimation du modèle semi paramétrique en multiniveaux (Tableau 38 en Annexe du chapitre 2)
5. Test d'Hausman (modèle semi paramétrique vs modèle semi-paramétrique en multiniveau (Tableau 38 en Annexe du chapitre 2))
6. **Simulation des courbes semi-paramétriques pour chacune des variables contextuelles (Graphique 30)**

5.2. Taille optimale de la classe en mathématiques

Les graphiques 31 présentent les simulations du score des élèves en mathématiques selon la taille des classes et l'ensemble des contextes où les élèves accumulent leurs acquis scolaires.

La taille optimale de la classe est plus faible pour les femmes (21 élèves) que pour les hommes (24 élèves) (Graphique 31.a).

Concernant le problème d'absentéisme des enseignants, nos résultats montrent que la taille optimale de la classe est de 23 à 26 élèves quel que soit le degré d'absentéisme des enseignants. En revanche, la présence d'un problème grave d'absentéisme réduit cette taille à 20 voire 22 élèves par classe (Graphique 31.b).

Par rapport au milieu de l'établissement, la taille optimale serait de 24 élèves dans le milieu urbain, et de 23 élèves dans le milieu rural (Graphique 31.c).

Pour les enseignants âgés de 25-29 ans, la taille optimale de la classe est de 26 élèves, tandis qu'elle est de 25 élèves et 23 élèves pour les enseignants âgés respectivement de 50-59 ans et de 30-39 ans (Graphique 31.d).

La formation continue des enseignants engendre une différence en termes de performance scolaire. Nos estimations révèlent que la taille optimale pour les enseignants ayant bénéficié d'une formation continue est de 22 élèves, tandis qu'elle est de 20 élèves pour les non bénéficiaires de la formation continue (Graphique 31.e).

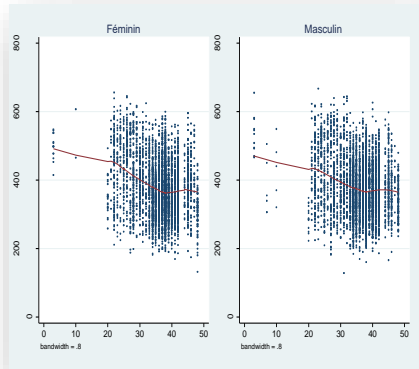
Les élèves assignés à des enseignants satisfaits de leur travail performant mieux que ceux dont les enseignants sont moins satisfaits de leur situation. Cependant, les résultats de notre modèle montrent que la taille optimale de classe varie de 21 à 24 élèves quelconques le degré de satisfaction des enseignants envers leur travail à l'école (Graphique 31.f).

Concernant l'influence du degré de la mixité sociale à l'établissement, les enseignants assignés à des classes avec une composition sociale d'élèves défavorisés peuvent accueillir au maximum 20 élèves. Cette taille est de 22 à 24 élèves pour une classe avec une composition sociale mixte, et de 21 à 23 élèves lorsqu'il s'agit d'une composition sociale des élèves favorisés (Graphique 31.g).

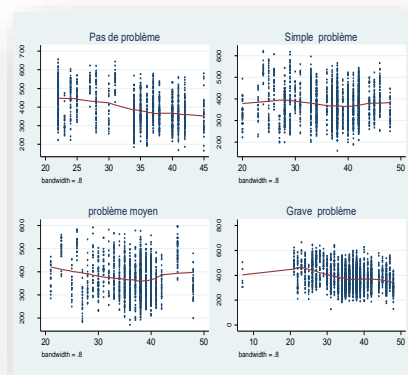
De plus, l'instruction des parents est un facteur déterminant de la réussite scolaire des élèves. En effet, les résultats de notre estimation montrent que la taille des classes optimale est de 21 élèves, quel que soit le niveau d'instruction des parents. Autrement dit, avoir des parents bien instruits ou peu instruits n'affecte pas la taille de la classe. Cela suggère que les mécanismes d'affectation des élèves et les politiques éducatives y afférentes ne doivent pas prendre en considération des critères de sélection (niveau d'instruction des parents) pour déterminer la taille de classe où les élèves sont assignés (Graphique 31.h).

Graphique 31 .Les courbes du modèle semi-paramétrique expliquant la performance en mathématiques

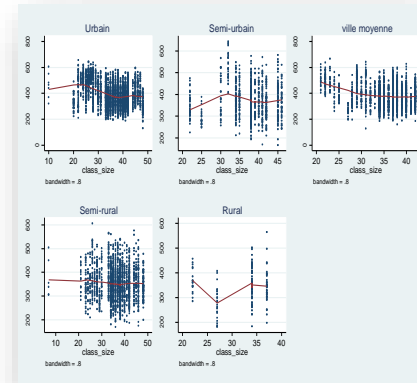
a. Genre de l'enseignant



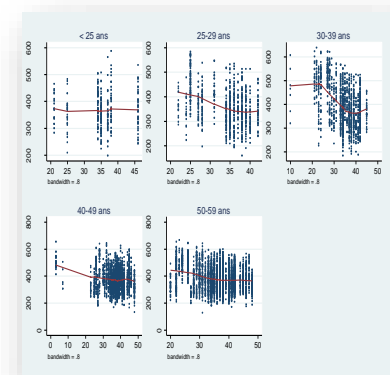
b. Absence des enseignants



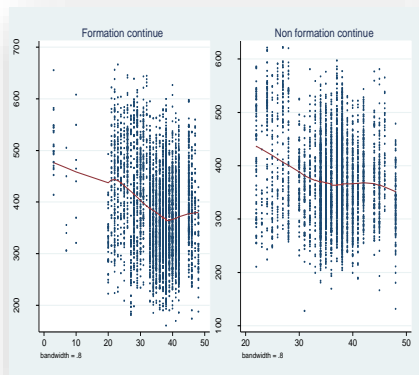
c. Milieu de l'établissement



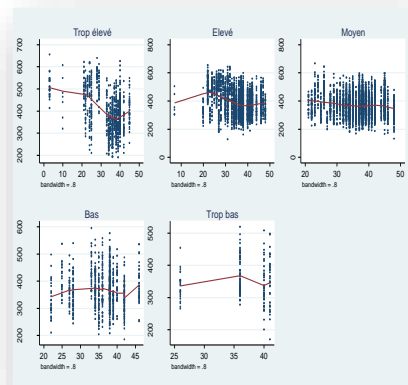
d. Age de l'enseignant



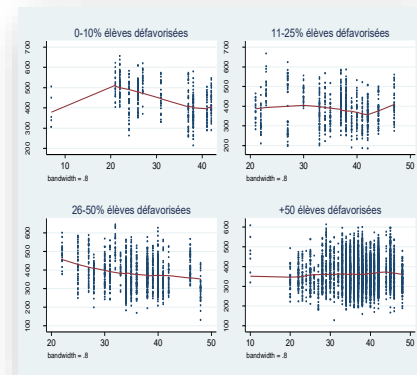
e. Formation continue



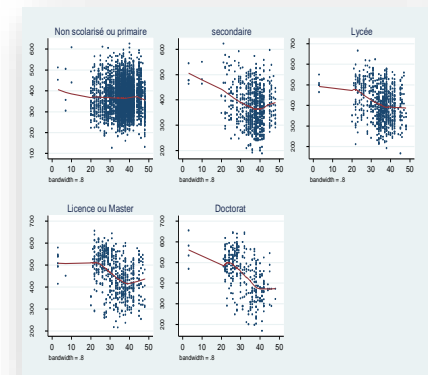
f. Satisfaction des enseignants



g. Mixité sociale



h. Niveau d'éducation des parents



Source : D'après l'auteur à partir des données TIMSS 2011

5.3. Taille optimale de la classe en sciences

Nous terminons par l'interprétation des résultats du modèle semi-paramétrique décrivant les tailles optimales de la classe en fonction du score des élèves dans les matières scientifiques.

Ainsi, nos résultats indiquent que la taille optimale de la classe pour les enseignants âgés de 25 à 29 ans est de 24 élèves par classe ; elle est de 21 élèves pour les enseignants âgés de 30-39 ans et seulement de 20 élèves pour les enseignants les plus âgés (50-59 ans) (Graphique 32.a).

S'agissant de la charge horaire, la taille des classes convenable est de 21 à 23 élèves, quelle que soit l'importance de ce problème. Les enseignants déclarant des problèmes graves de charge horaire ne devraient pouvoir accueillir au maximum 21 à 23 élèves, cela vaut également pour les enseignants qui ne sont pas concernés par ce problème (Graphique 32.b).

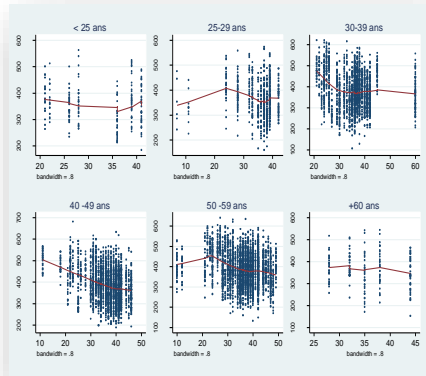
Concernant l'instruction des parents, la taille optimale des classes pour la réussite scolaire est de 22 élèves par classe pour tous les niveaux d'instruction des parents. De plus, à l'instar des résultats obtenus pour les mathématiques, ce résultat indique que la taille des classes comme composante du système éducatif ne doit pas être déterminée en fonction du niveau de vie de la famille ou bien de l'instruction des parents ou de tout autre facteur qui peut engendrer des faits discriminatoires (Graphique 32.d).

Nos résultats montrent également qu'un élève qui apprécie les matières scientifiques réussit d'autant mieux qu'il est affecté à une classe de petite taille. Notre courbe semi-paramétrique détermine une taille de classe de 21 à 22 élèves par classe abstraction faite du degré d'intérêt porté aux sciences (Graphique 32.e). En fait, pour que cette volonté persiste dans le temps et se traduise par une accumulation de connaissances et d'acquis scolaires, la classe doit être de taille réduite pour maintenir un niveau élevé des acquis scolaire.

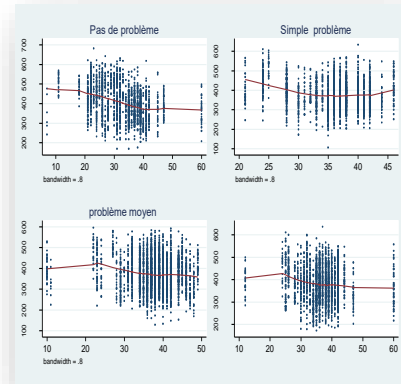
De même, selon le graphique 32.f, nous observons que la taille optimale pour les enseignants bénéficiaires d'une formation est de 22 élèves contre 20 élèves pour les non bénéficiaires. Le faible écart entre ces deux groupes d'enseignants qui n'est que de 2 élèves par classe traduit la faible valeur ajoutée de la formation continue en termes de maîtrise de l'environnement de la classe par les enseignants.

Graphique 32 .Les courbes du modèle semi-paramétrique expliquant le score en science

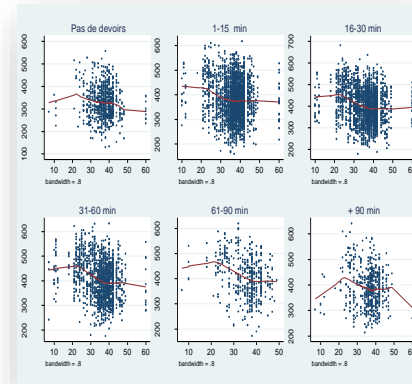
a. Age de l'enseignant



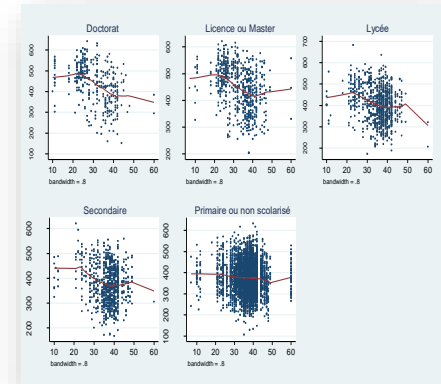
b. Charge horaire



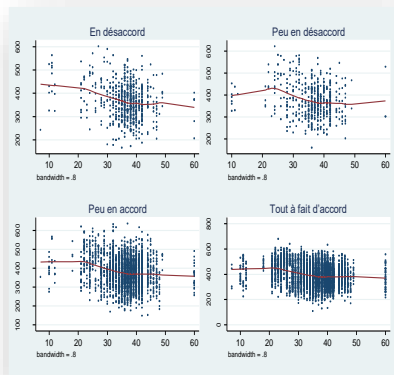
c. Devoirs à la maison



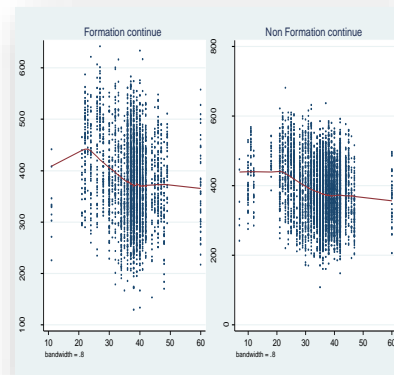
d. Education des parents



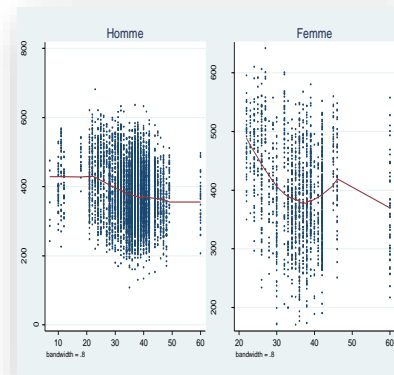
e. Apprécie les matières scientifiques



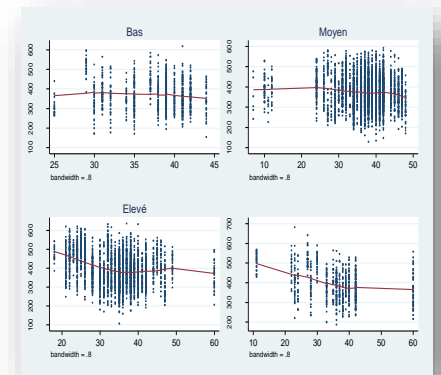
f. Formation continu



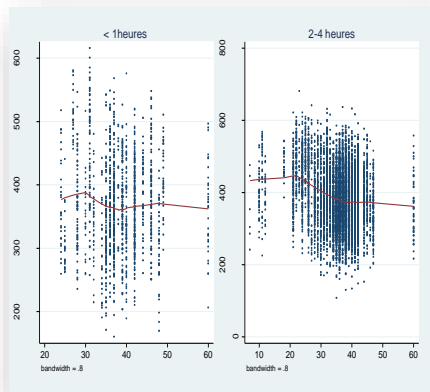
g. Genre de l'enseignant



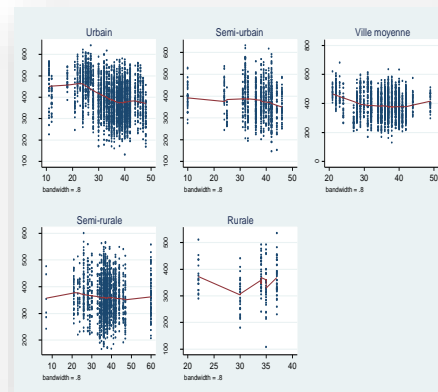
h. Satisfaction des enseignants



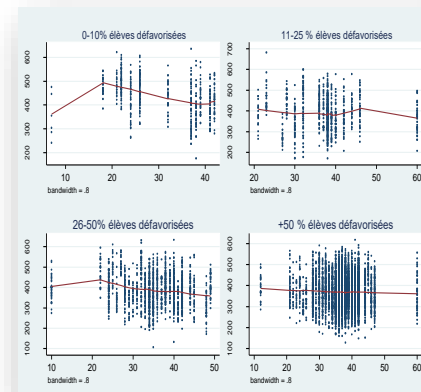
h. Temps consacré à la matière scientifique



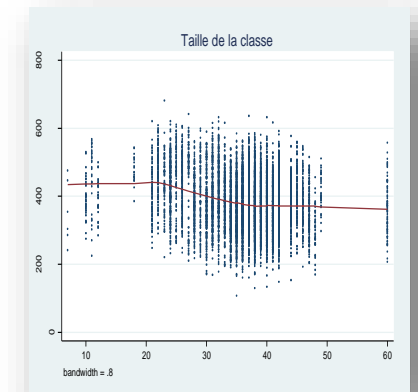
i. Milieu de l'établissement



j. Mixité sociale



k. Taille de la classe



Source : D'après l'auteur à partir des données TIMSS 2011

Encadré 3 : Etape d'estimation et interprétation des graphiques ci-dessus relatifs au Sciences (Graphique 32)

- Estimation du modèle MCO
- Estimation du modèle paramétrique
- Estimation du modèle semi paramétrique
- Estimation du modèle semi paramétrique en multiniveaux (Tableau 40 en Annexe du chapitre)
- Test d'Hausman (modèle semi paramétrique vs modèle semi-paramétrique en multiniveau (Tableau 40 en Annexe du chapitre 2))
- Simulation des courbes semi-paramétriques pour chaque variable contextuelle (Graphique 32)**

Nous retrouvons des résultats comparables à ceux obtenus pour la lecture et les mathématiques au regard du genre de l'enseignant. Ainsi, la taille optimale de la classe est de 21 élèves pour les femmes et 23 élèves pour les hommes (Graphique 32.g).

Les élèves dont l'enseignant se déclare satisfait de son métier ne réussissent pas forcément mieux que les élèves d'enseignants non satisfaits. Cela apparaît clairement sur les courbes du graphique 32.h qui indiquent une diminution du score des élèves lorsque la taille de la classe augmente, et ce quel que soit le niveau de satisfaction de l'enseignant envers son métier. Les résultats montrent une taille optimale de 22 à 25 élèves par classe quelle que soit le degré de satisfaction des enseignants (Graphique 32.h).

De même, en milieu urbain, la taille optimale de la classe est de 23 élèves et de 21 élèves en milieu rural (Graphique 32.i).

Ainsi, nos estimations révèlent que la taille de classe optimale se stabilise entre 22 et 23 élèves indépendamment de la composition sociale et économique de l'école où l'élève est assigné (Graphique 32.j).

Enfin, la taille optimale qui favorise la réussite de tous les élèves en matière scientifique est comprise entre 22 et 27 élèves. Ceci dit, quelconque le contexte où l'élève est assigné, la taille de la classe ne doit pas dépasser un effectif de 27 élèves pour assurer un environnement digne, une accumulation des acquis scolaires et par conséquent la correction des inégalités.

5.4. Discussions et recommandation

Suite à l'estimation des modèles semi-paramétriques, nos résultats montrent que la réduction de la taille de la classe contribue d'une manière considérable à l'amélioration de la performance scolaire et de la qualité de l'éducation telle qu'elle est mesurée par le score des élèves. En effet, notre simulation a permis d'identifier pour l'ensemble des matières évaluées un seuil optimal de 27 élèves par classe tout en mettant l'accent sur des tailles de classe, différenciées selon le contexte où l'élève est assigné

La littérature empirique examinant la relation entre taille de classe et performance scolaire révèle d'une façon unanime l'inexistence des études démontrant que l'augmentation de la taille des classes améliore de la performance scolaire. Bien que nos résultats confortent ceux des études antérieures, ils identifient avec plus de précision la taille optimale garantissant de meilleures performances pour tous les élèves.

De plus, nos résultats concluent que la réduction de la taille de la classe favorise la performance scolaire bien plus que toute autre politique éducative comme celle des pratiques pédagogiques, formation continue des enseignants, ou bien les transferts monétaires pour remédier aux décrochages scolaires. Ainsi, la réduction de la taille des classes est une politique éducative nécessaire et devrait figurer au rang des priorités de la réforme du système éducatif marocain, et par conséquent un impact indirect sur l'ensemble des composantes qui interviennent dans la production du capital humain.

Si la question de la taille des classes fait actuellement débat au niveau international, le débat au niveau national reste très modeste. Les résultats de ce travail recommandent d'instaurer des expérimentations randomisées dans l'objectif d'estimer l'impact réel d'une telle politique éducative sur la réussite des élèves, et aussi afin d'identifier les provinces et région nécessitant des interventions urgentes.

De plus, ce travail ouvre le débat et incite les pouvoirs publics chargé de l'évaluation du système éducatif de se pencher beaucoup plus sur l'évaluation de l'efficacité de cette politique éducative afin de confronter les résultats et propositions et anticiper des mécanismes et instruments efficaces et efficients pour offrir une éducation de qualité à tous les élèves scolarisés, et pour le bien-être de nos élèves marocains en particulier.

Conclusion

Dans ce chapitre, notre objectif a consisté à exploiter les résultats obtenus dans le chapitre 1 en partant de l'hypothèse que la taille de la classe exerce un effet significatif sur l'acquis scolaire des élèves. Afin de mettre en évidence le rôle joué par la taille de la classe sur les performances scolaires, nous nous sommes appuyés sur la littérature économétrique des modèles non-paramétriques et semi-paramétriques. Ces modèles sont d'une grande utilité pour évaluer les politiques publiques, et plus particulièrement la qualité de l'éducation car ils permettent d'identifier les différents seuils à partir desquelles l'acquis scolaire diminue de manière significative.

Pour ce faire, nous avons mobilisé des modèles semi-paramétriques prenant en considération la taille de classe dans la partie non paramétrique et les autres variables intrinsèques à l'élève et contextuelles à l'école dans la partie paramétrique de nos modèles.

Ces derniers ont permis d'identifier les seuils de la taille des classes d'une façon générale, mais aussi à partir conditionnellement aux facteurs contextuels. Cette stratégie a permis d'obtenir des résultats novateurs et inédits dans le contexte marocain. Nous avons pu ainsi identifier le seuil de la taille des classes dans le milieu urbain, pour les enseignants âgés, pour les enseignants du genre féminin, etc.

Dans ce chapitre, nous avons ainsi démontré l'intérêt d'une réduction de la taille des classes et son implication sur l'amélioration de la performance scolaire de tous les élèves. Globalement, les résultats obtenus aboutissent à des conclusions similaires pour les trois matières évaluées (lecture, mathématiques et sciences). Dans tous les cas, la réduction de la taille des classes augmente la performance scolaire pour l'ensemble des matières évaluées dans les dispositifs TIMSS et PIRLS. De même, nous avons montré que la taille optimale de la classe qui assure une performance scolaire assez élevée varie d'un contexte à un autre, mais cette variation est bornée à une taille de classe de 20 à 27 élèves. Ceci dit, abstraction faite du contexte dans lequel l'élève évolue, la taille optimale qui assure une performance scolaire juste et égalitaire pour l'ensemble des élèves du système éducatif marocain ne doit pas dépasser les 27 élèves par classe.

Si la politique de généralisation a réussi à ouvrir les portes des écoles pour les élèves marocains, cette généralisation doit être obligatoirement accompagnée par une égalité des résultats pour garantir la pertinence et l'atteinte des objectifs instaurés depuis des décennies par les pouvoirs

publics. D'après nos résultats de modélisation, la réduction de la taille des classes peut être une politique publique pertinente pour garantir une démocratisation basée sur l'égalité des résultats entre tous les élèves marocains quelles que soient leurs origines socioéconomiques. De même, cette réduction de la taille des classes engendrera un changement radical dans toutes les composantes du système éducatif. En effet, réduire la taille de la classe implique naturellement un environnement adéquat pour le travail de l'enseignant qui est le premier acteur de toute réussite scolaire.

Enfin, ce chapitre se conclut par deux recommandations sur le plan des protocoles d'évaluation à instaurer. Il s'agit en premier lieu de mobiliser une expertise outillée des méthodes d'évaluation avec pour objectif d'expérimenter la politique de réduction de la taille de la classe sur des régions ou provinces pilotes. De plus, une structure statistique solide demeure nécessaire pour suivre la dynamique des tailles de classe, leurs changements suite aux politiques publiques menées et les impacts observés sur l'amélioration de la réussite des élèves. Si les résultats du chapitre révèlent des constats concluants sur la pertinence de la politique de la réduction de la taille de la classe, cette dernière nécessitera inévitablement un suivi et des évaluations régulières en se référant aux protocoles nationales et à l'expertise internationale.

Annexes du chapitre 2

Tableau 36. Descriptions des variables sélectionnées du programme PIRLS 2011

Code PIRLS	Variables collectées à partir de PIRLS	Descriptions des variables
asrrea	Score des élèves en lecture	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 4 ^e année primaire en 2011 variant entre 300 à 600 points
btbg12	Le nombre d'élèves dans la classe	Une variable continue indiquant la taille de la classe
T1	La taille de la classe à plusieurs modalités	Une variable à plusieurs modalités : 1 : Taille <=24 ; 2 : 25<Taille<=35 ; 3 : Taille>=36
Variables intrinsèques à l'élève		
G1	Genre	Une variable dichotomique indiquant le genre de l'élève : 1 : garçon, 0 : fille.
G2	Age de l'élève en année de 2011	Une variable continue indiquant l'âge de l'élève
G3	Combien de fois l'élève parle la langue arabe à la maison	Une variable indicatrice à trois modalités : 1 : toujours ; 2 : parfois ; 3 : jamais
G4	Le nombre approximatif de livres dans votre maison	Une variable indicatrice à cinq modalités : 1 : entre 0 -10 ; 2 : entre 11-25 ; 3 : 26-100 ; 4 : entre 101-200 ; 5 : entre plus que 200
G5-D	Disponibilité d'une chambre propre à l'élève	Une variable dichotomique indiquant la disponibilité d'une chambre propre à l'élève : 1 : Oui, 0 : non.
G7-D	Les parents s'assurent si l'élève a bien accompli ses devoirs à la maison	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
G8-A	L'élève a le sentiment d'appartenance à l'école	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
G9-A	L'élève a été soumis à des situations d'insultes par d'autres élèves	Une variable indicatrice à quatre modalités : 1 : au moins une fois par semaine ; 2 : 1 à 2 fois par mois ; 3 : quelque fois pendant l'année ; 4 : jamais
R1	En une journée, combien de temps l'élève consacre à la lecture en dehors de l'école	Une variable indicatrice à quatre modalités : 1 : moins de 30 minutes ; 2 : de 30 minutes à une heure ; 3 : entre une heure et 2 heures ; 4 : plus que 2 heures
R2-A	Combien de fois l'élève lu pour le plaisir	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R4	Combien de fois l'élève emprunte des livres à la bibliothèque de l'école	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R5-A	L'élève aime ce qu'il lu à l'école	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
R5-B	L'enseignant donne à l'élève des choses importantes pour la lecture	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
R6-A	Combien de fois l'élève a-t-il réalisé une lecture silencieuse ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R7-F	L'élève trouve-t-il du plaisir à lire ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord

R8-A	L'élève lit très bien (niveau de performance en lecture)	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
Variables liées au contexte d'environnement familial (Parents)		
asbh02-A	Avant la rentrée à l'école primaire, combien de fois avez-vous lu des livres avec votre enfant ?	Une variable indicatrice à trois modalités : 1 : beaucoup ; 2 : parfois ; 3 : jamais
asbh02-H	Avant la rentrée à l'école primaire, combien de fois avez-vous écrit des lettres et des mots avec votre enfant ?	Une variable indicatrice à trois modalités : 1 : beaucoup ; 2 : parfois ; 3 : jamais
Asbh05	L'âge de rentrée scolaire à la première année du cycle primaire	Une variable indicatrice à quatre modalités : 1 : moins ou égal à 5 ans ; 2 : 6ans ; 3 : 7ans ; 4 : 8 ans et plus
Asbh04a	Est-ce que votre enfant a bénéficié de l'école maternelle ou du préscolaire ?	Une variable indicatrice à 2 modalités : 1 : Oui ; 2 : Non
Asbh04b	Combien d'années votre enfant a-t-il bénéficié de l'enseignement préscolaire ?	Une variable indicatrice à 5 modalités : 1 : 3 ans et plus ; 2 : entre 2 -3 ans ; 3 : 2 ans ; 4 : entre 1 – 2 ans ; 5 : moins ou égal à une année
Asbh08	En une journée, combien de temps votre enfant consacre-t-il pour accomplir les devoirs à la maison ?	Une variable indicatrice à 5 modalités : 1 : 0 minute ; 2 : moins ou égal à 15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31 -60 minutes ; 5 : plus que 60 minutes
Asbh10f	L'école fait-elle un bon travail pour améliorer le niveau de lecture de votre enfant ?	Une variable indicatrice à 4 modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
Asbh09a	Les parents discutent-ils avec les enfants de leurs devoirs à la maison ?	Une variable indicatrice à 4 modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
Asbh09f	Les parents aident-ils leurs enfants à lire à la maison ?	Une variable indicatrice à 4 modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
Asbh17	Quel est le plus haut niveau de scolarité atteint par les parents ?	Une variable indicatrice à 6 modalités : 1 : aucun diplôme ; 2 : diplôme primaire ; 3 : diplôme collégial ; 4 : diplôme lycéen ; 5 : licence ; 6 : Master ou Doctorat
Asbh19	Quel est le type de revenu des parents ?	Une variable indicatrice à 4 modalités : 1 : Emploi avec revenu ; 2 : Emploi avec revenu partiel ; 3 : sans revenu ; 4 : autre Emploi
Variables liées au contexte des pratiques d'enseignement (Enseignant)		
atbg01	Le nombre d'années d'expérience du professeur	Une variable continue indiquant le nombre d'années d'expérience de l'enseignant
atbg02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : homme ; 2 : femme
atbg03	L'âge de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
atbg04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collègue ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : doctorat.
atbg06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
atbg06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
atbg06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas

atbg07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
atbg08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
atbg08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
atbg11a	Le degré de satisfaction de l'enseignant dans son travail d'enseignement	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
atbg11e	L'enseignant envisage de continuer dans son travail d'enseignant pour une durée indéterminée	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
atbg16a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant ?	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
atbg16e	Est-ce que la présence d'élèves dissipés détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
atbg17a	La fréquence des rencontres parents enseignants pour discuter des progrès de l'élève	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
atbr02	Le temps consacré à la lecture pendant une semaine	Une variable continue indiquant le nombre d'heures consacré à la lecture en classe au cours de la semaine
atbr08f	L'enseignant apprend aux élèves du nouveau vocabulaire d'une façon régulière	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr08b	L'enseignant demande à chaque élève de lire à haute voix	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr09b	L'enseignant demande aux élèves de révéler les idées principales d'un texte	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr10a	L'enseignant demande aux élèves d'écrire quelques phrases sur un texte déjà traité	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr14	Combien de fois l'enseignant demande-t-il aux élèves de préparer un texte comme un devoir de lecture à la maison	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1-2 fois par semaine ; 4 : entre 3-4 fois par semaine ; 5 : chaque jour
atbr18c	L'enseignant consacre-t-il plus de temps avec un élève ayant des difficultés d'apprentissage	Une variable indicatrice à 2 modalités : 1 : Oui ; 2 : Non
atbr19a	Le degré d'importance consacré par l'enseignant pour l'évaluation régulière de l'apprentissage des élèves	Une variable indicatrice à trois modalités : 1 : une grande importance ; 2 : peu d'importance ; 3 : aucune importance
atbr21	Pendant les deux dernières années, combien d'heures ont été consacrées par l'enseignant pour la formation continue ?	Une variable indicatrice à cinq modalités : 1 : aucune heure ; 2 : moins de 6 heures ; 3 : entre 6-15 heures ; 4 : entre 16-35 heures ; 5 : plus de 35 heures
Variables liées au contexte de l'environnement de l'école (Directeur)		
acbg03a	Dans une classe, quel est le pourcentage d'élèves issus d'un milieu défavorisé ?	Une variable indicatrice à quatre modalités : 1 : entre 0-10 % ; 2 : entre 11-25% ; 3 : entre 26-50 % ; 4 : plus de 50 %

acbg05b	Dans quel milieu se situe l'école ?	Une variable indicatrice à cinq modalités : 1 : urbain ; 2 : semi urbain ; 3 : ville de taille moyenne ; 4 : semi rural ; 5 : rural
acbg06ba	Durant une journée, quel est le nombre d'heures consacré à l'éducation ?	Une variable continue indiquant le nombre d'heures consacré l'éducation pendant une journée de la semaine
acbg10ae	Quel est le degré d'existence des lieux consacrés à l'enseignement, et en premier lieu les classes ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne ; 4 : existence parfaite
acbg10ba	Quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiées à l'enseignement de la lecture ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne ; 4 : existence parfaite
acbg11aa	Informers les parents sur les progrès des élèves	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : une fois par an ; 3 : entre 2-3 fois par an ; 4 : plus que 3 fois par an
acbg13ab	Le problème d'absences des élèves	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13ac	Le problème de la violence dans la classe	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13ba	Le problème d'arrivée tardive à l'école ou de départ anticipé des enseignants	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13bb	Le problème d'absence des professeurs	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg15i	Le degré de la discussion du directeur avec les enseignants dont ceux qui ont des difficultés en classe	Une variable indicatrice à trois modalités : 1 : jamais ; 2 : parfois ; 3 : beaucoup
acbg17a	A quel niveau du primaire, l'élève a-t-il déjà des acquis suffisants pour la connaissance des lettres ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire
acbg17e	A quel niveau du primaire, l'élève a-t-il déjà des acquis suffisants pour la lecture d'un texte ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire
acbg17g	A quel niveau du primaire, l'élève a-t-il déjà des acquis suffisants pour identifier les idées principales d'un texte ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire

Source : D'après l'auteur à partir des données PIRLS 2011

Tableau 37. Descriptions des variables sélectionnées du programme TIMSS 2011

Code TIMSS	Variables collectées à partir de TIMSS	Descriptions des variables
bsmmat	Score des élèves en mathématiques	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2 ^e année collégiale en 2011 variant entre 300 à 600 points
bssbio	Score des élèves en science de la vie	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2 ^e année collégiale en 2011 variant entre 300 à 600 points
bssear	Score des élèves en science de la terre	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2 ^e année collégiale en 2011 variant entre 300 à 600 points
bssche	Score des élèves en chimie	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2 ^e année collégiale en 2011 variant entre 300 à 600 points
bssphy	Score des élèves en physique	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2 ^e année collégiale en 2011 variant entre 300 à 600 points
btbs12	Le nombre d'élèves dans la classe	Une variable continue indiquant la taille de la classe
T1	La taille de la classe à plusieurs modalités	Une variable à plusieurs modalités : 1 : Taille <=27 ; 2 : 28<Taille<=37 ; 3 : taille>=38
Variables intrinsèques à l'élève		
bsbg01	Genre	Une variable dichotomique indiquant le genre de l'élève : 1 : garçon, 0 : fille.
bsbg02b	Age de l'élève en année de 2011	Une variable continue indiquant l'âge de l'élève
bsbg04	Le nombre approximatif de livres dans votre maison	Une variable indicatrice à cinq modalités : 1 : entre 0 -10 ; 2 : entre 11-25 ; 3 : 26-100 ; 4 : entre 101-200 ; 5 : entre plus que 200
bsbg05d	Disponibilité d'une chambre propre à l'élève	Une variable dichotomique indiquant la disponibilité d'une chambre propre à l'élève : 1 : Oui, 0 : non.
bsbg06	Le niveau d'instruction des parents	Une variable indicatrice à sept modalités : 1 : niveau primaire ou analphabète ; 2 : primaire ; 3 : collège ; 4 : lycée ; 5 : licence ; 6 : master ; 7 : doctorat
bsbg10a	Combien de fois l'élève utilise-t-il l'ordinateur à la maison ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
bsbg11d	Les parents s'assurent-ils que l'élève a bien accompli ses devoirs à la maison ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
bsbg12a	L'élève a-t-il le sentiment d'appartenance à l'école ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
bsbg13a	L'élève a-t-il été soumis à des situations d'insultes par d'autres élèves ?	Une variable indicatrice à quatre modalités : 1 : au moins une fois par semaine ; 2 : 1 à 2 fois par mois ; 3 : quelque fois pendant l'année ; 4 : jamais
bsbg14a	L'élève ressent-t-il du plaisir à l'apprentissage des mathématiques	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
bsbg16a	Habituellement, l'élève travaille-t-il dur pour l'apprentissage des mathématiques	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
Bsbg17a	L'élève éprouve-t-il du plaisir pour l'apprentissage des sciences de la vie ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
Bsbg19a	Habituellement, l'élève travaille-t-il dur pour l'apprentissage des sciences de la vie	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord

Bsbg18a	L'élève éprouve-t-il du plaisir pour l'apprentissage des sciences de la terre ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg20a	Habituellement, l'élève travaille-t-il dur pour l'apprentissage des sciences de la terre	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
bsbg26a	L'élève éprouve-t-il du plaisir pour l'apprentissage de la chimie ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
bsbg28a	Habituellement, l'élève travaille-t-il dur pour l'apprentissage de la chimie ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : totalement en désaccord
Bsbg30a	L'élève éprouve-t-il du plaisir pour l'apprentissage de la physique ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg32a	Habituellement, l'élève travaille-t-il dur pour l'apprentissage de la physique ?	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg33ba	En une journée, combien de temps l'élève consacre-t-il aux devoirs en mathématiques ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 ; entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bb	En une journée, combien de temps l'élève consacre-t-il aux devoirs des sciences de la vie ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 ; entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bc	En une journée, combien de temps l'élève consacre-t-il aux devoirs des sciences de la terre ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 ; entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bd	En une journée, combien de temps l'élève consacre-t-il aux devoirs d en chimie ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 ; entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33be	En une journée, Combien de temps l'élève consacre aux devoirs de la physique	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 ; entre 61-90 minutes ; 6 : plus que 90 minutes.
Variables liées au contexte des pratiques d'enseignement (Enseignant en mathématiques)		
btdg01	Le nombre d'années d'expérience du professeur	Une variable continue indiquant le nombre d'années d'expérience de l'enseignant
btdg02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : homme ; 2 : femme
btdg03	Age de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
btdg04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collège ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : doctorat.
btdg06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btdg06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btdg06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btdg07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btdg08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème

btbg08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbg11a	L'enseignant a-t-il un sentiment de fierté dans son travail d'enseignement ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbg11e	L'enseignant envisage-t-il de poursuivre son travail d'enseignant pour une durée indéterminée ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbg15a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant ?	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbg15e	Est-ce que la présence d'élèves dissipés détériore les pratiques pédagogiques de l'enseignant ?	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbg16a	La fréquence des rencontres parents enseignants pour discuter des progrès des élèves concernant leurs apprentissages.	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
btbg17a	Durant la semaine, quel est le temps consacré à l'enseignement des mathématiques	Une variable continue indiquant le nombre d'heures consacré à l'enseignement des mathématiques durant la semaine
btbg19b	Combien de fois l'enseignant demande-t-il aux élèves d'apprendre les règles de mathématiques ?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
btbg19c	Combien de fois l'enseignant corrige-t-il les exercices en classe ?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
Btbg25a	Combien de fois l'enseignant demande-t-il aux élèves de préparer des devoirs de mathématiques à la maison ?	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1 -2 fois par semaine ; 4 : entre 3-4 fois par semaine ; 5 : chaque jour
Btbg27	Combien de fois l'enseignant programme-t-il des examens en mathématiques ?	Une variable indicatrice à cinq modalités : 1 : une fois par semaine ; 2 : une fois dans deux semaine ; 3 : une fois par mois ; 4 : quelque fois par ans ; 5 : jamais
Btbg29b	Pendant les deux dernières années ; l'enseignant a-t-il assisté à des formations continues pour l'amélioration des pratiques pédagogiques d'enseignement des mathématiques ?	Une variable dichotomique à deux modalités. 1 : Oui ; 2 : Non
Variables liées au contexte des pratiques d'enseignement des sciences (Enseignant des sciences)		
btbs01	Le nombre d'années d'expérience du professeur	Une variable continue indiquant le nombre d'année d'expérience de l'enseignant
btbs02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : Homme ; 2 : femme
btbs03	Age de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
btbs04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collège ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : doctorat.
btbs06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbs06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas

btbs06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbs07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbs08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbs08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbs11a	L'enseignant a-t-il un sentiment de fierté dans son travail d'enseignement ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbs11e	L'enseignant envisage-t-il de poursuivre son travail d'enseignant pour une durée indéterminée ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbs15a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant ?	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbs15e	Est-ce que la présence des élèves dissipés détériore les pratiques pédagogiques de l'enseignant ?	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbs16a	La fréquence des rencontres parents-enseignants pour discuter des progrès des élèves concernant leurs apprentissages	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
btbs17a	Pendant une semaine, quel est le temps consacré à l'enseignement des sciences ?	Une variable continue indiquant le nombre d'heures consacré à l'enseignement des mathématiques pendant une semaine
btbs19f	Combien de fois l'enseignant demande-t-il aux élèves d'apprendre les leçons des sciences ?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
Btbs24a	Combien de fois l'enseignant demande-t-il aux élèves de préparer des devoirs de sciences à la maison ?	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1 -2 fois par semaine ; 4 : entre 3-4 fois par semaine ; chaque jour
Btbs26	Combien de fois l'enseignant programme-t-il des examens en sciences ?	Une variable indicatrice à cinq modalités : 1 : une fois par semaine ; 2 : une fois dans deux semaine ; 3 : une fois par mois ; 4 : quelque fois par ans ; 5 : jamais
Btbs28b	Pendant les deux dernières années ; l'enseignant a-t-il assisté à des formations continues pour l'amélioration des pratiques pédagogiques d'enseignement des sciences ?	Une variable dichotomique à deux modalités. 1 : Oui ; 2 : Non
Variables liées au contexte de l'environnement de l'école (Directeur)		
Bcbg03a	Dans une classe, quel est le pourcentage des élèves issus d'un milieu défavorisé ?	Une variable indicatrice à quatre modalités : 1 : entre 0-10 % ; 2 : entre 11-25% ; 3 : entre 26-50 % ; 4 : plus de 50 %
bcbg05b	Dans quel milieu se situe l'école ?	Une variable indicatrice à cinq modalités : 1 : urbain ; 2 : semi urbain ; 3 : ville de taille moyenne ; 4 : semi rural ; 5 : rural
bcbg06ba	Durant une journée, quel est le nombre d'heures consacré à l'éducation ?	Une variable continue indiquant le nombre d'heures consacré l'éducation pendant une journée de la semaine

bcbg08a	Existe-t-il un laboratoire pour mener des expériences scientifiques par les élèves ?	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
bcbg09ae	Quel est le degré d'existence des lieux consacrés à l'enseignement, et en premier lieu les classes ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
bcbg09ba	Quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiés à l'enseignement des mathématiques ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
bcbg09ca	Quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiés à l'enseignement des sciences ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
Bcbg10ca	Informers les parents sur les progrès des élèves	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : une fois par an ; 3 : entre 2-3 fois par an ; 4 : plus que 3 fois par an
Bcbg12ab	Le problème d'absences des élèves	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12ac	Le problème de la violence dans la classe	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12ba	Le problème d'arrivée tardive à l'école ou de départ anticipé des enseignants	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12bb	Le problème d'absences des professeurs	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg16a	Le directeur de l'école motive-t-il les enseignants des mathématiques par des indemnités pour les encourager dans leurs travaux d'enseignements ?	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
Bcbg16b	Le directeur de l'école motive-t-il les enseignants des sciences par des indemnités pour les encourager dans leurs travaux d'enseignements	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
Bcbg17d	Combien de temps le directeur de l'école consacre-t-il pour contrôler l'évolution des apprentissages des élèves pour atteindre les objectifs scolaires ?	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : parfois ; 3 : toujours

Source : D'après l'auteur à partir des données TIMSS 2011

Tableau 38. Les résultats du modèle de la performance scolaire en lecture

Variable explicatives	OLS		Random effect		Semiparametric		Multilevel modeling Semiparametric	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
Féminin	30.74204	0.000	29.61692	0.000	30.41942	0.000	19.51309	0.000
Masculin (Ref)	-	-	-	-	-	-	-	-
Age <=2000	-11.83617	0.012	-1.631835	0.714	-10.54606	0.021	-1.348917	0.780
Age 2001-2003	48.09567	0.000	40.50728	0.000	45.10766	0.000	28.46975	0.000
Age >=2004 (Ref)	-	-	-	-	-	-	-	-
Sentiment d'appartenance à l'école (en accord)	55.32597	0.000	40.85874	0.000	51.97273	0.000	33.35112	0.000
Sentiment d'appartenance à l'école (peu d'accord)	28.95539	0.000	23.70187	0.000	25.47052	0.000	24.46831	0.000
Sentiment d'appartenance à l'école (en désaccord) (Ref)	-	-	-	-	-	-	-	-
Elève aime la lecture à l'école (tout à fait d'accord)	54.62612	0.000	41.32966	0.000	48.26346	0.000	33.48421	0.000
Elève aime la lecture à l'école (peu d'accord)	31.47427	0.000	23.75109	0.000	24.92536	0.000	22.3705	0.000
Elève aime la lecture à l'école (En désaccord) (Ref)	-	-	-	-	-	-	-	-
Expérience d'enseignant (<=10 ans)	37.9144	0.000	29.60619	0.008	24.20927	0.000	-	-
Expérience d'enseignant (10-20 ans)	-16.23737	0.000	-17.51626	0.077	-23.99415	0.000	-	-
Expérience d'enseignant (20-30 ans)	-5.193438	0.054	-4.735647	0.645	-2.490332	0.386	-	-
Expérience d'enseignant (>=31 ans) (Ref)	-	-	-	-	-	-	-	-
Enseignant (femme)	30.80461	0.000	36.02333	0.000	28.7824	0.000	-	-
Enseignant (homme) (Ref)	-	-	-	-	-	-	-	-
Satisfaction des enseignants (trop élevé)	37.64592	0.000	47.87832	0.000	31.01258	0.000	-	-
Satisfaction des enseignants (élevé)	15.28264	0.000	13.76681	0.114	12.42885	0.000	-	-
Satisfaction des enseignants (bas) (Ref)	-	-	-	-	-	-	-	-
Connaissance des programmes (trop élevé)	70.06752	0.000	61.21632	0.000	59.74059	0.000	-	-
Connaissance des programmes (élevé)	21.14414	0.000	19.48835	0.012	14.63619	0.000	-	-
Connaissance des programmes (bas) (Ref)	-	-	-	-	-	-	-	-
Charge horaire (pas de problème)	28.55891	0.000	29.37826	0.012	16.71691	0.000	-	-
Charge horaire (simple problème)	22.7982	0.000	1.064725	0.920	17.32236	0.000	-	-

Charge horaire (grave problème) (Ref)	-	-	-	-	-	-	-	-
Les élèves bavardes beaucoup	-6.983234	0.058	-3.750436	0.794	-2.722161	0.494	-	-
Les élèves bavardes peu	.116869	0.962	4.474093	0.609	-1.295159	0.622	-	-
Les élèves bavardes pas du tout (Ref)	-	-	-	-	-	-	-	-
Aucune formation continue	-26.26731	0.000	-25.50049	0.001	-23.10665	0.000	-	-
Ayant accès à une formation continue (Ref)	-	-	-	-	-	-	-	-
entre 0-10 % d'élèves défavorisées	39.48452	0.000	42.90067	0.001	28.1376	0.000	-	-
entre 11-25% d'élèves défavorisées	23.177	0.000	20.11078	0.181	12.08987	0.004	-	-
entre 26-50% d'élèves défavorisées	24.28806	0.000	26.39204	0.043	25.96315	0.000	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-	-	-
Ecole urbaine	40.84379	0.000	46.22205	0.000	37.99302	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-	-	-
Préscolaire >=3 ans	46.35206	0.000	26.23574	0.000	44.76863	0.000	22.9481	0.000
Préscolaire 2-3 ans	45.08356	0.000	25.15206	0.000	42.53727	0.000	18.80169	0.000
Préscolaire 2 ans	27.19128	0.000	13.08949	0.000	28.25373	0.000	6.824862	0.005
Préscolaire 1-2 ans	10.86443	0.020	1.727099	0.649	11.42994	0.011	-4.457657	0.219
Préscolaire < =1 ans (Ref)	-	-	-	-	-	-	-	-
Taille de la classe >=40 élèves	-28.50009	0.000	-21.82698	0.094	-77.40965	0.028	-	-
Taille de la classe 30-40 élèves	-19.36359	0.000	-19.52988	0.099	-61.91279	0.087	-	-
Taille de la classe 20-30 élèves	13.98158	0.002	11.65746	0.352	-165.5254	0.000	-	-
Taille de la classe < 20 élèves	-	-	-	-	-	-	-	-
Constante	304.525	0.000	306.4257	0.000				
Test d'Hausman	253,92		120.87		0.00		0.00	0.00

Source : D'après l'auteur à partir des données PIRLS 2011

Tableau 39. Les résultats du modèle de la performance en mathématiques

Variable explicatives	OLS		Random effect		Semiparametric		Multilevel modeling Semiparametric	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
masculin	1.521357	0.378	6.757123	0.000	2.510189	0.125	21.14064	0.000
Féminin (Ref)	-	-	-	-	-	-	-	-
Age >=1995	-56.84253	0.000	-40.90929	0.000	-50.79661	0.000	-33.38621	0.000
Age <1995 (Ref)	-	-	-	-	-	-	-	-
Education des parents –Doctorat	78.97422	0.000	21.19137	0.000	58.31072	0.000	.4676801	0.905
Education des parents –Master	90.77385	0.000	44.11059	0.000	74.56632	0.000	23.71804	0.000
Education des parents –Licence	42.81589	0.000	23.58623	0.000	37.80504	0.000	10.94154	0.000
Education des parents –lycée	10.16884	0.000	7.146328	0.004	9.216851	0.001	3.038688	0.196
Education des parents - secondaire, primaire ou analphabète(Ref)	-	-	-	-	-	-	-	-
Utilisation ordinateur chaque jour	47.02342	0.000	20.52347	0.000	37.78472	0.000	12.42387	0.000
Utilisation ordinateur 1-2 semaines	37.93694	0.000	15.1399	0.000	31.14932	0.000	7.136514	0.000
Utilisation ordinateur 1-2 mois	10.98317	0.009	.3229748	0.928	7.141151	0.077	-2.003442	0.550
Utilisation ordinateur jamais(Ref)	-	-	-	-	-	-	-	-
Sentiment plaisir pour les maths (accord)	44.69195	0.000	40.81698	0.000	42.65714	0.000	35.47957	0.000
Sentiment plaisir pour les maths (peu d'accord)	17.41163	0.000	15.40079	0.000	17.23904	0.000	14.49769	0.000
Sentiment plaisir pour les maths(en désaccord) (Ref)	-	-	-	-	-	-	-	-
Pas de devoirs mathématiques	-24.9401	0.000	-16.93702	0.000	-23.74051	0.000	-14.84844	0.000
Devoirs mathématiques 1 -15 min	21.99426	0.000	23.37827	0.000	22.3927	0.000	21.86412	0.000
Devoirs mathématiques 16-30 min	41.01788	0.000	33.01783	0.000	36.61232	0.000	29.19983	0.000
Devoirs mathématiques 31-60 min	53.63322	0.000	37.70766	0.000	47.75867	0.000	32.25445	0.000
Devoirs mathématiques 61-90 min	44.90338	0.000	28.16592	0.000	37.3498	0.000	24.91404	0.000
Devoirs mathématiques plus de 90 minutes (Ref)	-	-	-	-	-	-	-	-
entre 0-10 % d'élèves défavorisées	73.26818	0.000	79.15912	0.000	50.15381	0.000	-	-
entre 11-25% d'élèves défavorisées	12.68372	0.000	10.52248	0.355	10.52476	0.001	-	-
entre 26-50% d'élèves défavorisées	9.774711	0.000	9.867382	0.265	12.30509	0.000	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-	-	-

Ecole urbaine	19.90783	0.000	23.70874	0.000	17.79574	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-	-	-
Enseignants femme	-2.136956	0.275	-4.561074	0.548	-1.627668	0.391	-	-
Enseignants homme (Ref)	-	-	-	-	-	-	-	-
Age de l'enseignant <=25	-11.33588	0.020	-14.04169	0.458	-12.16002	0.010	-	-
Age de l'enseignant 26-29	-19.93087	0.000	-19.773	0.094	-25.77923	0.000	-	-
Age de l'enseignant 30-39	16.23968	0.000	22.21889	0.024	9.618203	0.000	-	-
Age de l'enseignant 40-49	-3.587636	0.090	-4.171343	0.611	-4.396603	0.034	-	-
Age de l'enseignant >=50	-	-	-	-	-	-	-	-
7-8 heures de maths par semaine	40.70134	0.000	52.45846	0.113	35.06865	0.000	-	-
4-6 heures de maths par semaine	11.405	0.000	16.54348	0.077	3.959818	0.095	-	-
<=3 heures de maths par semaine (Ref)	-	-	-	-	-	-	-	-
Présence des Formations continues	12.03072	0.000	14.41156	0.030	6.135196	0.000	-	-
Absence des Formations continues (Ref)	-	-	-	-	-	-	-	-
Taille de la classe >=38 élèves	-60.97185	0.000	-60.97185	0.000	117.8225	0.000	-	-
Taille de la classe 28-37 élèves	-65.14481	0.000	-65.14481	0.000	130.7174	0.000	-	-
Taille de la classe <=27 élèves (Ref)	-	-	-	-	-	-	-	-
Constante	376.1586	0.000	377.2089	0.000	-	-	-	-
Test d'Hausman	211.92		111.87		0.00		0.00	0.00

Source : D'après l'auteur à partir des données TIMSS 2011

Tableau 40. Les résultats du modèle de la performance scolaire en sciences

Variables explicatives	OLS		RE		Semiparametric		Semiparametric Multilevel modeling	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
masculin	3.682634	0.029	-1.649827	0.265	2.893382	0.073	-13.72843	0.000
Féminin (Ref)	-	-	-	-	-	-	-	-
Age >=1995	-56.14689	0.000	-43.69694	0.000	-50.52464	0.000	-34.69205	0.000
Age <1995 (Ref)	-	-	-	-	-	-	-	-
Education des parents –Doctorat	66.2042	0.000	23.15163	0.000	45.24603	0.000	2.185929	0.597
Education des parents –Master	78.68997	0.000	42.9338	0.000	63.86833	0.000	23.09439	0.000
Education des parents –Licence	34.98397	0.000	19.71945	0.000	30.87818	0.000	9.044273	0.000
Education des parents –lycée	7.81443	0.006	5.396039	0.036	7.99668	0.004	2.010922	0.419
Education des parents - secondaire, primaire ou analphabète(Ref)	-	-	-	-	-	-	-	-
Devoirs des sciences (pas de devoirs)	-13.82581	0.000	-13.58693	0.000	-14.74032	0.000	-10.45394	0.000
Devoirs des sciences (1-15 min par jour)	14.78997	0.000	8.311442	0.000	12.78463	0.000	6.984019	0.000
Devoirs des sciences (16-30 min par jour) (Ref)	23.90115	0.000	15.20471	0.000	21.506	0.000	12.88193	0.000
entre 0-10 % d'élèves défavorisées	58.85971	0.000	61.73702	0.000	33.83721	0.000	-	-
entre 11-25% d'élèves défavorisées	10.62323	0.000	8.613387	0.384	7.776911	0.010	-	-
entre 26-50% d'élèves défavorisées	9.298942	0.000	9.451129	0.219	6.344314	0.007	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-	-	-
Ecole urbaine	16.0326	0.000	18.88021	0.001	14.63258	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-	-	-
Présence d'un laboratoire	8.691656	0.000	12.25506	0.111	2.573442	0.252	-	-
Absence d'un laboratoire (Ref)	-	-	-	-	-	-	-	-
Encourager les enseignants par indemnités	63.3665	0.000	69.80307	0.000	45.47064	0.000	-	-
Pas d'encouragement des enseignants (Ref)	-	-	-	-	-	-	-	-
Enseignant femme	15.74753	0.000	17.93864	0.014	12.37733	0.000	-	-
Enseignant homme (Ref)	-	-	-	-	-	-	-	-
Age <=25 ans	-15.1535	0.024	-10.95138	0.615	-36.40838	0.000	-	-
Age 25 -29 ans	-4.457353	0.362	-3.195179	0.846	-12.32296	0.010	-	-

Age 30 -39 ans	10.75988	0.008	15.88054	0.245	1.409563	0.728	-	-
Age 40 -49 ans	15.75731	0.000	21.54152	0.110	9.292692	0.017	-	-
Age 50 -59 ans	15.85092	0.000	19.86424	0.128	13.35702	0.001	-	-
Age >=60 ans (Ref)	-	-	-	-	-	-	-	-
Satisfaction des enseignants (trop élevé)	23.14368	0.000	28.6317	0.007	14.58949	0.000	-	-
Satisfaction des enseignants (élevé)	19.36082	0.000	24.39383	0.000	12.6545	0.000	-	-
Satisfaction des enseignants (bas) (Ref)	-	-	-	-	-	-	-	-
Connaissance des programmes (trop élevé)	35.50256	0.000	42.79779	0.000	24.53813	0.000	-	-
Connaissance des programmes (élevé)	6.060399	0.001	7.569738	0.198	7.409123	0.000	-	-
Connaissance des programmes (bas) (Ref)	-	-	-	-	-	-	-	-
Demande aux élèves de préparer les devoirs des sciences (jamais)	-24.95286	0.000	-29.95469	0.062	-27.42432	0.000	-	-
Demande aux élèves de préparer les devoirs des sciences (beaucoup) (Ref)	-	-	-	-	-	-	-	-
Presence des Formations continues	3.387464	0.052	4.113184	0.488	-.6803638	0.690	-	-
Absence des Formations continues (Ref)	-	-	-	-	-	-	-	-
Taille de la classe >=38 élèves	-62.48297	0.000	-63.41242	0.000	-67.86544	0.000	-	-
Taille de la classe 28-37 élèves	-64.53258	0.000	-65.52218	0.000	-68.74453	0.000	-	-
Taille de la classe <=27 élèves (Ref)	-	-	-	-	-	-	-	-
Constante	379.5752	0.000	388.31424	0.000				
Test d'Hausman	118.2		90.87		0.00		0.00	0.00

Source : D'après l'auteur à partir des données TIMSS 2011

Chapitre 3 : Les effets de la composition sociale et économique de l'école sur la performance scolaire des élèves marocains

Introduction

Toutes les écoles se différencient entre elles notamment selon la composition sociale et économique des élèves qui les fréquentent. Ce constat a été au centre de nombreuses réflexions et de recherches dans le champ de l'économie de l'éducation qui s'intéresse au contexte social d'apprentissage des élèves. L'interaction entre l'appartenance sociale des élèves et leur réussite scolaire a été largement débattue en raison de la divergence des résultats empiriques obtenus selon les méthodes d'estimation et des systèmes éducatifs évalués.

Les premières études sur les déterminants des performances scolaires se sont focalisées sur le rôle de l'environnement scolaire et son impact sur la performance scolaire des élèves (Coleman, 1966). Or les études plus récentes mettent en avant le rôle joué par les facteurs liés à l'établissement scolaire et l'influence des pairs sur la performance scolaire.

L'origine sociale est considérée comme le facteur déterminant le plus débattue dans la littérature théorique et empirique de la sociologie et l'économie de l'éducation. En effet, la plupart des études convergent vers l'existence d'une relation étroite entre l'environnement familial de l'élève et la réussite scolaire. Les analyses empiriques estiment généralement l'environnement familial par le statut socio-économique des parents (SSE), mesuré par le niveau de scolarité, la profession des parents et le revenu familial (Coleman et al, 1966 ; hakkinen et al. 2003 ; Heyneman et Loxley, 1983). A partir des données du Programme international pour le suivi des acquis des élèves (PISA), Fuchs et Wößmann (2004) concluent que l'éducation des parents a un effet plus significatif sur la réussite des élèves de 15 ans notamment en lecture comparativement aux mathématiques et aux sciences. D'autres travaux trouvent que les élèves issus de familles dont le niveau d'éducation des parents est faible ont inévitablement de moins bon résultats (Schiller et al. ,2002). Les auteurs stipulent que les parents instruits apportent à leurs enfants un soutien pédagogique et social important pour la réussite scolaire, comparativement aux parents dont le niveau d'éducation est faible.

Néanmoins, dans certaines études l'effet de la famille peut être négatif ou non significatif suggérant que l'impact du statut professionnel des parents et de l'interaction parents-enfants n'est pas clair (Chevalier et Lanot, 2002 ; Ganzach, 2010), et donc le statut socioéconomique des parents n'a pas d'impact significatif sur les apprentissages des élèves et sur leur réussite scolaire. Les auteurs proposent une méthodologie qui sépare les effets financiers des effets familiaux, et par conséquent ils constatent que les élèves issus de familles plus pauvres sont

moins susceptibles d'investir dans l'éducation. Toutefois, un transfert financier n'entraînerait pas une augmentation significative des investissements dans la scolarité, ce qui démontre que les effets caractéristiques de la famille dominent les effets des contraintes financières. De plus Iverson et Walberg (1982) affirment que l'environnement psychosociologique et le degré de stimulation intellectuelle à la maison sont plus profondément liés à la performance scolaire que le statut socio-économique (Meuret et Morlaix, 2006).

Hanchane et Mostafa (2012) révèlent que les effets de pairs ont un impact significatif sur l'amélioration des apprentissages de tous les élèves, tandis que les autres facteurs liés aux ressources scolaires engendrent des effets moindres.

Pourtant, force est de constater que la plupart des systèmes éducatifs ne se caractérisent pas par une mixité sociale réelle au sein des établissements scolaires. L'émergence d'un enseignement privé, la ségrégation sociale liée au lieu d'habitation, et la mise en place d'une carte scolaire, sont autant de facteurs agissant à l'encontre des politiques éducatives visant l'hétérogénéité et la mixité sociale à l'école.

Les études empiriques consacrées aux effets de pairs³⁸ aboutissent à des résultats divergents, selon les systèmes éducatifs et les méthodes d'estimations utilisées pour tester l'effet de cette composition sur l'amélioration des apprentissages des élèves. Cette divergence des résultats empiriques met en évidence un débat profond sur les dispositifs d'évaluation et des méthodes d'estimation d'un tel effet.

Par ailleurs, les évaluations empiriques menées sur le système éducatif marocain restent très insuffisantes et généralistes en matière de recommandation et de désagrégation des résultats (Hanchane et al 2012) et (Mourji et Abbaia 2013). A notre connaissance, il n'existe aucune étude (pour le cas du Maroc) qui examine de manière détaillée l'effet de la composition sociale et économique sur l'apprentissage des élèves marocains. Les travaux existants s'appuient sur des modèles insuffisants en termes de spécification et de traitement des problèmes soulevés par leur estimation. En effet, les résultats de Hanchane et al. 2012 et de Mourji et Abbaia, 2013 se limitent à montrer l'existence d'un effet significatif de la composition sociale et économique sur la performance scolaire mesuré généralement par les caractéristiques intrinsèques de l'élève, à savoir l'éducation des parents, le revenu, la disponibilité des ressources à la maison, etc.

³⁸L'effet de composition est défini comme l'effet engendré par un élève sur les autres élèves du même groupe, inversement.

Cependant, elles n'apportent pas d'indication sur l'existence d'effets de seuil dans la composition sociale et économique au sein de l'école et en conséquence ne permettent d'identifier quel serait le degré de mixité sociale pour des divers contextes qui assurerait à la fois l'équité et la réussite scolaire.

Dans cette perspective, ce chapitre étudie l'effet de la composition sociale et économique de l'école sur la performance scolaire des élèves marocains de la 4^e année primaire et de la 2^e année collégiale. Dans notre étude, la composition sociale et économique de l'école est appréciée à partir de la déclaration du directeur de l'établissement³⁹. Afin d'évaluer les effets de la composition sociale et économique, nous mobiliserons une modélisation semi-paramétrique laquelle permettra d'identifier les degrés de mixité sociale les plus avantageux et bénéfiques pour une performance scolaire généralisée.

Ce chapitre est organisé 4 sections. La section1 est consacrée à la revue de la littérature. La section2 présente la méthodologie adoptée pour tester l'effet de la mixité sociale sur la performance et déterminer le niveau du degré de cette mixité qui assure une réussite scolaire de tous les élèves. La section 3 décrit les données et statistiques descriptives. Enfin, dans la section 4, nous présenterons les résultats à partir desquels nous formulerons des recommandations pour une politique éducative juste en matière d'inclusion sociale et réduction des inégalités.

1. Revue de la littérature

1.1. Littérature théorique et apport méthodologique

Le degré de ségrégation sociale à l'école varie sensiblement selon les systèmes éducatifs. Cela se manifeste par le niveau et degré de mixité que chaque école enregistre au cours du temps.

Les performances scolaires des élèves sont influencées par leurs caractéristiques intrinsèques, mais aussi, entre autres facteurs, par celles des autres élèves de leur établissement. Les camarades d'école peuvent se motiver et s'entraider pour surmonter les difficultés d'apprentissage ; mais ils peuvent aussi perturber les cours, nécessiter une attention disproportionnée de la part des enseignants et devenir une source d'inquiétude. Les modalités

³⁹Le directeur d'établissement déclare à travers le remplissage d'un questionnaire le niveau de composition sociale et économique de l'école. Il a quatre catégories au choix : 25 % des élèves favorisés, 25-50% des élèves défavorisés, 50-75% des élèves défavorisés, plus que 75% des élèves défavorisés

d'affectation des élèves dans les établissements peuvent contribuer à creuser les différences socio-économiques ou académiques (stratification) entre les établissements, peuvent donc et ainsi avoir une incidence sur les résultats d'apprentissage au niveau national (OCDE-PISA 2018). Selon les résultats de l'enquête PISA 2015, dans nombre de pays, les établissements sont plus susceptibles de présenter de fortes concentrations d'élèves favorisés sur le plan socio-économique que d'élèves défavorisés. De plus, la ségrégation sociale est plus prononcée dans les systèmes éducatifs inéquitables en matière d'apprentissage et au niveau des acquis scolaires (France, Maroc, etc.).

Gilotte et Girard (2005) ont étudié la question de la mixité sociale à partir d'un échantillon de 109 collèges publics de la région parisienne en France. Leurs résultats indiquent que sur 58 écoles issues (milieu urbain) d'un milieu favorisé, 46% des élèves sont issus de classes sociales aisées et seulement 18% sont issus d'un milieu modeste. En revanche, sur les 17 écoles issues d'un milieu défavorisé (milieu urbain), les rapports ne sont que de 7 à 51% respectivement.

Bien qu'un assouplissement de la carte scolaire puisse favoriser la mixité sociale, et par la suite une performance scolaire généralisée pour tous les élèves scolarisés, ce résultat n'est pas toujours atteint, en raison de des comportements de choix et de stratégies scolaires des parents.

Plusieurs études montrent que ces politiques au service de la mixité sociale sont favorables à l'amélioration de la performance des élèves issus de milieux sociaux défavorisés mais en même temps elles engendrent un déséquilibre en hétérogénéité spatiale en raison de la fuite de certaines familles aisées (Baum-Snow et Lutz, 2011) et par conséquent elles contribuent à l'augmentation des taux de pauvreté au sein des classes les plus défavorisées.

D'un point de vue des choix rationnels, le choix de l'école par les parents représente un facteur déterminant du choix de leur résidence. Afin de contourner la carte scolaire, les familles aisées sont parfois prêtes à déménager dans des quartiers situés à proximité ce qui contribue d'ailleurs à l'augmentation du prix de l'immobilier dans ces quartiers.

Ainsi, Fack et Grenet (2009, 2010) ont montré qu'une augmentation des prix de vente des appartements de 1,5% génère une augmentation de la note moyenne d'un écart type pour les élèves de la 3e année collégiale à Paris. En effet, La carte scolaire n'accentue pas seulement la ségrégation scolaire, mais elle configure également les choix résidentiels en fonction de la qualité des établissements de la commune.

Dans certains systèmes éducatifs (Maroc, Tunisie, Jordanie, etc), l'inscription dans des établissements scolaires privés, perçue comme un gage de réussite scolaire pour les parents, est une autre stratégie de contournement de la carte scolaire. Ce choix stratégique qui tend à se développer notamment au Maroc représente un véritable obstacle à la lutte contre la ségrégation scolaire.

Pour autant, selon une étude menée en France (Réforme de 2007 de l'assouplissement de la carte scolaire de 2007-DEPP⁴⁰), l'assouplissement de la carte scolaire n'a pas nécessairement d'impact significatif sur la réduction de la ségrégation scolaire. Ce résultat est expliqué en grande partie par le manque d'information parmi les classes les plus défavorisées et au niveau d'éducation limité pour rechercher et exploiter l'information utiles dans l'assignation des élèves aux établissements.

Les expériences internationales en la matière indiquent que les multitudes choix offerts dans les deux secteurs éducatifs privé et public n'ont pas eu d'impacts significatifs sur l'amélioration de la performance scolaire de l'ensemble des élèves.

A ce titre, il semble intéressant de mentionner le cas de l'expérience chilienne qui sert de référence en matière de dynamique des différentes formes de ségrégation sociale à l'école. La réforme éducative menée au Chili (1981) s'est traduite par le passage d'un système éducatif étatique fortement centralisé à un système décentralisé au niveau des municipalités et avec une participation croissante du secteur privé. Les établissements privés s'engageant à appliquer le programme scolaire du ministère bénéficiaient de subventions. L'Etat restait garant de l'éducation primaire et prévoyait l'attribution de bourses aux élèves issus des milieux défavorisés pour favoriser leur accès aux écoles privées.

De ce fait, la décentralisation des décisions publiques a donné lieu à plus de souplesse pour une gestion revisitée de l'établissement scolaire. Ainsi, le secteur privé a conservé tous ses droits (frais d'inscription, nombre d'élève par classe, gestion autonome) sans exiger des frais d'inscription supérieurs au montant de la bourse étatique octroyée aux élèves sélectionnés à partir de critères socio-économiques.

⁴⁰Direction de l'évaluation, de la prospective et de la performance

Entre 1980 à 1990, cette réforme a entraîné une expansion du secteur privé chilien passant de 1000 à 2000 écoles, et une augmentation des effectifs des élèves inscrits de 20 à 40 % notamment en milieu urbain et dans les municipalités les plus riches.

En exploitant les enquêtes internationales d'évaluation des connaissances au Chili, Hsieh et Urquiola (2006) ont montré que l'extension des aides financières pour l'inscription des élèves ciblés dans des écoles privées n'a pas eu d'effet significatif sur l'amélioration des résultats scolaires moyens des élèves. Cependant, le programme a créé un phénomène de tri des élèves car les meilleurs élèves des écoles publiques sont partis dans les écoles privées.

Parmi d'autres expériences menées dans ce sens, Söderström et Uusitalo (2010) ont évalué l'impact de la réforme de la suppression de la carte scolaire dans la ville de Stockholm en 2000 sur la ségrégation. Depuis 2000, les parents peuvent inscrire leurs enfants dans n'importe quelle école de la ville tandis que les décisions d'admission dans les établissements s'appuient sur les résultats scolaires.

Les auteurs montrent que la liberté de choix des écoles s'est traduite par une modification significative de la répartition des élèves entre les établissements. Ainsi, la proportion des élèves qui ont changé d'établissements scolaires pour une école d'un autre arrondissement est passée de 48% à 55% entre 1999 à 2000. Cette mobilité des élèves entre les établissements scolaires a engendré une stratification socio-économique des établissements.

De même, Söderström (2006) montre que cette réforme n'a pas eu d'effet significatif sur l'amélioration de la performance des élèves scolarisés. En revanche, La réforme a accru la sélection des élèves selon leurs aptitudes mais aurait eu également pour effet inattendu d'accroître la ségrégation selon le statut socio-économique, entre les autochtones et les immigrants.

En revanche, l'étude menée par Lavy (2012) sur les établissements publics de Tel-Aviv montre que l'instauration d'un système de choix de l'établissement scolaire a eu des impacts positifs sur l'amélioration de la performance scolaire de tous les élèves sans augmenter la stratification au sein des écoles. Le système instauré a été bien confectionné puisqu'il assure une certaine homogénéité au sein des différentes écoles de Tel-Aviv.

Cependant, la liberté de choix de l'établissement n'est pas sans soulever certaines difficultés car la répartition des élèves suppose la mise en place d'un mécanisme qui permette de rendre compatible les critères relatifs à l'admission des élèves dans les établissements à la liberté de

choix des parents. Ainsi, les travaux de recherche fondateurs d'Alvin Roth et Lloyd Shapley (2012) sur les processus d'affectation ont été exploités pour permettre de définir des politiques éducatives efficaces permettant un choix scolaire optimal et une réussite scolaire de tous les élèves scolarisés.

D'un point de vue théorique, le processus d'affectation des élèves doit valider trois critères. Le critère de l'allocation efficace qui repose sur le principe d'amélioration de la performance d'un élève sans dégrader celle d'un autre. En d'autres termes, le processus doit produire des allocations efficaces selon le principe de Pareto.

Le second critère appelé la situation stable où le respect des priorités des pouvoirs publics doit être vérifié. Et le dernier critère du mécanisme d'affectation est d'être exempt de toute manipulation. Il doit inciter les parents d'élèves à révéler leurs préférences concernant le choix optimal de l'école pour leurs enfants.

De ce fait la littérature théorique sur les processus d'affectation des élèves mentionne que ces trois critères cités sont loin de pouvoir être vérifiés simultanément.

Abdulkadiroğlu et Sönmez (2003) démontrent que le processus d'affectation des élèves le plus utilisé repose sur le principe des incitations rationnelles (Boston Mechanism⁴¹) par lequel les pouvoirs publics incitent les familles d'élèves à choisir des comportements stratégiques de façon à se retrouver dans une situation d'équilibre sous optimal, car il est qualifié de stable incitant par conséquent à la déclaration sincère des préférences.

En exploitant l'algorithme développé par Gale et Shapley (1962)⁴², les travaux d'évaluation de l'académie de Paris (Hiller et Tercieux, 2014) concernant les processus d'affectation des élèves, montrent que l'extension du choix scolaire et l'assignation des élèves selon des critères boursiers ont engendré une réduction significative des phénomènes de ségrégation sociale à l'école.

En effet, la prise en compte des critères sociaux a permis de réunir les deux objectifs attendus à savoir la généralisation du choix scolaire et l'atteinte des objectifs de la mixité sociale à l'école

⁴¹Le mécanisme de Boston est un processus de placement d'étudiants dans les programmes de choix d'école à travers le monde

⁴²L'algorithme de Gale et Shapley (1962) est considéré parmi les processus combinant à la fois une situation stable et le principe de révélation des préférences. Cet algorithme qualifié de sous optimal a donné lieu à des réformes importantes dans le choix d'un tel processus d'affectation des élèves.

parisienne. Cependant alors que le secteur public est soumis à l'adoption de règles pour l'affectation des élèves, le secteur privé n'est pas concerné.

Du point de vue méthodologique, les problèmes que soulève l'estimation de l'effet de la composition sociale et économique sur la performance scolaire sont multiples. Ils concernent la mesure des indices de composition sociale et économique de l'école, la spécification des modèles économétriques et les biais induits dans la prédiction des acquis scolaires des élèves.

En effet, la mise en évidence de l'effet de composition doit permettre de révéler les écarts de performance des élèves ayant les mêmes caractéristiques en début de la formation, mais qui sont affectés à des écoles qui se différencient selon la composition sociale et économique des élèves.

Il s'agit de mesurer l'effet engendré par les caractéristiques moyennes de tous les élèves du groupe sur le profil de chaque élève, tout en contrôlant la modélisation par les caractéristiques intrinsèques aux élèves et les variables contextuelles liées à l'environnement scolaire. Selon Bruce Sacerdote, (2001), il existe certainement une certaine covariance entre les caractéristiques moyennes des élèves et les caractéristiques propres de chaque élève qui précisément peut être à l'origine d'un biais dans l'estimation de l'effet de composition. Il est possible de remédier à ce problème en utilisant la démarche proposée par Mundlak (1978), en introduisant dans la spécification du modèle à effets aléatoires la moyenne des variables intrinsèques à l'élève (Hanchane et al 2012).

Cependant, Thrupp, Lauder et Robinson (2002) suggèrent une démarche méthodologique qui permet de valider les conditions préalables et les hypothèses de la modélisation économétriques.

Premièrement, l'échantillon d'élèves et d'écoles doit être représentatif de la population. L'échantillon doit donc inclure des écoles issues des zones favorisées et défavorisées selon la méthode d'échantillonnage stratifiée. Deuxièmement, le modèle proposé doit incorporer le maximum de variables individuelles intrinsèques à l'élève afin d'estimer une matrice de variance covariance qui soit la moins biaisée possible (Harker et Tymms, 2004⁴³). Troisièmement, le modèle doit incorporer les variables contextuelles. De ce fait, la performance scolaire sera déterminée à partir de l'impact direct de la composition et d'un impact indirect en relation avec les caractéristiques propres à l'école.

⁴³Discussion des auteurs des Problème de spécification du modèle de prédiction individuelle

Enfin, la coexistence de différents niveaux emboîtés dans les données (élèves, classe, école, etc), justifie une modélisation hiérarchique qui permettra une décomposition de la variance totale en deux composantes : la variance inter établissement et intra établissement afin d'extraire le coefficient de corrélation intra-classe⁴⁴ qui sert d'indicateur pour valider la pertinence de recourir à une telle modélisation.

Les études sur l'effet de composition sont nombreuses, mais l'ampleur de cet effet reste ambiguë. Les résultats sont divergents et cela tient à la présence de biais non contrôlés et à la non consistance des modèles économétriques estimés (biais de spécification ou d'identification). Par exemple, le recours à des modèles reposant sur une spécification en coupe transversale peuvent engendrer des effets non-significatifs voire même contradictoire avec les signes et la validation des hypothèses attendues. La revue de la littérature empirique semble suggérer que les travaux exploitant les dispositifs hiérarchiques semblent les plus pertinents pour une estimation de l'effet propre de composition sociale de l'école sur les apprentissages des élèves.

1.2. Littérature empirique

En exploitant des données longitudinales du système éducatif en Nouvelle Zélande, Lauder et al (1999) ont montré que 40 % de la variance totale est dû aux variables liées à l'environnement scolaire et surtout à la composition sociale et économique de l'école.

Keeves et Saha (1992) ont montré que la relation entre le statut socioéconomique et performance scolaire est beaucoup plus élevé dans les pays développés. Ce résultat a connu une amplification entre les années 1970 et les derniers résultats empiriques.

Duru-Bellat et al. (2004) ont étudié l'effet de la composition sociale et scolaire sur l'amélioration des apprentissages des élèves en France. Les résultats montrent que les effets de composition sur l'apprentissage restent très minimes dans le système éducatif français. En revanche, cet effet de composition est significatif au niveau des élèves de la première année du cours élémentaire (CE1). En fait, la composition sociale de l'école impacte directement les pratiques pédagogiques des enseignants, les attitudes des élèves ainsi que leurs progressions.

⁴⁴L'utilité de la modélisation hiérarchique et la spécification du modèle ont été présentés en détail dans le chapitre 1 de ce rapport de thèse

Le même constat et résultat a été prouvé et démontré par l'OCDE à travers les données du PISA 2015 et 2018.

Van der Slik, Driessen et De Bot (2006) ont tenté de modéliser l'effet de composition sur l'apprentissage des élèves en lecture entre la 4e et la 6e année primaire à travers des variables liées à la composition socioéconomique et ethniques des établissements. Les résultats ont montré que la ségrégation scolaire engendre certainement des inégalités d'acquis scolaires entre les différents groupes d'élèves. De plus, selon cette étude, l'apprentissage des élèves en lecture est lié beaucoup plus à la composition socioéconomique plutôt qu'à la composition ethnique.

L'analyse des écoles secondaires par Opdenakker et Van Damme (2001) montre un effet significatif de la composition des établissements sur l'amélioration des acquis scolaires des élèves en Belgique.

Exploitant la base de données longitudinale du National Educational longitudinal Survey de 1988, Rumberger et Palardy (2005) analysent le processus d'apprentissage en mathématiques, en lecture et en sciences sur une période d'enseignement allant de la 8e année à la 12e année d'enseignement. Sur la base d'un échantillon de 14217 élèves et 913 écoles représentatives des écoles secondaires américaines, les résultats de cette étude montrent que 25% de la variance totale est expliquée par les différences entre les écoles. Par ailleurs, le 1/3 de la variance totale est expliquée par les caractéristiques intrinsèques des élèves. Ainsi, la composition sociale et économique a un effet significatif sur l'amélioration des acquis scolaires des élèves.

De même, cet effet de composition sur l'apprentissage des élèves de l'enseignement primaire du Hampshire en Angleterre a été également traité par Lauder et al. (2007). Recourant à différents indices de composition tenant compte entre autres du sexe, de l'origine ethniques et sociale, les auteurs montrent que la composition des écoles explique significativement la progression des élèves en mathématiques et en lecture. Ainsi, les effets de composition expliquent respectivement 24% et 27 % des écarts d'apprentissage des élèves en lecture et en mathématiques

Vandenberrghe (2001) a révélé que la meilleure performance scolaire des élèves en mathématiques et en sciences est conditionnée par l'appartenance à une classe où les élèves sont issus d'une catégorie sociale élevée. De même, Duru-Bellat (2003) a démontré que la composition sociale de la classe a plus d'impact sur la performance scolaire que l'effet isolé de l'origine sociale des élèves.

Ce résultat a également été confirmé par Crahay et Monseur (2006) à partir de données sur un large panel de pays, et d'une méthodologie en deux étapes. Selon ces auteurs, l'impact de la composition sociale de l'école sur la performance des élèves en lecture est bien observé dans presque l'ensemble des pays de l'échantillon étudié mais avec une certaine variabilité entre les pays. Le résultat le plus remarquable est que l'impact de la composition sociale de la classe est très minime dans les pays où les systèmes éducatifs sont les plus performants comme la Finlande et le Danemark. (Crahay et Monseur, 2006).

D'après les conclusions de la littérature théorique et empirique, il apparaît clairement que les phénomènes d'agrégation sociale sont très nuisibles aux élèves ayant des difficultés d'apprentissage. Ainsi, plus un système éducatif produit des phénomènes d'agrégation sociale, plus il augmente les inégalités entre les différents groupes d'élèves et plus il accroît les inégalités sociales entre les différentes écoles.

En effet, exploitant les données de PISA, TIMSS et PIRLS, Crahay (2000 et 2007) met en évidence une corrélation significative entre l'égalité du système éducatif et son efficacité. Ce résultat est également confirmé par Grisay (1993) qui exploite des données longitudinales afin de vérifier l'hypothèse d'agrégation sociale sur l'amélioration des apprentissages des élèves du cycle primaire et du collège en France. Les conclusions de cette étude révèlent que plus l'environnement scolaire est formé de groupes hétérogènes d'élèves, plus il est bénéfique à l'amélioration des apprentissages de tous les élèves⁴⁵. De plus, au cours des deux années qui ont suivi l'évaluation menée auprès des élèves au collège, Grisay montre que les profils des élèves tendent à se ressembler lorsque les classes ont été fréquentées par des élèves ayant des caractéristiques socioéconomiques différentes. Au contraire, les classes homogènes engendrent plus d'inégalité d'apprentissage entre les élèves notamment au collège.

Ainsi, plus les politiques éducatives sont caractérisées par les phénomènes d'agrégation plus les différences de performance entre les différents groupes sociaux d'élèves seront importantes.

Généralement, les récents travaux menés dans plusieurs pays confirment également l'hypothèse d'une corrélation négative entre l'apprentissage des élèves et les phénomènes d'agrégation sociale des systèmes éducatifs (OCDE-PISA 2015 et 2018).

⁴⁵Ce résultat est observé sur l'ensemble de disciplines évaluées par l'auteur à savoir les mathématiques et le français

A partir d'un panel de 28 pays, Monseur et Marcel Crahay (2008) ont réussi à produire un travail de recherche quasiment exhaustif, examinant l'effet de la composition sociale des écoles sur la performance des élèves dans plusieurs systèmes éducatifs. Dans cette perspective, les auteurs ont regroupé les écoles selon le statut social et économique des élèves en trois catégories appelées « écoles favorisées », « écoles mixtes » et « écoles défavorisées ». L'exploitation de la base de données PISA 2006 a montré que les écoles mixtes sont plus concentrées dans les pays de l'Europe du Nord. En revanche, dans certains systèmes éducatifs comme ceux de la France, la Belgique, et du Luxembourg, les écoles mixtes représentent moins de 40% du total des écoles. Les résultats des estimations économétriques montrent que la faible performance scolaire des élèves est due essentiellement à l'appartenance à une école où le pourcentage des élèves issus des milieux défavorisés est élevé, quel que soit le statut socioéconomique de l'élève. Par contre, la meilleure performance scolaire des élèves est conditionnée par la fréquentation d'une école dont le pourcentage d'élèves favorisés est important, quel que soit le statut socioéconomique de l'élève. Concernant les écoles mixtes, les phénomènes d'hétérogénéité sont plutôt bénéfiques pour les élèves issus des milieux défavorisés, tandis que les élèves favorisés n'améliorent pas leurs apprentissages de manière significative. Concernant les deux disciplines évaluées à savoir la lecture et les sciences), Crahay et Monseur (2008) soulignent que les phénomènes d'agrégation sociale sont parmi les déterminants les plus significatifs de l'augmentation des inégalités de performance entre les élèves défavorisés et les élèves favorisés.

Les résultats de Crahay et Monseur (2008) confirment ainsi l'existence d'une relation étroite entre la composition sociale et économique des écoles et la performance scolaire des élèves. Cependant, l'impact des phénomènes d'agrégation sur l'apprentissage des élèves varie selon les systèmes éducatifs. Il est présent en Allemagne, en France, en Belgique, en Italie, au Japon, au Portugal, au Luxembourg, aux États-Unis, et dans d'autres pays. Les seuls pays ne recourant pas aux politiques d'agrégation éducative sont la Finlande, l'Islande, la Norvège et la Suède.

Or, les systèmes éducatifs appliquant les politiques de regroupement des élèves selon leur statut socioéconomique sont les systèmes éducatifs les plus inégalitaires en termes de performance scolaire.

Dans l'ensemble, la littérature empirique montre que plus le système éducatif est orienté vers les politiques d'homogénéité des élèves selon leur statut socioéconomique, moins il sera bénéfique à tous les élèves et plus les inégalités de performance scolaire seront très prononcées

entre les différents groupes sociaux d'élèves. De ce fait, l'équité et l'efficacité d'un système éducatif passe directement par la minimisation des phénomènes de ségrégation (Crahay et Monseur, 2008).

De même, il est évident de constater que l'école n'est qu'une structure sociale dans un espace plus large, qui peut être le quartier ou bien la ville, ce qui suggère la nécessité de tenir compte des paramètres de l'environnement scolaire.

La section suivante sera dédiée à la méthodologie adoptée pour l'estimation du modèle économétrique.

2. Méthodologie d'estimation

Dans cette section, nous décrivons comment un modèle semi-paramétrique connu dans la littérature théorique des modèles en coupe instantanée, peut être généralisé à une situation où les données sont structurées en multiniveaux. Plus précisément, nous proposons une généralisation des modèles semi-paramétriques en multiniveaux sans imposer une forme spécifique à notre fonction d'intérêt censée capturer les effets de pairs.

Cette proposition est un essai pour revisiter la fonction de production du capital humain, à travers une forme fonction non linéaire. Il s'agit de la relation de non linéarité entre le degré de composition sociale et économique à l'école et la performance scolaire des élèves. Pour ce faire, nous proposons une estimation économétrique en faisant appel aux estimateurs non paramétrique et semi-paramétrique fondamentalement basés sur la fonction de Kernel.

Sans imposer aucune forme fonctionnelle sur la relation du modèle à estimer, cette technique renvoie premièrement à l'exploitation des variables observées, avec bien évidemment un intérêt primaire celui de l'absence du nombre de paramètre estimé, et peu d'hypothèses à tester comme il est connu et débattu dans la littérature sur les modèles de régression paramétrique. Ainsi, $f(z)$ devient la fonction de lien dans un modèle non et semi-paramétrique, et qui remplace le paramètre $\hat{\beta}$ d'une estimation avec GLS⁴⁶, OLS⁴⁷ ou IV⁴⁸.

⁴⁶ Moindres carrées généralisées

⁴⁷ Moindres carrées ordinaires

⁴⁸ Variables instrumentales

Premièrement, on va consacrer cette section à la description méthodologique des modèles non paramétriques. Cette dernière concerne l'estimateur de Kernel avec la discussion du travail pionnier de Nadaraya et Watson (1964), ainsi que la contribution de Cleveland (1979) sur le lissage local polynomial. Ensuite, nous exposons la technique d'estimation et les étapes à suivre pour l'estimation de la courbe non-paramétrique et semi-paramétrique dans un cadre de données hiérarchique.

La contribution de Nadaraya et Watson (1964) pour l'estimation d'un modèle non paramétrique a une inspiration directe des travaux sur le lissage par moyenne mobile.

En effet, la méthode de moyenne mobile souffre de plusieurs limites, car la pondération des points z_i est souvent la même. En revanche, l'estimateur Kernel permet de calculer des pondérations relatives à la distance observée entre z_i et z . Cette technique est basée sur un paramètre de lissage appelé aussi Bandwidth parameter. Techniquement parlant, le poids associé à une estimation d'une fonction de lien $g(z)$ est donnée par :

$$w_i(z) = \frac{K([z_i - z]/\lambda)}{\sum_{j=1}^N K([z_j - z]/\lambda)} \quad (47)$$

avec λ est le paramètre de lissage (Bandwidth parameter)

Ainsi, l'estimateur Kernel de la fonction de lien $g(z)$ est donnée par :

$$\hat{g}(z) = \sum_{i=1}^N w_i(z) y_i \quad (48)$$

Une seconde technique repose sur le lissage local polynomial. Cette dernière consiste à estimer un modèle de régression multiple à chaque voisinage $V(z)$ d'un point z . De plus, cette technique développée par Cleveland (1979) permet l'introduction de plusieurs composantes en puissances quadratique et cubique. Celle-ci se formalise comme suit :

$$\left\{ \begin{array}{l} \hat{g}(z) = \hat{\beta}_1(z) + z\hat{\beta}_2(z) \\ \{\hat{\beta}_1(z), \hat{\beta}_2(z)\} = \underset{\{\beta_1(z), \beta_2(z)\}}{\text{MIN}} \sum_{z_i \in V(z)} [y_i - \beta_1(z) - z_i\beta_2(z)]^2 \end{array} \right. \quad (49)$$

Cette méthode a comme objectif la minimisation des résidus quadratique comme la méthode des moindres carrés ordinaire (MCO), mais en tenant compte de chaque voisinage du point z , afin d'identifier les paramètres $\beta_1(z)$. Par conséquent, les paramètres estimés auront des valeurs

différentes de celles issues des moindres carrés ordinaires, puisqu'ils reposent sur une variation selon un point de référence z .

La régression « lowess » développée par Cleveland et Devlin (1988) généralise la proposition faite par Cleveland (1979) en pondérant chaque point z_i par une densité de type Kernel. Cette dernière peut être de type Epanchnikov, Cosine, normal ou autres. Les résultats de paramétrages des $\beta_1(z)$ et $\beta_2(z)$ de la régression polynomiale locale sont identifiés par la programmation suivante :

$$\left\{ \begin{array}{l} \hat{g}(z) = \hat{\beta}_1(z) + z\hat{\beta}_2(z) \\ \{\hat{\beta}_1(z), \hat{\beta}_2(z)\} = \underset{\{\beta_1(z), \beta_2(z)\}}{\text{MIN}} \sum_{z_i \in V(z)} K\left(\frac{[z_i - z]}{\lambda}\right) [y_i - \beta_1(z) - z_i\beta_2(z)]^2 \end{array} \right. \quad (50)$$

Cette littérature théorique des modèles non et semi-paramétriques a connu par la suite une autre extension, avec une spécification dans le cadre de l'économétrie des données de panel (Rodríguez-Poo et Soberon, 2017 ; Lee et al., 2017). Li et Stengos (1996) Li et Ullah (1998). Ces derniers sont d'ailleurs les premiers à avoir développé cette généralisation aux données longitudinales.

Henderson et al. (2008) ont par la suite développé une méthode d'estimation itérative ⁴⁹ des modèles non paramétriques et semi-paramétriques sur des données longitudinales que Zhou et Li (2011) ont étendu au cas de panels non cylindrés.

Nous nous appuyons sur la méthodologie de Zhou et Li (2011), car leur méthode ainsi que l'estimateur qui en découle nous semblent être les plus en adéquation avec notre matrice de données qui se caractérise par une structure de données hiérarchisés.

Nous devons ainsi estimer une relation non paramétrique et semi-paramétriques avec ($p = 1$ et $p > 1$)⁵⁰ afin de tester la non linéarité existante entre la composition sociale et économique de l'école et la performance scolaire des élèves.

$$y_{ij} = g(X_{ij}) + u_j + \varepsilon_{ij} \quad i = 1, 2, \dots, n ; j = 1, 2, \dots, k \quad (1) \quad (51)$$

$$y_{ij} = g(X_{ij}) + g(Z'_j) + Z''_j\gamma + u_j + \varepsilon_{ij} \quad i = 1, 2, \dots, n ; j = 1, 2, \dots, k \quad (2) \quad (52)$$

⁴⁹La méthode d'estimation itérative des paramètres est utilisée en présence des données à valeurs manquantes. Elle équivalente à une méthode d'imputation multiple dans les modèles linéaires.

⁵⁰P désigne la puissance quadratique et cubique à travers laquelle on teste la présence d'un tel effet non linéaire

Les spécifications (51) et (52) correspondent respectivement aux modèles non paramétrique et semi-paramétrique avec des données en structure hiérarchique.

Afin d'estimer ce modèle, nous nous appuyons sur la méthode développée par Rodriguez-Poo et Soberon (2017).

Nous désignons par y la variable dépendante et X le vecteur de p variables explicatives où la fonction de lien $G(\cdot)$ qui relie le vecteur X avec la variable y est une fonction non spécifiée à estimer définit $\mathfrak{R}^p \rightarrow \mathfrak{R}$. Pour le cas du modèle semi-paramétrique, « q » d'autres variables de contrôle Z sont considérées dont γ est un vecteur de q paramètres à estimer. Nous considérons le cas d'une matrice non cylindrée où chaque école « j » dispose de m_j observations. Les effets individuels u_j sont supposés fixes et corrélés avec X où la forme de cette corrélation est non spécifiée. Les termes d'erreurs ε_{ij} sont supposés être i.i.d, de moyenne nulle et de variance égale σ_ε^2 où $E(\varepsilon_{ij}|X_{ij}) = 0$. On note I_k la matrice identité de dimension k et e_k un vecteur

unitaire $k \times 1$. Si on définit $\tilde{\varepsilon}_i = (\tilde{\varepsilon}_{i2}, \dots, \tilde{\varepsilon}_{im_i})'$ où $\tilde{\varepsilon}_{ij} = \varepsilon_{ij} - \varepsilon_{j1}$, on peut exprimer la matrice de covariance de $\tilde{\varepsilon}_j$ « Σ_j » et son inverse « Σ_j^{-1} » en fonction de σ_ε^2 comme suit :

$$\begin{cases} \Sigma_j = \sigma_\varepsilon^2 (I_{m_j-1} + e_{m_j-1} e_{m_j-1}') \\ \Sigma_j^{-1} = \sigma_\varepsilon^{-2} (I_{m_j-1} - e_{m_j-1} e_{m_j-1}' / m_j) \end{cases} \quad (53)$$

Si on note $g_{ij} = g(z_{ij})$, le modèle (1) devient donc $y_{ij} = g_{ij} + u_j + \varepsilon_{ij}$. Pour le cas où « $i=1$ » on a $y_{i1} = g_{i1} + u_j + \varepsilon_{i1}$.

Si on note de plus $\tilde{y}_j = (\tilde{y}_{j2}, \dots, \tilde{y}_{jm_j})'$ et $g_j = (g_{j2}, \dots, g_{jm_j})'$ où $\tilde{y}_{ij} = y_{ij} - y_{j1}$, on peut exprimer $\tilde{\varepsilon}_j$ en fonction de \tilde{y}_j , g_j et g_{j1} comme suit :

$$y_{ij} = g(z_{ij}) + u_j + \varepsilon_{ij} = g_{ij} + u_j + \varepsilon_{ij} \quad (54)$$

$$\Rightarrow \tilde{y}_{ij} + y_{j1} = g_{j1} + u_j + \tilde{\varepsilon}_{ij} + \varepsilon_{j1} \quad (55)$$

$$\Rightarrow \tilde{y}_{ij} = g_{ij} - g_{j1} + \tilde{\varepsilon}_{ij} \quad (56)$$

$$\Rightarrow \tilde{y}_j = g_j - g_{j1} e_{m_j-1} + \tilde{\varepsilon}_j \quad (57)$$

D'où $\tilde{\varepsilon}_j = \tilde{y}_j - g_j + g_{j1} e_{m_j-1}$

A l'instar de Wang (2003), Lin et Carroll (2006) et Henderson et al. (2008), nous utilisons l'approche itérative basée sur la vraisemblance profilée appelée aussi « vraisemblance de profile ». Nous développons dans un premier temps, la dérivée première de la vraisemblance individuelle « $L_j(\cdot)$ » par rapport à g_{ij} notée $L_{ij}^g = \partial L_j(\cdot) / \partial g_{ij}$. Ainsi on peut développer la vraisemblance individuelle à partir de la formulation de $\tilde{\varepsilon}_j$ comme suit :

$$L_j(\cdot) = -\frac{1}{2} \tilde{\varepsilon}_j' \Sigma_j^{-1} \tilde{\varepsilon}_j \quad , \quad j = 1, 2, \dots, k. \quad (58)$$

$$L_i(\cdot) = -\frac{1}{2} (\tilde{y}_i - g_i + g_{i1} e_{m_i-1})' \Sigma_j^{-1} (\tilde{y}_j - g_j + g_{j1} e_{m_j-1}) \quad , \quad j = 1, 2, \dots, k. \quad (59)$$

$$L_{ij}^g = \frac{\partial L_j(\cdot)}{\partial g_{ij}} = \begin{cases} -e'_{m_j-1} \Sigma_j^{-1} (\tilde{y}_j - g_j + g_{j1} e_{m_j-1}) \quad , \quad i = 1 \\ c'_{j,i-1} \Sigma_j^{-1} (\tilde{y}_j - g_j + g_{j1} e_{m_j-1}) \quad , \quad i \geq 2 \end{cases} \quad (60)$$

où $c_{j,i-1}$ est un vecteur de dimension $(m_i - 1) \times 1$ dont tous les éléments sont nuls à l'exception du $(i - 1)$ ème élément qui est égale à 1

Si on définit

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} g(z) \\ \partial g(z) / \partial z \end{pmatrix} = \begin{pmatrix} g(z) \\ g^{(1)}(z) \end{pmatrix} \quad \text{et} \quad G_{ij} = \begin{pmatrix} 1 \\ (z_{ij} - z) / h \end{pmatrix}$$

L'estimation de $\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix}$ se fait en résolvant la condition du premier ordre de la vraisemblance profilée d'une manière itérative comme suit :

$$\sum_{j=1}^n \frac{1}{m_i} \sum_{i=1}^{m_j} K_h(z_{ij} - z) G_{ij} L_{ij}^g \left(\hat{g}_{[l-1]}(z_{j1}), \dots, G_{ij}(\alpha_0, \alpha_1)', \dots, \hat{g}_{[l-1]}(z_{imj}) \right) = 0 \quad (61)$$

où $\hat{g}_{[l-1]}(z_{ij})$ est l'estimation de $g(z_{is})$ pour la $(l - 1)$ ème itération et $k_h(v) = h^{-1} k(v/h)$ et $k(\cdot)$ est la fonction de Kernel.

Nous pouvons par la suite définir l'estimation pour la l ème itération en fonction de la $(l - 1)$ ème itération :

$$\begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} = \begin{pmatrix} \hat{g}_{[l]}(z) \\ \hat{g}_{[l]}^{(1)}(z) \end{pmatrix} = \frac{(A_1 + A_2)}{A_3}$$

Telle que :

$$\begin{cases} A_1 = \sum_{j=1}^k \frac{1}{m_j} \left(e'_{m_j-1} \Sigma_j^{-1} e_{m_j} K_h(z_{j1} - z) G_{j1} \hat{g}_{[l-1]}(z_{j1}) + \sum_{i=2}^{m_j} c'_{j,i-1} \Sigma_j^{-1} c_{j,i-1} K_h(z_{ij} - z) G_{ij} \hat{g}_{[l-1]}(z_{ij}) \right) \\ A_2 = \sum_{j=1}^k \frac{1}{m_j} \left(-K_h(z_{j1} - z) G_{j1} e'_{m_j-1} \Sigma_j^{-1} H_{j,[l-1]} + \sum_{i=2}^{m_j} K_h(z_{j1} - z) G_{ij} c'_{j,i-1} \Sigma_j^{-1} H_{j,[l-1]} \right) \\ A_3 = \sum_{j=1}^k \frac{1}{m_j} \left(e'_{m_j-1} \Sigma_j^{-1} e_{m_j-1} K_h(z_{j1} - z) G_{j1} G'_{j1} + \sum_{i=2}^{m_j} c'_{j,i-1} \Sigma_j^{-1} c_{j,i-1} K_h(z_{ij} - z) G_{ij} G'_{ij} \right) \end{cases} \quad (62)$$

où $H_{j,[l-1]}$ est un vecteur de dimension $(m_j - 1) \times 1$ dont les éléments notés « $h_{i_s,[l-1]}$ » sont tels que $h_{i_s,[l-1]} = (\tilde{y}_{ij} - (\hat{g}_{[l-1]}(z_{ij}) - \hat{g}_{[l-1]}(z_{j1})))$, $i = 1, 2, \dots, m_j$.

L'estimateur initial de $g(\cdot)$ est obtenu en se basant sur données de l'école tandis que la dernière itération est retenue dès lors que le critère de convergence est vérifié :

$$\frac{\sum_{j=1}^k \frac{1}{m_j} \sum_{i=2}^{m_j} (\hat{g}_{[l]}(z_{ij}) - \hat{g}_{[l-1]}(z_{ij}))^2}{\sum_{j=1}^k \frac{1}{m_j} \sum_{i=2}^{m_j} \hat{g}_{[l-1]}^2(z_{ij})} \leq 0.01 \quad (63)$$

Par ailleurs la variance σ_ε^2 est estimée par :

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{2n} \sum_{j=1}^k \frac{1}{m_j - 1} \sum_{i=2}^{m_j} (y_{ij} - y_{j1} - (\hat{g}_{[l]}(z_j) - \hat{g}_{[l-1]}(z_{ij})))^2 \quad (64)$$

La variance de l'estimateur $\hat{g}(z)$ est calculé par : $\kappa (nh\hat{\Omega}(z))^{-1}$

$$\text{où } \kappa = \int k^2(v) dv \quad \text{net} \quad \hat{\Omega}(z) = \sum_{j=1}^k \frac{m_j - 1}{m_j} \sum_{i=2}^{m_j} K_h(z_{j1} - z) / \hat{\sigma}_\varepsilon^2 \quad (65)$$

Pour l'estimation du modèle semi-paramétrique, on définit l'estimateur non paramétrique des « q » variables de contrôle $\hat{g}_x(\cdot) = (\hat{g}_{x,1}(\cdot), \dots, \hat{g}_{x,q}(\cdot))$ et de la variable dépendante $\hat{g}_y(\cdot)$ définit l'estimateur. Ainsi l'estimation de \mathcal{Y} qui est de dimension $q \times 1$ est :

$$\hat{\mathcal{Y}} = \left(\sum_{j=1}^k \frac{\tilde{x}_j^* \Sigma_j^{-1} \tilde{x}_j^*}{m_j} \right)^{-1} \left(\sum_{j=1}^k \frac{\tilde{x}_j^* \Sigma_j^{-1} \tilde{y}_j^*}{m_j} \right) \quad (66)$$

où \tilde{x}_{js^*} et \tilde{y}_{js^*} sont respectivement des matrices de dimension $(m_j - 1) \times q$ et $(m_j - 1) \times 1$ telles que la sème ligne est définie par : $\tilde{x}_{js^*} = \tilde{x}_{js} - (\hat{g}_x(z_{js}) - \hat{g}_x(z_{j1}))$ et $\tilde{y}_{js^*} = \tilde{y}_{js} - (\hat{g}_y(z_{js}) - \hat{g}_y(z_{j1}))$. La composante non-paramétrique du modèle semi-paramétrique est déduite en remplaçant \tilde{y}_{ij} par $\tilde{y}_{ij} - x'_{ij}\hat{\gamma}$.

Cependant, l'introduction des variables de contrôle va nous obliger d'estimer un modèle semi-paramétrique, tandis que l'estimation d'un modèle non paramétrique exige l'existence seulement de la partie non paramétrique contrôlée par C_j (Composition sociale et économique de l'école).

$$\{ y_{ij} = \beta(X_{ij}) + \gamma(Z_j) + u_j + \varepsilon_{ij} \quad (67)$$

$$\{ y_{ij} = \beta(X_{ij}) + g(C_j) + \gamma(Z_j) + u_j + \varepsilon_{ij} \quad (68)$$

Et donc le modèle que nous cherchons à estimer se compose d'une partie paramétrique associée aux variables de l'élève, et deux composantes des variables de l'école. La première est une partie non paramétrique identifiée à partir de l'estimation de la fonction $g(\cdot)$ et une partie paramétrique identifiée à travers le paramètre γ .

3. Données utilisées et faits stylisés

Nous proposons d'utiliser la méthodologie que nous venons de présenter afin d'estimer l'impact de la composition sociale et économique de la classe sur la performance scolaire des élèves de la 4^e année primaire en lecture et de la 2^e année collégiale en mathématiques et en sciences à partir des données TIMSS et PIRLS. Ces deux dispositifs d'évaluation comportent une variable indiquant la composition sociale et économique de l'école qui correspond au pourcentage des élèves issus d'une classe défavorisée au sein de l'école. Ainsi, à partir d'une approche purement descriptive, la décomposition des scores des élèves selon cette variable nous permet au préalable de rendre compte des écarts de performance entre les différentes catégories sociales à savoir « classes défavorisées », « classes mixtes » et « classes favorisées ».

Afin de contrôler l'effet de composition sociale et économique de l'école sur les performances scolaires, nous avons sélectionné un certain nombre de variables de contrôle en nous appuyant sur la littérature existante.

Ainsi, les tableaux 47 et 48 en annexe présentent les données utilisées pour l'évaluation de l'impact de la composition sociale et économique de l'école sur la performance scolaire en lecture des élèves marocains de la 4^e année primaire (PIRLS 2011) et en sciences et mathématiques pour les élèves de la 2^e collégiale (TIMSS 2011).

Les résultats obtenus dans le chapitre 1 ont permis de montrer que la composition sociale et économique de l'école représentait un facteur déterminant de l'apprentissage des élèves. Ce résultat a été obtenu à partir d'une modélisation multiniveau avec correction du biais d'endogénéité. Dans ce qui vient, nous examinons la relation entre les résultats scolaires des élèves selon la composition sociale et économique des établissements, tout en creusant davantage dans l'analyse. En fait, le présent chapitre ne cherche pas seulement à vérifier l'existence d'un effet causal entre la composition sociale et économique de l'établissement et la performance scolaire, mais aussi il a pour objectif de déceler les contextes dans lesquels un degré de mixité sociale à l'école peut contribuer à une amélioration généralisée de la performance scolaire.

Afin de rendre compte des écarts de performance scolaire des élèves selon la composition sociale et économique de l'école, nous avons choisi de classer cette dernière en quatre modalités. La variable composition sociale et économique de l'école est déduite directement des informations collectées auprès du directeur de l'établissement. En fait, il s'agit des réponses fournies à la question suivante : quel est le pourcentage des élèves issus d'une classe défavorisée au sein de votre établissement. Le premier type d'école « écoles favorisées » correspond à un pourcentage d'élèves défavorisés ne dépasse pas les 10%. Le second type « école moyennant mixte » se caractérise par un pourcentage de 11 à 25 % d'élèves issus d'une classe défavorisée.

Le troisième type d'école « école mixte » concerne un pourcentage d'élèves issus d'un milieu défavorisé variant entre 26 à 50%. Et enfin, une école est qualifiée « d'école défavorisée » si elle comporte un pourcentage de plus de 50% des élèves issus d'un milieu défavorisé.

Tableau 41. Répartition des écoles selon leur composition sociale

	4 ^e année primaire en %	2 ^e année collège en %
Favorisée : -10%	12 %	6%
Mixte moyen : 11-25%	9%	9%
Mixte 25-50%	14%	17%

Défavorisée : +50%	65%	68%
---------------------------	-----	-----

Source : D'après l'auteur à partir des données TIMSS et PIRLS 2011

Le tableau 41 présente pour les deux années d'études, la répartition des écoles selon la composition sociale des élèves. Nous observons que la répartition des écoles est comparable dans le primaire et au collège. Ainsi, pour la 4e année primaire, les écoles défavorisées représentent 65% de la totalité des écoles. Les écoles favorisées ne représentent que 12%, tandis la mixité concerne moins d'un quart des écoles. Une répartition quasiment comparable peut être observée au collège où 68% des écoles se caractérisent par un pourcentage élevé d'élèves issus de milieu défavorisé.

3.1. Les écarts d'acquis scolaire en lecture à l'école primaire selon la composition sociale et économique de la classe

Le tableau 42 présente les écarts de performance scolaire en lecture selon un croisement des variables intrinsèques à l'élève et contextuelles. Le choix des variables est justifié à travers nos résultats du 1^e chapitre sur les déterminants de la performance scolaire au Maroc.

Tableau 42. Score en lecture et composition sociale et économique de l'école

Score Variables	Favorisée	Mixte moyen	Mixte	Défavorisée
Moyenne	350	340	340	300
Genre	Garçon	340	330	290
	File	350	345	310
Retard Scolaire	<= 1 an	385	350	340
	Plus de 2 années	360	340	290
Préscolarisation	Oui	345	340	310
	Non	300	280	290
Absentéisme des élèves	Simple problème	420	400	355
	Grave problème	360	330	300
	Sans diplôme	370	350	295
	Moins que secondaire	390	380	360
Education des parents	Secondaire et plus	430	390	350
	<=20 ans	420	400	350
Expérience de l'enseignant	>20 ans	390	360	330
	Simple problème	410	380	360
Charge horaire	Grave problème	360	340	330
	Urbain	380	360	330
Milieu de l'établissement	Rural	330	320	280

Source : D'après l'auteur à partir des données PIRLS 2011

Les élèves issus d'une école favorisée performant beaucoup mieux que ceux issus d'un établissement scolaire défavorisé avec un écart de performance s'élevant à 50 points. Les élèves issus d'une école mixte enregistrent un score de 340 points, inférieur seulement de 10 points au score observé chez les élèves issus d'une école favorisée.

Les filles obtiennent de meilleurs résultats que les garçons quelle que soit la composition sociale et économique de l'école. Cependant les écarts de performance entre les écoles favorisées et défavorisées sont quasiment comparables pour les filles et les garçons.

De même, le fait d'avoir bénéficié d'une préscolarisation contribue à de meilleurs résultats pour les élèves. La préscolarisation a d'autant plus d'incidence sur les résultats que l'élève suit sa scolarité dans une école favorisée. Ainsi, les écarts de performance entre écoles favorisées et défavorisées s'élèvent à 35 points pour les élèves ayant été préscolarisés. Il est intéressant de noter que la fréquentation d'une école mixte permet une performance proche de celle obtenue dans une école favorisée.

Ces résultats descriptifs indiquent que de manière générale, la performance scolaire des élèves est très corrélée avec l'appartenance à une école à composition favorisée, et ceci quel que soit le contexte (urbain-rural ; expérience de l'enseignant ; éducation des parents) ou les difficultés auxquelles peuvent être confrontés les enseignants (charge horaire, absentéisme des élèves) (Tableau 42).

3.2. Les écarts d'acquis scolaire en mathématique et en science selon la composition sociale et économique de la classe

Le tableau 43 présente les moyennes de score selon le type de composition sociale et économique et les mêmes caractéristiques que nous avons sélectionnées précédemment pour la lecture.

Tableau 43. Scores en mathématique/science et composition sociale et économique de l'école

Score variables	Mathématique				Science			
	Favorisée	Mixte moyen	Mixte	Défavorisée	Favorisée	Mixte moyen	Mixte	Défavorisée
Moyenne	440	380	380	367	444	384	385	364

Genre	Garçon	430	390	384	370	427	390	387	372
	Fille	430	380	378	370	432	383	385	367
Retard Scolaire	<= 1 an	500	440	420	410	480	460	420	360
	>=2 ans	370	350	320	300	350	340	290	280
Absentéisme des élèves	Simple problème	460	430	390	350	450	405	380	370
	Grave problème	410	380	320	310	340	320	290	280
Genre de l'enseignant	F	500	370	380	375	450	400	398	380
	M	440	390	385	358	430	385	380	370
Expérience de l'enseignant	<=20 ans	480	420	390	320	390	380	380	340
	>20 ans	500	390	390	370	430	410	380	360
Milieu de l'établissement	Urbain	450	410	370	350	460	430	420	380
	Rural	350	320	320	280	340	300	280	270
Disponibilité du laboratoire	Oui	-	-	-	-	430	385	385	367
	Non	-	-	-	-	390	386	384	370

Source : D'après l'auteur à partir des données TIMSS 2011

Nous observons que les élèves sont moins performants s'ils fréquentent une école défavorisée plutôt qu'une école favorisée. Ainsi, les écarts de performance sont de 73 points et de 80 points respectivement pour les mathématiques et les sciences. Ces résultats restent également valables pour les autres caractéristiques identifiées. Il en résulte que manifestement le contexte au sein duquel l'élève évolue engendre des écarts de performance entre écoles favorisées et écoles défavorisées dans les deux matières. Cet écart se creuse d'avantage lorsque la composition sociale et économique de l'école intervient comme élément d'analyse des différences de performance scolaire des élèves (tableau 43).

4. Résultats de la modélisation

Cette section a pour objectif d'interpréter et d'analyser les résultats issus du modèle semi-paramétrique global et des modèles semi-paramétriques catégoriels conditionnellement au contexte d'apprentissage de l'élève. Plus précisément, les modèles estimés sont les modèles (67) et (68) que nous avons présentés dans la section 2. Ces modèles ont été estimés pour les trois matières à savoir : lecture, mathématiques et sciences. Nous commençons par l'interprétation des résultats obtenus pour les scores en lecture, puis les mathématiques et enfin les sciences. Ainsi, l'objectif primordial du chapitre est de vérifier si la mixité sociale est une politique publique pertinente pour améliorer les acquis scolaires des élèves marocains. Cette approche s'avère utile pour rendre compte de la situation de la mixité sociale au sein des établissements scolaires, et pour proposer des politiques publiques qui pourraient être mise en œuvre pour améliorer la qualité du système éducatif marocain.

Par souci de comparaison, nous avons estimé 3 modèles. Le premier est un modèle multiniveaux paramétrique, le second est un modèle semi-paramétrique, et le dernier un modèle multiniveaux

semi-paramétrique. Les tableaux (44 ,45 et 46) en annexe présentent l'ensemble des résultats. Dans ce qui vient, nous focalisons notre attention sur l'interprétation du modèle validé par le test d'Hausman. Il s'agit du modèle multiniveaux semi-paramétrique conditionnel aux contextes, et à travers lequel nous représentons les courbes simulées pour chacune des matières évaluées. En fait, nous avons estimé en premier un modèle non paramétrique pour identifier la relation entre les niveaux de compositions sociale et économique de l'établissement et la performance scolaire. Par la suite, cette courbe estimée a été revisitée à travers l'estimation d'un modèle semi-paramétrique conditionnel au contexte afin d'identifier la relation composition sociale et économique de l'établissement et performance scolaire pour chaque contexte ou l'élève est assigné (modèle 68).

Pour ce faire, l'estimation des courbes globales et des courbes contextualisés permettront d'identifier la composition sociale et économique de l'école à partir de laquelle les scores chutent de manière significative. L'encadré 5 présente les étapes d'estimation et de simulation des courbes, et faciliter la démarche d'interprétation des résultats. Cet encadré est valable pour l'ensemble des matières (lecture, mathématiques et sciences).

Encadré 5 : Etape d'estimation du modèle et simulation des courbes

1. Estimation du modèle paramétrique
2. Estimation du modèle non paramétrique
3. Estimation du modèle semi-paramétrique avec la composition sociale et économique comme la partie non paramétrique
4. Validé le modèle à travers un test d'Hausman (Modèle non paramétrique vs modèle semi paramétrique)
5. Simulation de la courbe du modèle non paramétrique
6. Simulation de la courbe du modèle semi paramétrique
7. Simulation des courbes du modèle semi paramétrique conditionnelle au contexte. Il s'agit de déterminer pour chaque variable contextuelle, l'allure de la courbe reliant la composition sociale et économique de l'établissement et la performance scolaire

4.1. Résultats pour les performances en lecture

Les résultats de notre modèle non-paramétrique révèlent que le score des élèves décroît en fonction du type de la composition sociale et économique à l'école. La performance scolaire est plus élevée chez les élèves appartenant à une école favorisée avec un score de 380 points. Ce score diminue pour les élèves fréquentant les établissements de composition mixte avoisinant un score de 360 à 370 points. Les élèves des écoles défavorisées enregistrent un score seulement de 300 points. Ce résultat montre clairement que la composition sociale mixte est beaucoup plus bénéfique pour une performance scolaire généralisée pour l'ensemble des élèves du système éducatif marocain (Graphique 33.a).

Ce résultat qui reste général mérite d'être approfondi pour vérifier s'il reste valable selon le contexte de l'élève. Pour ce faire, les graphiques 33b à 33h présentent les résultats associés à différents contextes.

Par rapport aux problèmes d'absentéisme des enseignants, nos résultats indiquent que les élèves affectés à un établissement favorisé réussissent beaucoup mieux lorsque ce problème est déclaré comme absent ou simple. En effet, si on considère uniquement le groupe des enseignants qui ne s'absentent jamais, le score des acquis scolaires en lecture des élèves est de 400 et 300 respectivement pour les écoles favorisés et défavorisés. Les élèves des écoles mixtes réussissent d'autant mieux que la mixité sociale est importante avec un pourcentage de 25 à 50% des élèves issus des classes défavorisées dans l'établissement (Graphique 33.b).

De même, selon le graphique 33.d illustrant l'effet de la charge horaire des enseignants, les courbes estimées montrent que les enseignants déclarant subir un grave problème de la charge horaire enregistrent des scores médiocres quelle que soit la composition sociale et économique de l'école fréquentée. Par conséquent, les performances ne se distinguent pas entre les élèves d'une école favorisée, mixte ou défavorisée lorsqu'un problème de charge horaire est présent. En revanche, les performances scolaires se dégradent en passant d'une école mixte à une école défavorisée pour les enseignants ayant déclaré l'absence d'un problème de charge horaire. Autrement dit, même si la charge horaire n'est pas un problème pour les enseignants, la performance scolaire des élèves détériore pour ceux appartenant à une école défavorisée (Graphique 33.d).

S'agissant de l'expérience des enseignants, les élèves performant mieux lorsqu'ils sont affectés à des écoles ayant une composition mixte ou favorisée comparativement aux élèves assignés à des écoles défavorisées. Ce résultat n'est valable que pour une certaine catégorie des enseignants ayant moins de 10 ans d'expérience. En fait, les scores passent de 400 points lorsqu'il s'agit d'une école à composition favorisée à 300 points pour les écoles défavorisées. Ce résultat n'est cependant pas vérifié pour la catégorie d'enseignants ayant accumulé plus de 30 années d'expérience. Globalement, nos résultats démontrent que la composition mixte est d'autant plus importante pour les enseignants jeunes (Graphique 33.e).

Au regard du genre de l'enseignant, les femmes sont plus performantes dans les écoles favorisées avec un score de 390 points. Ce score diminue à 370 points dans une composition sociale mixte et s'avère être d'un niveau médiocre dans les compositions sociales défavorisées

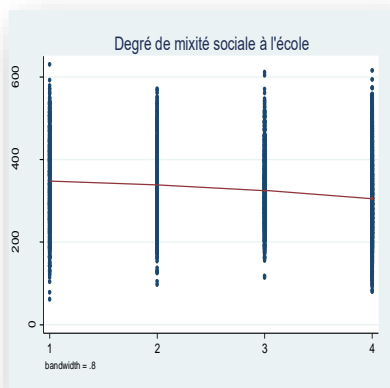
avec un score de seulement 300 points. Pour les hommes, ces derniers sont plus ou moins performants dans les classes à composition sociale mixte (Graphique 33.f).

En ce qui concerne la relation entre l'instruction des parents et la composition sociale et économique des écoles, on a pu démontrer dans les chapitres précédents que l'instruction des parents joue un rôle déterminant pour l'amélioration de la performance scolaire des élèves. En revanche, les résultats de notre modèle semi-paramétrique conditionnel au contexte montre que même si les élèves ont la chance d'avoir des parents bien instruits, ils réussissent moins bien lorsqu'ils fréquentent des classes à composition sociale défavorisées. En fait, les enfants dont les parents ont au moins un niveau d'éducation secondaire réussissent mieux au regard des autres élèves dont les parents ont un niveau d'éducation inférieur. Ces élèves réussissent d'autant mieux qu'ils fréquentent une école favorisée avec un certain degré de mixité sociale. Pour les autres élèves, on peut noter que la mixité sociale peut s'avérer favorable aux élèves dont les parents ont un niveau d'éducation correspondant au primaire bien que le score de réussite soit plus faible que pour les élèves dont les parents ont un niveau supérieur d'éducation. En considérant uniquement les parents ayant un niveau d'instruction élevé, les estimations montrent que le score baisse de 400 à 340 points, si on passe d'une composition sociale et économique favorisée à une composition défavorisée. La composition sociale mixte avec un degré de 25 à 50% des élèves défavorisés montre un score moyen de 375 points (Graphique 33.g).

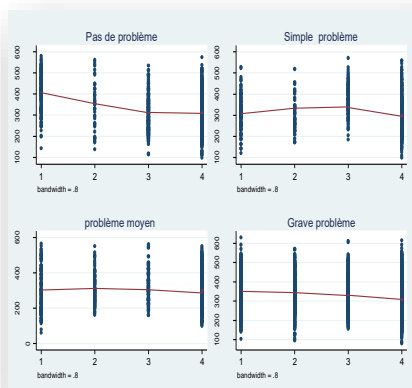
Les élèves préscolarisés ne performant pas tous de la même façon. Les résultats de nos simulations révèlent que les élèves préscolarisés assignés à des classes à composition sociale défavorisées sont les moins performants, enregistrant le score le plus bas (290 points) comparativement aux élèves assignés à des classes à composition sociale mixte (360 points) ou favorisée (380 points). Ces résultats apportent un éclairage complémentaire à nos résultats du modèle multiniveaux paramétrique estimé dans le chapitre 1, et aux résultats des études empiriques existantes au Maroc (Hanchane et al., 2012 ; Ibrouk, 2013 ; Mourji, 2013 et al.) (Graphique 33.h).

Graphique 33. Les courbes simulées du modèle multiniveaux semi-paramétrique conditionnel aux contextes en lecture

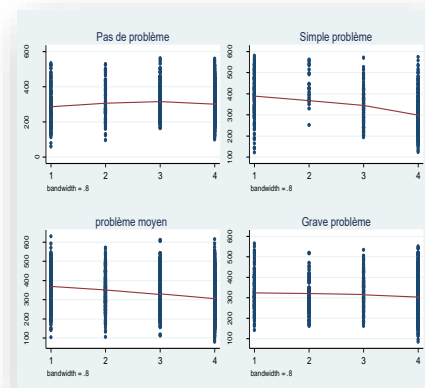
a. Degré de mixité sociale



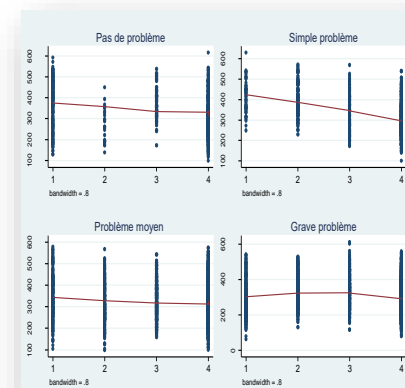
b. Absence des enseignants



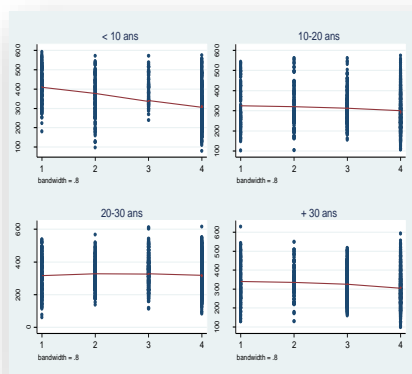
c. Absence des élèves



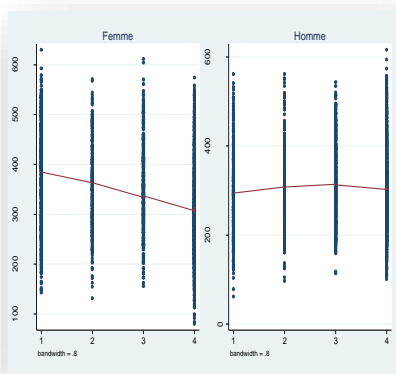
d. Charge horaire



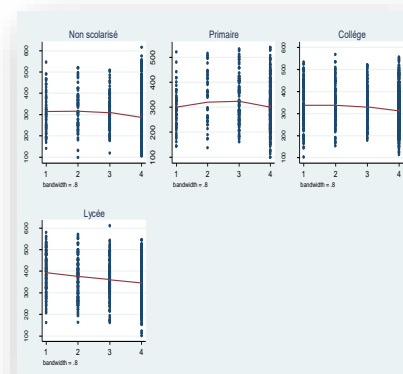
e. Expérience des enseignants



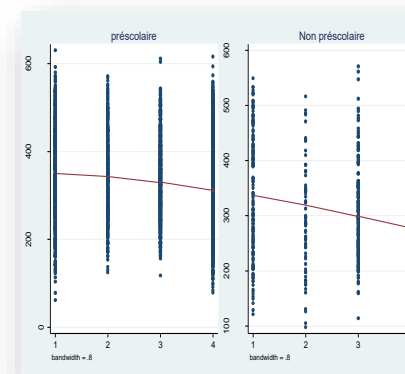
f. Genre de l'enseignant



g. Education des parents



h. Préscolarisation



Source : D'après l'auteur à partir des données PIRLS 2011

4.2. Résultats pour les mathématiques

Les élèves issus des classes à composition favorisée performant beaucoup mieux que les élèves issus des classes défavorisées, tandis que les élèves assignés à une composition mixte sont d'autant plus performants que ceux des écoles défavorisées. Les courbes simulées à partir du modèle multiniveaux semi paramétrique conditionnel au contexte nuance en partie ce résultat qui reste généraliste (Graphique 34).

En ce qui concerne l'âge de l'enseignant, les courbes simulées révèlent que les élèves assignés à des classes défavorisées performant beaucoup moins que les élèves issus des classes mixtes ou favorisées, et ceci pour l'ensemble des catégories d'âge des enseignants, sauf pour la catégorie d'âge 25-29 ans où on observe une courbe à escalier ascendant. Autrement dit, un élève appartenant à une école à composition mixte ou favorisée, et affecté à un enseignant âgé, est beaucoup plus performant qu'un élève affecté à un enseignant jeune, et à une école à composition sociale et économique défavorisée. Ce résultat affine la nature de l'effet causal de l'âge de l'enseignant sur la performance scolaire dans la mesure où cet effet est expliqué par une composante importante celle de la composition sociale et économique de l'établissement (Graphique 34.a).

L'analyse de la relation existante entre le nombre d'heures d'éducation par jour et composition sociale et économique à l'école démontre que les scores les plus élevés sont toujours enregistrés par les élèves appartenant à des classes favorisées quel que soit le degré de la charge horaire déclarée par les enseignants. En effet, les scores chutent de 450 (école favorisée) à 380 points (école défavorisée) pour la catégorie des enseignants avec une charge horaire de 5 à 6 heures par jour, tandis que la courbe simulée associée à la catégorie d'enseignant avec moins de 5h d'éducation par jour passe de 390 (école favorisée) à 350 points (école défavorisée) (Graphique 34.b). Cette analyse confirme une fois de plus que la composition mixte est bénéfique pour l'amélioration de la performance scolaire des élèves.

Au regard des problèmes d'absentéisme des enseignants, les directeurs déclarant n'avoir aucun problème d'absentéisme des enseignants, le score des élèves est de 450 points, lorsqu'il s'agit d'une école à composition favorisée, et de 370 points seulement pour les élèves affectés à des écoles à composition défavorisée. En revanche, la courbe simulée présente un score de 390 points chez les élèves issus des écoles à composition sociale et économique mixte (Graphique 34.c). En effet, la performance scolaire se stabilise à un score de 380 points pour les différentes

compositions sociales et économiques de l'établissement lorsqu'il s'agit des directeurs ayant déclaré un simple problème d'absentéisme des enseignants (Graphique 34.c).

Enfin, concernant le milieu de l'établissement, les résultats sont similaires à ceux des résultats obtenus pour les mathématiques. Les élèves issus du milieu urbain ne sont pas tous traités de la même manière. Les performances scolaires diffèrent considérablement selon la composition sociale de l'école. En résumé, quel que soit le milieu de l'établissement, la composition sociale et économique mixte est toujours plus avantageuse qu'une école défavorisée (Graphique 34.d).

Ainsi, les élèves assignés à des enseignantes (femme) performant mieux lorsqu'il s'agit d'une école à composition sociale favorisée (score de 430 points). En revanche, les élèves affectés à des enseignantes d'une école défavorisée n'enregistrent qu'un score de 360 points. La courbe simulée se stabilise par un score de 390 pour les écoles mixtes. Ces résultats valent pour les enseignants du genre masculin (Graphique 34.e).

S'agissant de l'éducation des parents, les résultats montrent que les élèves dont les parents ont un faible niveau d'instruction réalisent des scores avancés lorsqu'ils fréquentent des écoles à composition sociale et économique favorisée (400 points). La courbe simulée chute à 360 points pour les élèves affectés à des classes à compositions sociale et économique défavorisée. Cependant, les élèves issus d'une classe à composition sociale et économique mixte voient leur score avoisiner les 390 points. D'un autre côté, nous remarquons que les élèves ayant des parents bien instruits, et appartenant à des classes à composition sociale favorisée performant beaucoup mieux (490 points) que les élèves issus des écoles à composition sociale mixte (445 points), ou défavorisée (390 points). Ces résultats suggèrent l'existence d'une corrélation assez étroite entre le niveau d'éducation des parents et la composition sociale et économique de l'établissement ou l'élève est affecté, tout en révélant qu'une école mixte reste beaucoup plus bénéfique qu'une école à composition défavorisée pour la réussite scolaire des élèves. (Graphique 34.f).

En considérant le temps consacré aux devoirs, les élèves des écoles favorisées qui consacrent plus de 90 minutes pour les devoirs des mathématiques à la maison sont les plus performants avec un score de 420 points. La courbe simulée passe de 380 à 400 points lorsqu'il s'agit d'une école mixte et seulement à 360 points pour les élèves des écoles défavorisées. Globalement, la trajectoire de la courbe est vérifiée pour l'ensemble des contextes en relation avec le temps consacré aux devoirs des mathématiques à la maison (Graphique 34.g).

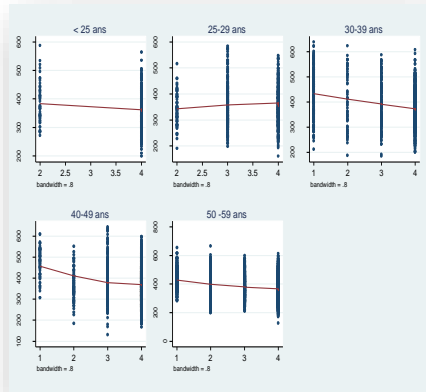
L'intérêt pour les mathématiques est un facteur qui révèle des écarts significatifs de performance scolaire entre les élèves. Ce résultat reste conditionné par la composition sociale et économique de l'école. La courbe simulée indique que les élèves affectés à une école à composition favorisée et qui apprécient les mathématiques performant beaucoup mieux que des élèves d'une école à composition sociale et économique défavorisée. Par ailleurs, la composition sociale mixte reste plus bénéfique qu'une école à composition sociale défavorisée. Autrement dit, avoir la passion envers les mathématiques et être affecté à une école à composition mixte est plus bénéfique pour l'élève que être assigné à une école à composition défavorisée (Graphique 34.h).

La satisfaction très élevée des enseignants augmente la performance scolaire des élèves, conditionnellement à leur appartenance à une école à composition sociale et économique favorisée. Pour le même groupe d'enseignant déclarant une satisfaction très élevée envers le métier d'enseignement, la courbe simulée connaît une trajectoire à la baisse lorsqu'il s'agit d'une école à composition sociale défavorisée. Pour ce même groupe d'enseignant, la performance scolaire est beaucoup plus implorante chez les élèves assignés à une école à composition sociale mixte comparativement à une école défavorisée.

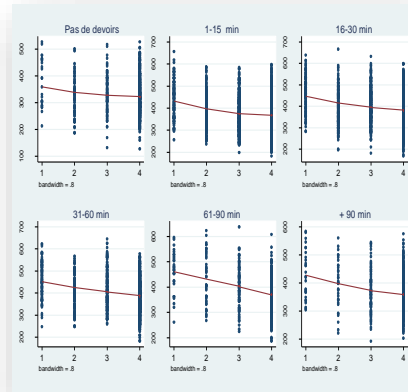
Concernant la formation continue des enseignants, les courbes simulées présentent une trajectoire à la baisse en fonction du degré de la composition sociale et économique de l'établissement, et ceci pour les enseignants ayant bénéficié ou non de cette formation continue.

Graphique 34. Les courbes simulées du modèle multiniveaux semi-paramétrique en mathématique

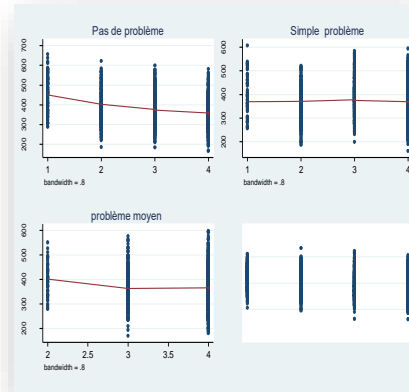
a. Age de l'enseignant



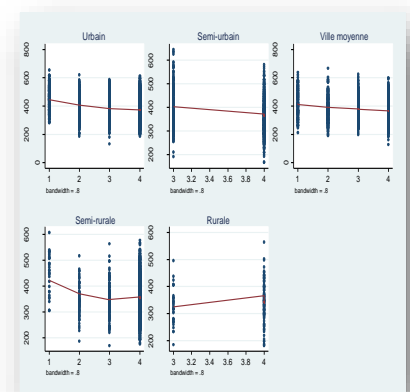
b. Nombre d'heures de devoirs



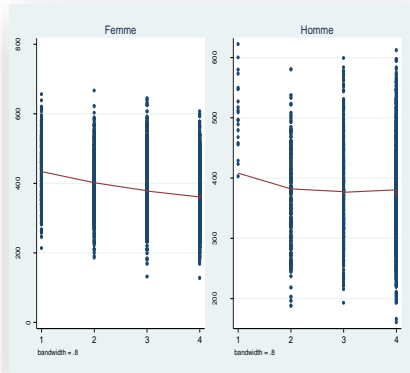
c. Absence des enseignants



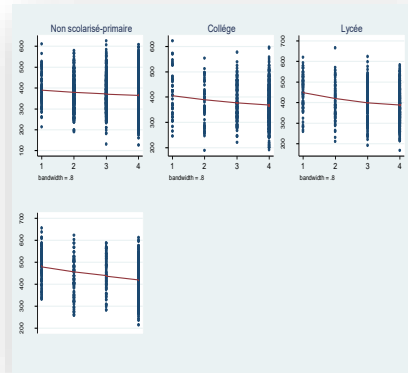
d. Milieu de l'établissement



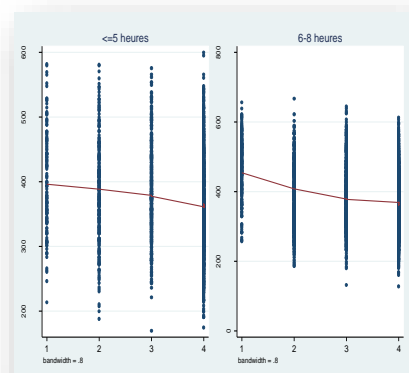
e. Genre de l'enseignant



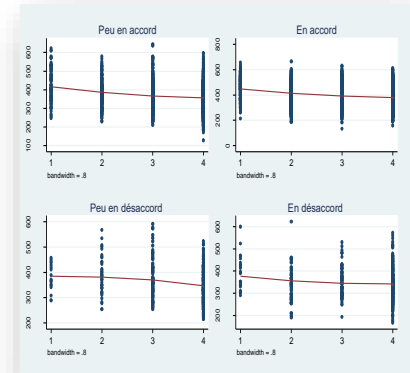
f. Education des parents



g. Nombre d'heure d'éducation par jour



h. Aime les mathématiques



Source : D'après l'auteur à partir des données TIMSS 2011

4.3. Résultats pour les sciences

Suivant la même démarche d'interprétation des résultats, cette sous-section présente les résultats pour les sciences en se référant à la même logique d'interprétation présentée pour la lecture et les mathématiques.

Concernant l'âge de l'enseignant, la trajectoire des courbes simulées rejoint les résultats démontrés pour la lecture et les mathématiques. En fait, les élèves assignés à des enseignants jeunes ou âgés performant beaucoup mieux quand ils appartiennent à des écoles à une composition sociale et économique favorisée. En effet, la courbe simulée indique une chute considérable pour la catégorie d'élèves issus des écoles défavorisées. Une lecture synthétique du graphique montre que le score passe de 430 (écoles favorisés) à 370 points (école défavorisés), tout en précisant encore une fois qu'une école à composition sociale mixte est plus avantageuse que qu'une école à composition défavorisée, car elle se situe entre les scores enregistrés dans les deux catégories d'école favorisée et défavorisée (Graphique 35.a).

Par rapport à l'éducation des parents, les scores les plus élevés (450 points) sont atteints pour les élèves d'écoles favorisées ayant des parents éduqués. La trajectoire de la courbe simulée révèle une dégradation des scores en passant d'une école favorisée à une école défavorisée, et ceci pour l'ensemble des niveaux d'éducation des parents. Pour autant avoir des parents avec un niveau d'éducation élevé n'engendre pas automatiquement une performance scolaire élevée car manifestement l'école et son environnement à travers sa composition sociale et économique interviennent d'une manière indirecte dans la détermination du niveau des acquis scolaires des élèves (Graphique 35.b).

Concernant le milieu de l'établissement, les simulations montrent que les courbes estimées présentent d'une manière globale une allure à la baisse. En effet, les élèves issus du milieu urbain appartenant aux écoles à composition sociale favorisée réussissent beaucoup mieux que les élèves assignées à des écoles mixtes ou défavorisées. Cependant, les résultats indiquent bien que la performance scolaire dans une école mixte en milieu urbain est plus élevée que dans une zone présentant une ségrégation scolaire élevée. En milieu rural, cette analyse reste limitée en raison de la concentration des élèves issus généralement des classes défavorisées avec un niveau très faible d'éducation des parents (Graphique 35.c).

Les élèves affectés à des enseignants déclarant avoir un grave problème de charge horaire n'enregistrent pas tous les mêmes scores, car les élèves d'une école à composition sociale favorisée sont les seuls qui obtiennent des scores élevés comparativement à ceux d'une école défavorisée (370 points). De plus, il apparaît que les élèves issus des écoles à composition sociale mixte marquent des scores assez comparables avec les élèves issus des écoles favorisées. Ainsi, quelle que soit la perception des enseignants envers les problèmes liés à la charge horaire (grave problème ou absence de problème), les élèves des écoles à composition sociale favorisée et mixte performant beaucoup mieux que les élèves issus des écoles défavorisées (Graphique 35.d).

La littérature empirique montre d'une façon unanime les différences de performances scolaires enregistrées entre les garçons et les filles. Nos résultats corroborent ce résultat par le fait que les filles affectées à des écoles à composition favorisée et mixte sont beaucoup plus performantes que celles fréquentant une école défavorisée. La courbe simulée démontre ce résultat aussi bien pour le genre féminin et masculin (Graphique 35.e).

L'intérêt de l'élève pour les matières scientifiques est un facteur déterminant du niveau des acquis scolaires. En effet, nos simulations montrent qu'à partir du moment où l'élève est passionné pour la matière, le degré de mixité sociale de l'école ne semble pas influencer ses performances scolaires (Graphique 35.f).

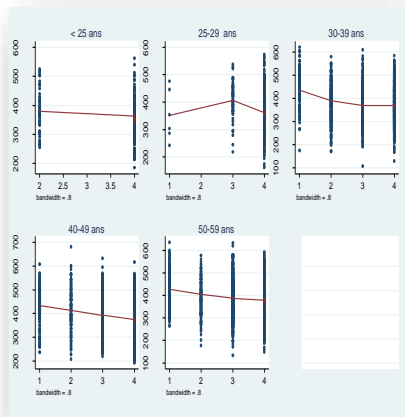
En considérant les enseignants ayant bénéficié d'une formation continue, il apparaît que cette formation continue a un impact plus important sur les performances des élèves (400 points) dans les écoles favorisées. La courbe montre une allure décroissante pour les enseignants affectés dans des écoles mixtes, et des performances scolaires faibles pour la catégorie des enseignants affectés à des écoles défavorisées (Graphique 35.g). Ces résultats rejoignent nos résultats sur la performance scolaire en mathématique, et mettre en question les résultats des modèles économétriques linéaires (Hanchane et al, 2012).

Concernant le genre, la trajectoire de notre courbe laisse apparaître que la performance scolaire diminue en fonction de la composition sociale et économique de l'établissement pour les enseignants du genre féminin et masculin. Pour les premières, la trajectoire de performance des élèves passe de 450 points en faveur des écoles à composition favorisée à 420 points pour les élèves affectés à une école à composition sociale et économique mixte, et à 380 points pour les enseignantes des écoles défavorisées. Ces mêmes tendances apparaissent pour les enseignants masculins (Graphique 35.h).

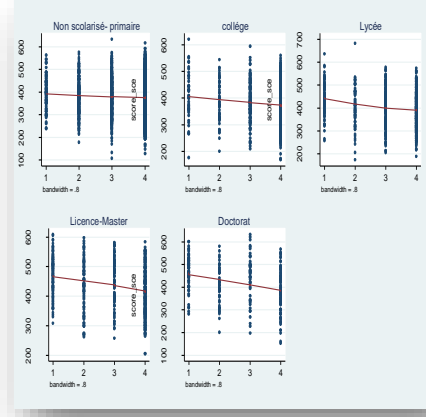
En conclusion, nos résultats pour les trois matières évaluées (lecture, mathématiques et sciences) révèlent qu'un degré de mixité sociale à l'école est bénéfique pour l'amélioration du niveau des acquis scolaires des élèves. Ce résultat est démontré dans le cadre de notre un modèle multiniveaux semi-paramétrique conditionnel aux contextes dans lesquels les élèves étudient. Cette modélisation nous a permis par le biais de simulations que cette conclusion reste valable dans différents contextes tenant compte des dysfonctionnements et des problèmes que les enseignants peuvent rencontrés quotidiennement au sein de l'école marocaine

Graphique 35. Les courbes du modèle multiniveaux semi-paramétrique en science

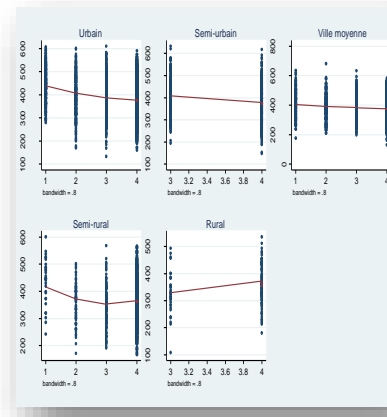
a. Age de l'enseignant



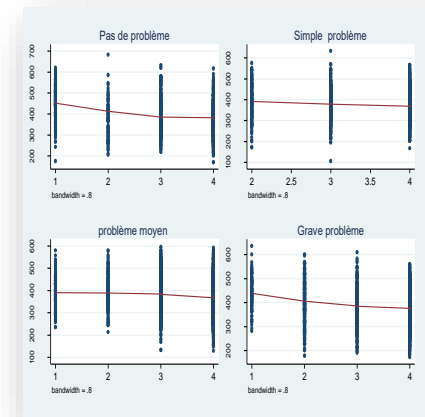
b. Education des parents



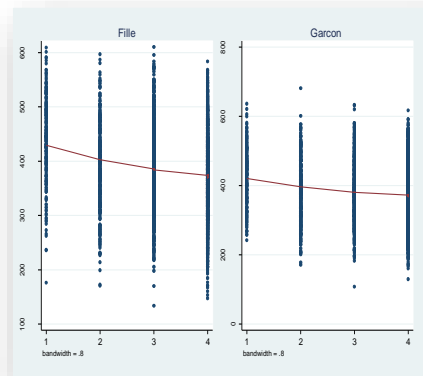
c. Milieu de l'établissement



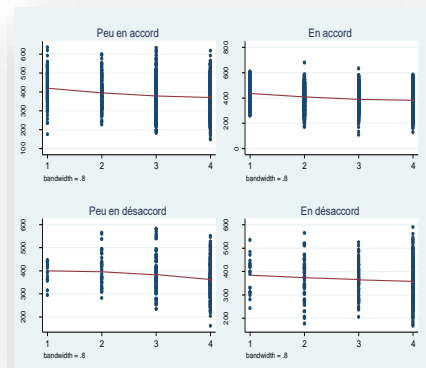
d. Charge horaire



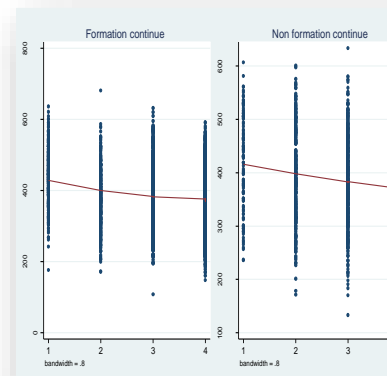
e. Genre de l'élève



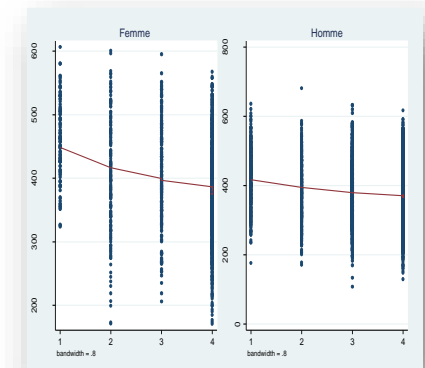
f. Aime les sciences



g. Formation continue des enseignants



h. Genre de l'enseignant



Source : D'après l'auteur à partir des données TIMSS 2011

4.4. Discussion et recommandations

Les inégalités d'acquis scolaire demeurent une question centrale des défaillances qu'a connu le système éducatif marocain depuis les années 2000. En effet, ces inégalités s'expliquent par plusieurs facteurs qui interviennent pour la détermination du niveau des acquis scolaires d'un élève, d'un établissement scolaire, d'une région ou d'un quartier. La faiblesse de la mixité sociale dans les écoles marocaines est inévitablement en lien étroit avec la concentration des difficultés sociales dans les quartiers défavorisés et à la ségrégation urbaine.

Les élèves appartenant aux milieux les plus défavorisés et qui n'ont qu'accès à une école publique gratuite, ont manifestement moins de chance de réussir leur parcours scolaire, et de s'insérer facilement à un marché de travail qui devient de plus en plus compétitif. Ce constat met en question la cohésion sociale qui ne se réalisera jamais sans une école qui favorise l'égalité des chances.

Dans ce contexte, le Maroc ne dispose d'aucun outil statistique permettant de contrôler l'ampleur et les évolutions de la ségrégation sociale et scolaire des écoles. Peu de travaux de l'effet de la mixité sociale sur la réussite scolaire des élèves ont été réalisés à notre connaissance.

Ce chapitre a tenté d'apporter un éclairage sur les relations entre la composition sociale et économique de l'établissement et le niveau de la performance scolaire des élèves. En fait, si la littérature empirique démontre la relation dans un cadre linéaire, notre apport réside dans l'examen de la non linéarité de cette relation.

A travers l'estimation des modèles semi-paramétriques conditionnels au contexte dans lequel les élèves étudient, nous avons démontré que la composition sociale et économique de l'établissement mixte leur est bien plus bénéfique comparativement à une école à composition défavorisée quel que soit le contexte ou l'élève est assigné. En revanche, nos résultats indiquent que les meilleures performances sont atteintes par les élèves des écoles favorisées.

Nos résultats suggèrent que l'instauration de mécanismes innovants visant un assouplissement de la carte scolaire pourrait modérer les problèmes de ségrégation sociale et par conséquent une amélioration du niveau des acquis scolaires des élèves. Cet instrument largement utilisé au niveau international a été à l'écart de nos politiques éducatives marocaines.

Pour en connaître mieux les canaux et mécanismes de cet effet, des expérimentations régionales, provinciales ou locales doivent être instaurés en urgence. De plus, le système d'évaluation reste modeste et moins innovants par rapport aux politiques éducatives qui ont été aménagés depuis des décennies. Les expériences internationales en matière de politique éducative favorisant la mixité sociale à l'école mettent en évidence un retard marocain en la matière.

De plus, si l'institut des politiques publiques en France évalue régulièrement l'effet d'un degré de mixité sociale à l'école sur la réussite scolaire, les institutions publiques marocaines chargées du dossier ne produisent que peu d'évaluation de même nature.

En somme, ce travail conclue que la performance scolaire des élèves est conditionnée naturellement et intuitivement par un niveau de mixité sociale à l'école. Cette dernière renvoie à l'assouplissement de la carte scolaire en passant par des expérimentations dans le but de cibler les contextes où cette politique éducative réussisse d'une manière efficace et efficiente. De même, et compte tenu de l'importance du sujet de la mixité dans le débat public, des politiques innovantes doivent être instaurés en urgence, tout en les accompagnant par des évaluations rigoureuses et une expertise solide en la matière. Le développement de ces recherches et travaux engage aussi l'implication des institutions chargées de la production statistique pour la mise en place des systèmes d'informations dynamiques et spatiales en lien avec la mixité sociale en générale et la mixité scolaire en particulier.

Conclusion

L'originalité de ce travail réside dans le fait d'évaluer la performance scolaire des élèves marocains selon la composition sociale et économique de l'école. En partant de la catégorisation définie par les programmes TIMSS et PIRLS, nous avons identifié les écarts de performance scolaire des élèves selon les différents niveaux de la composition sociale et économique de l'école (écoles favorisée, mixte et défavorisée). Pour ce faire, nous avons présenté plusieurs modèles économétriques qui permettent d'identifier cet effet en mettant comme objectif de démontrer qu'un degré de plus de mixité sociale et économique au niveau des établissements marocains entraînera une amélioration du niveau des acquis scolaires des élèves. Pour ce faire, nous avons choisi de focaliser notre attention sur le modèle multiniveaux semi-paramétrique conditionnelle aux contextes dont les élèves sont assignés. A travers ce dernier, nous avons simulé les courbes mettant en relation le score des élèves pour chaque matière évaluée et le niveau de la composition sociale et économique de l'école, en mettant en relief les différents contextes d'accumulation des acquis scolaire. Cette modélisation a démontré des résultats très intéressants dans un cadre assez général et innovants, tout en mettant en question les travaux antérieurs (Hanchane et al.,2012 ; Ibourk,2013 ; Mourji et al ,2013). En fait, si la modélisation paramétrique permet d'identifier des coefficients associés à chaque variable du modèle, la modélisation semi-paramétrique simule des courbes reliant le score des élèves et la composition sociale et économique de l'école pour un ensemble de contexte (Milieu de l'établissement, problème de charge horaire, éducation des parents, etc).

Selon les résultats de notre modélisation, nous avons réussi à démontrer que la composition sociale et économique mixte au sein de l'école est plus bénéfique pour améliorer l'apprentissage des élèves en premier, et aussi pour s'assurer d'une réussite scolaire généralisée de tous les élèves. En effet, l'exercice empirique a démontré d'une manière claire que quel que soient les contextes ou l'élève est assigné, la composition sociale et économique de l'école qui garantit une réussite scolaire généralisée est celle d'une école à composition sociale et économique mixte. Ce type d'école reste beaucoup plus bénéfique pour l'ensemble des élèves en matière d'acquisition scolaire qu'une école caractérisée par une ségrégation sociale élevée.

Le chapitre 1 de la présente thèse a permis de révéler les déterminants de la performance scolaire des élèves marocains, tout en montrant une relation significative entre le niveau de composition sociale et économique de l'école et la performance scolaire des élèves. Le présent chapitre a mis le point sur la relation existante entre la composition sociale et économique de

l'école et performance scolaire, tout en concluant à travers une modélisation semi-paramétrique que la mixité sociale à l'école est une politique publique pertinente pour favoriser la réussite de tous les élèves.

Ce chapitre met en évidence les effets très négatifs de la ségrégation sociale sur les acquis scolaires des élèves et par conséquent l'absence d'une mixité sociale à l'école a également un impact négatif sur le mode d'organisation au sein des établissements en général, et sur les conditions d'apprentissage des élèves.

Une politique éducative favorisant la mixité sociale à l'école impactera naturellement les comportements aux seins des établissements, les pratiques pédagogiques, les relations maître-élève et les relations maître-directeur, etc, dans le but de garantir des performances scolaires comparables à l'échelle internationale (TIMSS, PIRLS, PISA 2018, etc). En effet, la mixité sociale à l'école marocaine s'impose aujourd'hui comme une évidence et comme un levier de démocratisation pour plus de justice scolaire et par conséquent une justice sociale et une démocratisation basé sur l'égalité des résultats et non pas l'égalité de chance.

Enfin, il faut mentionner qu'une dose de mixité sociale à l'école est un instrument pertinent pour favoriser la réussite et, plus largement, l'intégration future des élèves dans la société. En effet, les conclusions du chapitre mettent en évidence la nécessité d'un débat public profond sur l'assouplissement de la carte scolaire. Cet instrument de politique éducative très débattu et instauré par les systèmes éducatifs les plus performants nécessite d'être expérimenté d'une manière intuitive et rigoureuse à travers des programmes d'expérimentations régionaux, accompagnés par un suivi et des évaluations indépendantes, scientifiques et pertinentes afin d'en tirer les enseignements et les leçons nécessaires pour une généralisation sur l'ensemble du système éducatif marocain. Comme cité auparavant, il est incontournable que la réussite scolaire des élèves est en grande partie liée aux phénomènes de ségrégation sociale et à l'absence des politiques innovantes favorisant la mixité sociale et économique à l'école. L'intervention du système d'évaluation de manière régulière, et à travers des protocoles adaptés permettra de redresser les dysfonctionnements et arriver à remédier la non qualité de l'éducation au Maroc telle qu'elle est mesurée par le niveau des acquis scolaire à travers les programmes internationaux TIMSS, PIRLS, PISA ou nationaux comme PNEA.

Annexes du chapitre 3

Tableau 44. Les résultats des modèles économétriques en lecture

Variables explicatives	Multilevel modeling		Semiparametric modeling		Semiparametric Multilevel modeling	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
Féminin	29.61921	0.000	30.97202	0.000	19.71573	0.000
Masculin (Ref)	-	-	-	-	-	-
Age <=2000	-1.594974	0.720	-13.97908	0.005	.3878123	0.941
Age 2001-2003	40.44646	0.000	38.54216	0.000	30.50914	0.000
Age >=2004 (Ref)	-	-	-	-	-	-
Sentiment d'appartenance à l'école (en accord)	40.86914	0.000	50.46239	0.000	33.08702	0.000
Sentiment d'appartenance à l'école (peu d'accord)	23.70564	0.000	23.74746	0.000	23.85044	0.000
Sentiment d'appartenance à l'école (en désaccord) (Ref)	-	-	-	-	-	-
Plaisir de lecture (tout à fait d'accord)	51.1661	0.000	62.67342	0.000	40.25373	0.000
Plaisir de lecture (peu en accord)	26.34437	0.000	32.26314	0.000	22.3693	0.000
Plaisir de lecture (peu en désaccord)	13.18622	0.004	7.985906	0.165	14.33545	0.004
Plaisir de lecture (en désaccord) (Ref)	-	-	-	-	-	-
Expérience d'enseignant (<=10 ans)	29.81419	0.006	22.62683	0.000	-	-
Expérience d'enseignant (10-20 ans)	-17.58101	0.071	-14.89066	0.000	-	-
Expérience d'enseignant (20-30 ans)	-4.722124	0.639	.5667024	0.839	-	-
Expérience d'enseignant (>=31 ans) (Ref)	-	-	-	-	-	-
Enseignant (femme)	35.97568	0.000	20.6969	0.000	-	-
Enseignant (homme) (Ref)	-	-	-	-	-	-
Satisfaction des enseignants (trop élevé)	47.89923	0.000	4.044587	0.285	-	-
Satisfaction des enseignants (élevé)	13.90719	0.104	3.716361	0.136	-	-
Satisfaction des enseignants (bas) (Ref)	-	-	-	-	-	-
Satisfaction des enseignants (tout à fait d'accord)	24.38166	0.019	12.81175	0.000	-	-
Satisfaction des enseignants (peu d'accord)	.7163456	0.950	7.986772	0.024	-	-
Satisfaction des enseignants (en désaccord) (Ref)	-	-	-	-	-	-
Aucune formation continue	-25.56895	0.001	-22.63581	0.000	-	-

Ayant accès à une formation continue (Ref)	-	-	-	-	-	-
entre 0-10 % d'élèves défavorisées	42.89571	0.001	-	-	-	-
entre 11-25% d'élèves défavorisées	20.15998	0.171	-	-	-	-
entre 26-50% d'élèves défavorisées	26.38593	0.038	-	-	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-
Ecole urbaine	46.20681	0.000	30.79336	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-
Préscolaire >=3 ans	26.0977	0.000	41.18314	0.000	23.20151	0.000
Préscolaire 2-3 ans	25.00444	0.000	38.65933	0.000	19.83715	0.000
Préscolaire 2 ans	12.97391	0.000	26.90959	0.000	8.705147	0.001
Préscolaire 1-2 ans	1.643816	0.664	7.701831	0.102	-4.56617	0.234
Préscolaire < =1 ans (Ref)	-	-	-	-	-	-
Education des parents (Doctorat)	31.87252	0.000	46.17723	0.000	16.50113	0.010
Education des parents (Master)	40.7963	0.000	70.85365	0.000	21.48841	0.000
Education des parents (licence)	33.46745	0.000	55.39971	0.000	28.44719	0.000
Education des parents (lycée)	15.75041	0.000	21.7334	0.000	11.24392	0.000
Education des parents (secondaire collégial)	10.01326	0.001	7.791735	0.038	6.036153	0.051
Education des parents (primaire ou analphabète) (Ref)	-	-	-	-	-	-
Taille de la classe >=40 élèves	-21.88215	0.087	-21.10049	0.000	-	-
Taille de la classe 30-40 élèves	-19.36009	0.096	-11.35547	0.010	-	-
Taille de la classe 20-30 élèves	11.92766	0.332	4.846684	0.302	-	-
Taille de la classe < 20 élèves	-	-	-	-	-	-
Constante	306.4271	0.000	319.7382	0.000		
Random-effects Parameters	Estimate	Interval				
idschool: Identity var (constante)	1734.539	1960.015				
idschool: Identity var(Residual)	4455.893	4732.833				
Test d'Hausman	87.22		121.11		0.00	

Source : D'après l'auteur à partir des données PIRLS 2011

Tableau 45. Les résultats des modèles économétriques en mathématique

Variables explicatives	Multilevel modeling		Semiparametric modeling		Semiparametric Multilevel modeling	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
masculin	6.764282	0.000	3.072083	0.066	24.4886	0.000
Féminin (Ref)	-	-	-	-	-	-
Age >=1995	-40.6529	0.000	-47.21667	0.000	-32.65906	0.000
Age <1995 (Ref)	-	-	-	-	-	-
Education des parents –Doctorat	14.04107	0.000	43.00657	0.000	-2.00343	0.650
Education des parents –Master	38.57882	0.000	66.32563	0.000	25.6806	0.000
Education des parents –Licence	21.34604	0.000	32.18051	0.000	10.91038	0.000
Education des parents –lycée	6.681557	0.006	8.573817	0.002	2.618187	0.286
Education des parents - secondaire, primaire ou analphabète(Ref)	-	-	-	-	-	-
l'élève travaille dur les maths (accord)	36.19399	0.000	40.38936	0.000	30.65662	0.000
l'élève travaille dur les maths (peu d'accord)	8.598405	0.000	12.12013	0.000	8.878913	0.000
l'élève travaille dur les maths (en désaccord) (Ref)	-	-	-	-	-	-
Pas de devoirs mathématiques	-16.77014	0.000	-19.38653	0.000	-14.92612	0.000
Devoirs mathématiques 1 -15min	23.39799	0.000	24.99336	0.000	21.42601	0.000
Devoirs mathématiques 16-30min	32.81213	0.000	39.40849	0.000	29.16136	0.000
Devoirs mathématiques 31-60min	37.318	0.000	46.9285	0.000	32.92823	0.000
Devoirs mathématiques 61-90min	27.73571	0.000	40.27959	0.000	26.06312	0.000
Devoirs mathématiques plus de 90 minutes (Ref)	-	-	-	-	-	-
entre 0-10 % d'élèves défavorisées	79.16372	0.000	-	-	-	-
entre 11-25% d'élèves défavorisées	10.5292	0.348	-	-	-	-
entre 26-50% d'élèves défavorisées	9.86767	0.258	-	-	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-
Ecole urbaine	23.70393	0.000	10.67269	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-
Encourager les enseignants par indemnités	80.10149	0.000	35.16923	0.000	-	-
Pas d'encouragement des enseignants (Ref)	-	-	-	-	-	-

Enseignants femme	-4.557967	0.544	5.005487	0.008	-	-
Enseignants homme (Ref)	-	-	-	-	-	-
Satisfaction enseignants (trop bas)	-43.98948	0.070	-16.50395	0.006	-	-
Satisfaction enseignants (bas)	-28.27502	0.019	-6.865314	0.022	-	-
Satisfaction enseignants (moyen)	-26.61863	0.000	-9.442881	0.000	-	-
Satisfaction enseignants (élevé) (Ref)	-	-	-	-	-	-
7-8 heures de maths par semaine	52.44432	0.108	-41.22851	0.000	-	-
4-6 heures de maths par semaine	16.53618	0.074	-8.778038	0.707	-	-
<=3 heures de maths par semaine (Ref)	-	-	-	-	-	-
Presence des Formations continues	14.4077	0.029	6.463831	0.000	-	-
Absence des Formations continues (Ref)	-	-	-	-	-	-
Taille de la classe >=38 élèves	-60.97203	0.000	-16.26407	0.000	-	-
Taille de la classe 28-37 élèves	-65.14421	0.000	-15.44218	0.000	-	-
Taille de la classe <=27 élèves (Ref)	-	-	-	-	-	-
Random-effects Parameters						
idschool: Identity						
idschool: Identity Var (_cons)	2930.155					
idschool: Identity var(Résiduel)	4199.151					
LR test vs. linear regression:	chibar2(01)					
Test d'Hausman	177.32		132.98		0.00	

Source : D'après l'auteur à partir des données TIMS 2011

Tableau 46. Les résultats des modèles économétriques en science

Variables explicatives	Multilevel modeling		Semiparametric modeling		Semiparametric Multilevel modeling	
	$\hat{\beta}$	p value	$\hat{\beta}$	p value	$\hat{\beta}$	p value
masculin	-1.673559	0.258	2.817085	0.094	-15.84815	0.000
Féminin (Ref)	-	-	-	-	-	-
Age >=1995	-43.40825	0.000	-48.51976	0.000	-34.1613	0.000
Age <1995 (Ref)	-	-	-	-	-	-
Disponibilité d'une chambre	-3.717729	0.014	1.109587	0.513	-6.998161	0.000
Chambre non disponible(Ref)	-	-	-	-	-	-
Education des parents –Doctorat	16.80856	0.000	37.46644	0.000	-1.119155	0.810
Education des parents –Master	37.92809	0.000	58.23201	0.000	24.75928	0.000
Education des parents –Licence	17.63163	0.000	26.46567	0.000	8.988026	0.001
Education des parents –lycée	4.982376	0.050	7.022555	0.014	2.268987	0.381
Education des parents - secondaire, primaire ou analphabète(Ref)	-	-	-	-	-	-
Sentiment plaisir pour les sciences (accord)	15.60936	0.000	23.32847	0.000	12.00924	0.000
Sentiment plaisir pour les sciences (peu d'accord)	4.929877	0.038	11.1628	0.000	-.1910986	0.939
Sentiment plaisir pour les sciences (en désaccord) (Ref)	-	-	-	-	-	-
Devoirs des sciences (pas de devoirs)	-13.58317	0.000	-12.87819	0.000	-9.811046	0.001
Devoirs des sciences (1-15 min par jour)	8.238654	0.000	14.72896	0.000	7.116326	0.000
Devoirs des sciences (16-30 min par jour)	15.10568	0.000	23.1868	0.000	12.96074	0.000
Devoirs des sciences (plus de30 min par jour) (Ref)	-	-	-	-	-	-
entre 0-10 % d'élèves défavorisées	61.75116	0.000	-	-	-	-
entre 11-25% d'élèves défavorisées	8.622634	0.375	-	-	-	-
entre 26-50% d'élèves défavorisées	9.450199	0.211	-	-	-	-
plus de 50 %d'élèves défavorisées (Ref)	-	-	-	-	-	-
Ecole urbaine	18.87458	0.001	8.816495	0.000	-	-
Ecole rural (Ref)	-	-	-	-	-	-
Encourager les enseignants par indemnités	69.79294	0.000	36.15248	0.000	-	-
Pas d'encouragement des enseignants (Ref)	-	-	-	-	-	-

Enseignant femme	17.93351	0.013	12.67579	0.000	-	-
Enseignant homme (Ref)	-	-	-	-	-	-
Age <=25 ans	-10.96613	0.608	-15.4165	0.018	-	-
Age 25 -29 ans	-3.189306	0.843	-9324979	0.848	-	-
Age 30 -39 ans	15.86276	0.237	.2682514	0.948	-	-
Age 40 -49 ans	21.523	0.104	9.010258	0.025	-	-
Age 50 -59 ans	19.85279	0.121	7.821319	0.046	-	-
Age >=60 ans (Ref)	-	-	-	-	-	-
Satisfaction des enseignants (trop élevé)	28.61983	0.006	12.99222	0.000	-	-
Satisfaction des enseignants (élevé)	24.38032	0.000	9.478079	0.000	-	-
Satisfaction des enseignants (bas)	-	-	-	-	-	-
Connaissance des programmes (trop élevé)	42.78106	0.000	17.42384	0.000	-	-
Connaissance des programmes (élevé)	7.565689	0.193	3.0198	0.088	-	-
Connaissance des programmes (bas) (Ref)	-	-	-	-	-	-
Presence des Formations continues	4.111977	0.483	-.2462935	0.888	-	-
Absence des Formations continues (Ref)	-	-	-	-	-	-
Taille de la classe >=38 élèves	-59.65715	0.000	-27.06474	0.000	-	-
Taille de la classe 28-37 élèves	-62.4552	0.000	-27.5305	0.000	-	-
Taille de la classe <=27 élèves (Ref)	-	-	-	-	-	-
Random-effects Parameters						
idschool: Identity var (constante)	1612.539	1938.015				
idschool: Identity var(Résiduel)	4572.893	4710.828				
Test d'Hausman	123.76			0.000		0.000

Source : D'après l'auteur à partir des données TIMS 2011

Tableau 47. Descriptions des variables du programme PIRLS 2011

Code PIRLS	Variables	Descriptions des variables
asrrea	Score des élèves en lecture	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 4eme année primaire en 2011 variant entre 300 à 600 points
acbg03a	Dans une classe, quel est Le pourcentage des élèves issus d'un milieu défavorisé ?	Une variable indicatrice à quatre modalités : 1 : entre 0-10 % ; 2 : entre 11-25% ; 3 : entre 26-50 % ; 4 : plus de 50 %
Variables intrinsèques à l'élève		
G1	Genre	Une variable dichotomique indiquant le genre de l'élève : 1 : garçon, 0 : fille.
G2	Age de l'élève en année de 2011	Une variable continue indiquant l'âge de l'élève
G3	Combien de fois l'élève parle la langue arabe à la maison	Une variable indicatrice à trois modalités : 1 : Toujours ; 2 : parfois ; 3 : jamais
G4	Le nombre approximatif de livres dans votre maison	Une variable indicatrice à cinq modalités : 1 : entre 0 -10 ; 2 : entre 11-25 ; 3 : 26-100 ; 4 : entre 101-200 ; 5 : entre plus que 200
G5-D	Disponibilité d'une chambre propre à l'élève	Une variable dichotomique indiquant la disponibilité d'une chambre propre à l'élève : 1 : Oui, 0 : non.
G7-D	Les parents s'assurent si l'élève a bien accompli ses devoirs à la maison	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
G8-A	L'élève a le sentiment d'appartenance à l'école	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalemment en désaccord
G9-A	L'élève a été soumis dans des situations d'insultes par d'autres élèves	Une variable indicatrice à quatre modalités : 1 : au moins une fois par semaine ; 2 : 1 à 2 fois par mois ; 3 : quelque fois pendant l'année ; 4 : jamais
R1	En une journée, Combien de temps l'élève consacre-t-il à la lecture en dehors de l'école	Une variable indicatrice à quatre modalités : 1 : moins de 30 minutes ; 2 : de 30 minutes à une heure ; 3 : entre une heure et 2 heures ; 4 : plus que 2 heures
R2-A	Combien de fois l'élève a lu pour le plaisir ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R4	Combien de fois l'élève emprunte-t-il des livres de la bibliothèque de l'école ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R5-A	L'élève aime-t-il ce qu'il a lu à l'école ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout a fait en désaccord
R5-B	L'enseignant donne-t-il à l'élève des choses importantes pour la lecture ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout a fait en désaccord
R6-A	Combien de fois l'élève a lu une lecture silencieuse ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : une à deux fois par semaine ; 3 : une à deux fois par mois ; 4 : jamais
R7-F	L'élève trouve-t-il du plaisir à lire ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
R8-A	L'élève lit-il très bien (niveau de performance en lecture) ?	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
Variables liées au contexte d'environnement familial (Parents)		
asbh02-A	Avant la rentrée à l'école primaire, Combien de fois vous avez lu des livres avec votre enfant ?	Une variable indicatrice à trois modalités : 1 : beaucoup ; 2 : parfois ; 3 : jamais

asbh02-H	Avant la rentrée à l'école primaire, Combien de fois vous avez écrit des lettres et des mots avec votre enfant ?	Une variable indicatrice à trois modalités : 1 : beaucoup ; 2 : parfois ; 3 : jamais
Asbh05	L'âge de rentrée scolaire à la première année du cycle primaire	Une variable indicatrice à quatre modalités : 1 : moins ou égal à 5 ans ; 2 : 6ans ; 3 : 7ans ; 4 : 8 ans et plus
Asbh04a	Est-ce que votre enfant a bénéficié de l'école maternelle ou du préscolaire ?	Une variable indicatrice à 2 modalités : 1 : Oui ; 2 : Non
Asbh04b	Combien d'année votre enfant a bénéficié de l'enseignement du préscolaire ?	Une variable indicatrice à 5 modalités : 1 : 3 ans et plus ; 2 : entre 2 -3 ans ; 3 : 2 ans ; 4 : entre 1 – 2 ans ; 5 : moins ou égal à une année
Asbh08	En une journée, Combien de temps consacrez-vous à votre enfant pour accomplir les devoirs à la maison ?	Une variable indicatrice à 5 modalités : 1 : 0 minute ; 2 : moins ou égal à 15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31 -60 minutes ; 5 : plus que 60 minutes
Asbh10f	L'école fait un bon travail pour améliorer le niveau de lecture de mon enfant	Une variable indicatrice à 4 modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
Asbh09a	Les parents discutent avec les enfants de leurs devoirs à la maison	Une variable indicatrice à 4 modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
Asbh09f	les parents aident leurs enfants à lire à la maison	Une variable indicatrice à 4 modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
Asbh14	Combien existe-t-il de livres à votre maison ?	Une variable indicatrice à 5 modalités : 1 : 0-10 livres ; 2 : 11-25 livres ; 3 : 26-100 livres ; 4 : 101-200 livres ; 5 : plus de 200 livres
Asbh17	Quel est le plus haut niveau de scolarité atteint par les parents ?	Une variable indicatrice à 6 modalités : 1 : aucun diplôme ; 2 : diplôme primaire ; 3 : diplôme collégial ; 4 : diplôme lycéen ; 5 : licence ; 6 : Master ou Doctorat
Asbh19	Quel est le type de revenu des parents ?	Une variable indicatrice à 4 modalités : 1 : Fonction avec revenu ; 2 : fonction avec revenu partiel ; 3 : sans revenu ; 4 : autre fonction
Variables liées au contexte des pratiques d'enseignement (Enseignant)		
atbg01	Le nombre d'année d'expérience du professeur	Une variable continue indiquant le nombre d'année d'expérience de l'enseignant
atbg02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : Homme ; 2 : femme
atbg03	L'âge de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
atbg04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collège ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : Doctorat.
atbg06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
atbg06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
atbg06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
atbg07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
atbg08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
atbg08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
atbg11a	Le degré de satisfaction de L'enseignant dans son travail d'enseignement	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord

atbg11e	L'enseignant envisage de continuer dans son travail d'enseignant pour une durée indéterminée	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
atbg12	Le nombre d'élève dans la classe	Une variable continue indiquant la taille de classe
atbg16a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
atbg16e	Est-ce que la présence des élèves qui font du bruit détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
atbg17a	Le niveau de rencontre des parents pour discuter du progrès des élèves concernant leurs apprentissages	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
atbr02	Le temps consacré à la lecture pendant une semaine	Une variable continue indiquant le nombre d'heures consacré à la lecture en classe pendant une semaine
atbr08f	L'enseignant apprend aux élèves du nouveau vocabulaire d'une façon régulière	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr08b	L'enseignant demande à chaque élève de lire à haute voix	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr09b	L'enseignant demande aux élèves de révéler les idées principales d'un texte	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr10a	L'enseignant demande aux élèves d'écrire quelque phrase sur un texte déjà traité	Une variable indicatrice à cinq modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
atbr14	Combien de fois l'enseignant demande aux élèves de préparer un texte comme un devoir de lecture à la maison	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1-2 fois par semaine ; 4 : entre 3-4 fois par semaine ; 5 : chaque jour
atbr18c	L'enseignant Consacre plus de temps avec un élève qui a un niveau faible	Une variable indicatrice à 2 modalités : 1 : Oui ; 2 : Non
atbr19a	Le degré d'importance consacré par L'enseignant pour l'évaluation régulière de l'apprentissage des élèves	Une variable indicatrice à trois modalités : 1 : une grande importance ; 2 : peu d'importance ; 3 : aucune importance
atbr21	Pendant les deux dernières années, combien d'heures ont été consacré par l'enseignant pour la formation continue	Une variable indicatrice à cinq modalités : 1 : aucune heure ; 2 : moins de 6 heures ; 3 : entre 6-15 heures ; 4 : entre 16-35 heures ; 5 : plus de 35 heures
Variables liées au contexte de l'environnement de l'école (Directeur)		
acbg05b	Quel type du milieu où se trouve l'école ?	Une variable indicatrice à cinq modalités : 1 : urbain ; 2 : semi urbain ; 3 : ville de taille moyenne ; 4 : semi rural ; 5 : rural
acbg06ba	Pendant une journée, Quel est le nombre d'heures consacré à l'éducation ?	Une variable continue indiquant le nombre d'heures consacré l'éducation pendant une journée de la semaine
acbg10ae	Quel est le degré d'existence des lieux consacrés à l'enseignement, et en premier lieu les classes ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
acbg10ba	quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiés à l'enseignement de la lecture	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
acbg11aa	Informers les parents sur le progrès des élèves	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : une fois par an ; 3 : entre 2-3 fois par an ; 4 : plus que 3 fois par an

acbg13ab	Le problème d'absences des élèves	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13ac	Le problème de violence dans la classe	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13ba	Le problème d'arrivée tardive à l'école ou de départ anticipé des enseignants	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg13bb	Le problème d'absence des professeurs	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
acbg15i	Le degré de la discussion du directeur avec les enseignants dont ceux qui ont des difficultés en classe	Une variable indicatrice à trois modalités : 1 : jamais ; 2 : parfois ; 3 : beaucoup
acbg17a	A quel niveau primaire, l'élève a déjà des acquis suffisants pour la connaissance des lettres ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire
acbg17e	A quel niveau primaire, l'élève a déjà des acquis suffisants pour la lecture d'un texte ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire
acbg17g	A quel niveau primaire, l'élève a déjà des acquis suffisants pour déterminer les idées principales du texte ?	Une variable indicatrice à cinq modalités : 1 : préscolaire ou première année primaire ; 2 : deuxième année primaire ; 3 : troisième année primaire ; 4 : quatrième année primaire ; 5 : pas au niveau primaire

Source : D'après l'auteur à partir des données PIRLS 2011

Tableau 48. Descriptions des variables du programme TIMSS 2011

Code TIMSS	Variables	Descriptions des variables
bsmmat	Score des élèves en mathématiques	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2eme année collégial en 2011 variant entre 300 à 600 points
bssbio	Score des élèves en science de la vie	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2eme année collégial en 2011 variant entre 300 à 600 points
bssear	Score des élèves en science de la terre	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2eme année collégial en 2011 variant entre 300 à 600 points
bssche	Score des élèves en chimie	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2eme année collégial en 2011 variant entre 300 à 600 points
bssphy	Score des élèves en physique	Une variable quantitative mesurant le score des acquis scolaire des élèves de la 2eme année collégial en 2011 variant entre 300 à 600 points
Bcbg03a	Dans une classe, quel est le pourcentage des élèves issus d'un milieu défavorisé ?	Une variable indicatrice à quatre modalités : 1 : entre 0-10 % ; 2 : entre 11-25% ; 3 : entre 26-50 % ; 4 : plus de 50 %
Variables intrinsèques à l'élève		
bsbg01	Genre	Une variable dichotomique indiquant le genre de l'élève : 1 : garçon, 0 : fille.
bsbg02b	Age de l'élève en année de 2011	Une variable continue indiquant l'âge de l'élève
bsbg04	Le nombre approximatif de livres dans votre maison	Une variable indicatrice à cinq modalités : 1 : entre 0 -10 ; 2 : entre 11-25 ; 3 : 26-100 ; 4 : entre 101-200 ; 5 : entre plus que 200
bsbg05d	Disponibilité d'une chambre propre à l'élève	Une variable dichotomique indiquant la disponibilité d'une chambre propre à l'élève : 1 : Oui, 0 : non.

bsbg06	Le niveau d'instruction des parents	Une variable indicatrice à sept modalités : 1 : niveau primaire ou analphabète ; 2 : primaire ; 3 : collège ; 4 : lycée ; 5 : licence ; 6 : Master ; 7 : doctorat
bsbg10a	Combien de fois l'élève utilise-t-il l'ordinateur à la maison ?	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : entre 1-2 fois par semaine ; 3 : entre 1-2 fois par mois ; 4 : jamais
bsbg11d	Les parents s'assurent si l'élève a bien accompli ses devoirs à la maison	Une variable indicatrice à quatre modalités : 1 : chaque jour ; 2 : 1 à 2 fois par semaine ; 3 : 1 à 2 fois par mois ; 4 : jamais
bsbg12a	L'élève a le sentiment d'appartenance à l'école	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
bsbg13a	L'élève a été soumis dans des situations d'insultes par d'autres élèves	Une variable indicatrice à quatre modalités : 1 : au moins une fois par semaine ; 2 : 1 à 2 fois par mois ; 3 : quelque fois pendant l'année ; 4 : jamais
bsbg14a	L'élève a un sentiment de plaisir pour l'apprentissage des mathématiques	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
bsbg16a	Habituellement, l'élève travaille dur pour l'apprentissage des mathématiques	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg17a	L'élève a un sentiment de plaisir pour l'apprentissage des sciences de la vie	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg19a	Habituellement, l'élève travaille dur pour l'apprentissage des sciences de la vie	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg18a	L'élève a un sentiment de plaisir pour l'apprentissage des sciences de la terre	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg20a	Habituellement, l'élève travaille dur pour l'apprentissage des sciences de la terre	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
bsbg26a	L'élève a un sentiment de plaisir pour l'apprentissage de la chimie	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
bsbg28a	Habituellement, l'élève travaille dur pour l'apprentissage de la chimie	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg30a	L'élève a un sentiment de plaisir pour l'apprentissage de la physique	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg32a	Habituellement, l'élève travaille dur pour l'apprentissage de la physique	Une variable indicatrice à quatre modalités : 1 : d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : Totalement en désaccord
Bsbg33ba	En une journée, Combien de temps l'élève consacre-t-il aux devoirs des mathématiques	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 : entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bb	En une journée, Combien de temps l'élève consacre-t-il aux devoirs des sciences de la vie ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 : entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bc	En une journée, Combien de temps l'élève consacre-t-il aux devoirs des sciences de la terre ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 : entre 61-90 minutes ; 6 : plus que 90 minutes.
Bsbg33bd	En une journée, Combien de temps l'élève consacre-t-il aux devoirs de la chimie ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 : entre 61-90 minutes ; 6 : plus que 90 minutes.

Bsbg33be	En une journée, Combien de temps l'élève consacre-t-il aux devoirs de la physique ?	Une variable indicatrice à six modalités : 1 : pas de devoir ; 2 : entre 1-15 minutes ; 3 : entre 16-30 minutes ; 4 : entre 31-60 minutes ; 5 : entre 61-90 minutes ; 6 : plus que 90 minutes.
Variables liées au contexte des pratiques d'enseignement (Enseignement des mathématiques)		
btbg01	Le nombre d'année d'expérience du professeur	Une variable continue indiquant le nombre d'année d'expérience de l'enseignant
btbg02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : Homme ; 2 : femme
btbg03	L'âge de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
btbg04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collège ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : Doctorat.
btbg06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbg06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbg06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbg07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbg08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbg08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbg11a	L'enseignant a un sentiment de fierté dans son travail d'enseignement	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbg11e	L'enseignant envisage de continuer dans son travail d'enseignant pour une durée indéterminée	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbg12	Le nombre d'élève dans la classe	Une variable continue indiquant la taille de classe
btbg15a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbg15e	Est-ce que la présence des élèves qui font du bruit détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbg16a	Le niveau de rencontre des parents pour discuter du progrès des élèves concernant leurs apprentissages	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
btbg17a	Pendant une semaine, Le temps consacré à l'enseignement des mathématiques	Une variable continue indiquant le nombre d'heures consacré à l'enseignement des mathématiques pendant une semaine

btbg19b	Combien de fois l'enseignant demande aux élèves d'apprendre les règles mathématiques ?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
btbg19c	Combien de fois l'enseignant corrige les exercices en classe ?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
Btbg25a	Combien de fois l'enseignant demande aux élèves de préparer des devoirs de mathématique à la maison	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1 -2 fois par semaine ; 4 : entre 3-4 fois par semaine ; chaque jour
Btbg27	combien de fois l'enseignant programme des examens en mathématique	Une variable indicatrice à cinq modalités : 1 : une fois par semaine ; 2 : une fois dans deux semaine ; 3 : une fois par mois ; 4 : quelque fois par ans ; 5 : jamais
Btbg29b	Pendant les deux dernières années ; combien de fois l'enseignant a assisté à des formations continue pour l'amélioration des pratiques pédagogiques d'enseignement des mathématiques	Une variable dichotomique à deux modalités. 1 : Oui ; 2 : Non
Variabiles liées au contexte des pratiques d'enseignement (Enseignement des sciences)		
btbs01	Le nombre d'année d'expérience du professeur	Une variable continue indiquant le nombre d'année d'expérience de l'enseignant
btbs02	Le genre de l'enseignant	Une variable indicatrice à 2 modalités : 1 : Homme ; 2 : femme
btbs03	L'âge de l'enseignant	Une variable indicatrice à 6 modalités : 1 : moins de 25 ans ; 2 : entre 25-29 ans ; 3 : entre 30 -39 ans ; 4 : entre 40-49 ans ; 5 : entre 50 -59 ans ; 6 : plus que 60 ans.
btbs04	Quel est le plus haut niveau de scolarité atteint par l'enseignant ?	Une variable indicatrice à 6 modalités : 1 : niveau primaire ; 2 : collège ; 3 : lycée ; 4 : licence ; 5 : Master ; 6 : Doctorat.
btbs06a	Le niveau de satisfaction des enseignants dans leur travail	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbs06b	Le degré de connaissance des objectifs concernant les programmes scolaires.	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbs06h	Le niveau de désir des élèves pour une réussite scolaire	Une variable indicatrice à 5 modalités : 1 : trop élevé ; 2 : élevé ; 3 : moyen ; 4 : bas ; 6 : trop bas
btbs07e	Le respect des élèves envers leurs enseignants	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout a fait en désaccord
btbs08b	L'avis des enseignants par rapport aux problèmes d'encombrement des classes	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbs08c	L'avis des enseignants par rapport aux problèmes de la charge horaire	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
btbs11a	L'enseignant a un sentiment de fierté dans son travail d'enseignement	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbs11e	L'enseignant envisage de continuer dans son travail d'enseignant pour une durée indéterminée	Une variable indicatrice à quatre modalités : 1 : tout à fait d'accord ; 2 : un peu d'accord ; 3 : un peu en désaccord ; 4 : tout à fait en désaccord
btbs12	Le nombre d'élève dans la classe	Une variable continue indiquant la taille de classe

btbs15a	Est-ce que le manque de connaissances préalables chez les élèves détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbs15e	Est-ce que la présence des élèves qui font du bruit détériore les pratiques pédagogiques de l'enseignant	Une variable indicatrice à trois modalités : 1 : pas du tout ; 2 : un peu ; 3 : beaucoup
btbs16a	Le niveau de rencontre des parents pour discuter du progrès des élèves concernant leurs apprentissages	Une variable indicatrice à cinq modalités : 1 : au moins une fois par semaine ; 2 : 1-2 fois par mois ; 3 : 4-6 fois par an ; 4 : 1-3 fois par an ; 5 : jamais
btbs17a	Pendant une semaine, Le temps consacré à l'enseignement des sciences	Une variable continue indiquant le nombre d'heures consacré à l'enseignement des mathématiques pendant une semaine
btbs19f	Combien de fois l'enseignant demande aux élèves d'apprendre les règles sciences?	Une variable indicatrice à cinq modalités : 1 : pour l'ensemble des cours ; 2 : pour ½ des cours ; 3 : quelque cours ; 4 : jamais
Btbs24a	Combien de fois l'enseignant demande aux élèves de préparer des devoirs de science à la maison	Une variable indicatrice à cinq modalités : 1 : jamais ; 2 : moins d'une fois par semaine ; 3 : entre 1 -2 fois par semaine ; 4 : entre 3-4 fois par semaine ; chaque jour
Btbs26	combien de fois l'enseignant programme des examens en mathématique	Une variable indicatrice à cinq modalités : 1 : une fois par semaine ; 2 : une fois dans deux semaine ; 3 : une fois par mois ; 4 : quelque fois par ans ; 5 : jamais
Btbs28b	Pendant les deux dernières années ; combien de fois l'enseignant a assisté à des formations continue pour l'amélioration des pratiques pédagogiques d'enseignement des mathématiques	Une variable dichotomique à deux modalités. 1 : Oui ; 2 : Non
Variables liées au contexte de l'environnement de l'école (Directeur)		
bcbg05b	Quel type du milieu où se trouve l'école ?	Une variable indicatrice à cinq modalités : 1 : urbain ; 2 : semi urbain ; 3 : ville de taille moyenne ; 4 : semi rural ; 5 : rural
bcbg06ba	Pendant une journée, Quel est le nombre d'heures consacré à l'éducation ?	Une variable continue indiquant le nombre d'heures consacré l'éducation pendant une journée de la semaine
bcbg08a	Existe-t-il un laboratoire pour mener des expériences scientifiques par les élèves ?	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
bcbg09ae	Quel est le degré d'existence des lieux consacrés à l'enseignement, et en premier lieu les classes ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
bcbg09ba	Quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiés à l'enseignement des mathématiques	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
bcbg09ca	Quel est le degré d'existence des enseignants spécialisés dans les pratiques pédagogiques dédiés à l'enseignement des sciences ?	Une variable indicatrice à quatre modalités : 1 : inexistant ; 2 : un peu existant ; 3 : existence moyenne; 4 : existence parfaite
Bcbg10ca	Informers les parents sur le progrès des élèves	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : une fois par an ; 3 : entre 2-3 fois par an ; 4 : plus que 3 fois par an

Bcbg12ab	Le problème d'absences des élèves	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12ac	Le problème de violence dans la classe	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12ba	Le problème d'arrivée tardive à l'école ou de départ anticipé des enseignants	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg12bb	Le problème d'absence des professeurs	Une variable indicatrice à quatre modalités : 1 : pas un problème ; 2 : simple problème ; 3 : problème moyen ; 4 : grave problème
Bcbg16a	Le directeur de l'école motive les enseignants des mathématiques par des indemnités pour les encourager dans leurs travaux d'enseignements	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
Bcbg16b	Le directeur de l'école motive les enseignants des sciences par des indemnités pour les encourager dans leurs travaux d'enseignements	Une variable indicatrice à deux modalités : 1 : Oui ; 2 : non
Bcbg17d	Combien de temps consacre le directeur de l'école pour contrôler l'évolution des apprentissages des élèves pour atteindre les objectifs scolaires.	Une variable indicatrice à quatre modalités : 1 : jamais ; 2 : parfois ; 3 : toujours

Source : D'après l'auteur à partir des données TIMS 2011

Conclusion générale

Dans un contexte mondialisé de plus en plus concurrentiel aussi bien pour le secteur privé que pour les gouvernements et les organisations de la société civile, la question de la valorisation du capital humain est devenu incontournable dans les schémas de développement, et ce aussi bien pour les pays industrialisés que pour les pays en développement. La question de la formation et l'accumulation du capital humain fait nécessairement appel à la notion de rendements des systèmes éducatifs.

Dans tous les pays du monde, le secteur de l'éducation nationale revêt une importance cruciale. Il permet d'abord d'endiguer le phénomène de l'analphabétisme, mais il permet aussi et surtout l'éducation et la formation de générations futures et leur préparation aux défis du monde du travail. La réussite de tout système éducatif se mesure à l'aune du niveau de son rendement, lequel peut être lui-même appréhendé à travers la mesure des acquis scolaires des élèves.

Au Maroc, le dernier discours Royal ⁵¹a constitué une déclaration sans équivoque de la faillite du système éducatif dans notre pays. C'est une déclaration courageuse qui ouvre la voie à une véritable réforme du secteur de l'éducation nationale sur la base d'un diagnostic objectif qui énumère les faiblesses et les contraintes dont souffre le secteur et qui propose les solutions adaptées pour y faire face dans un cadre de réflexion plus globale celui de l'instauration d'un nouveau modèle de développement.

Les politiques publiques dans le domaine de l'éducation nationale et de réformes du système éducatif peuvent être très diversifiées et dépendent des ressources budgétaires disponibles mais aussi et surtout des objectifs arrêtés par les pouvoirs publics. Faut-il par exemple améliorer les conditions de fonctionnement des établissements scolaires existants ou bien construire de nouvelles unités pour augmenter le taux d'accès à l'éducation à l'échelle nationale ? La gestion des établissements scolaires devrait-elle être centralisée ou, au contraire, faudrait-il donner plus d'autonomie aux proviseurs et directeurs d'établissements scolaires au niveau local ? Les réponses à ces interrogations seraient à même de cerner les contours des politiques publiques à adopter dans l'objectif ultime d'améliorer les acquis scolaires des élèves et du rendement du système éducatif. Mais au-delà du cadre national, et au vu de la grande similitude des

⁵¹ Les deux discours adressés par le Souverain à l'occasion du 20e anniversaire de l'Accession de S.M. le Roi au Trône et du 66^e anniversaire de la Révolution du Roi et du Peuple. (20 aout 2019)

problématiques au niveau régional et parfois même au niveau mondial, l'étude de l'expérience internationale peut être très intéressante à cet égard.

La comparaison internationale en termes de rendement des systèmes éducatifs est en effet très édifiante, notamment en présence de résultats d'enquêtes internationales en la matière telles que par exemple les enquêtes PIRLS, TIMSS, PISA et bien d'autres.

Les résultats de ces enquêtes peuvent être utiles à plusieurs égards. Ils permettent tout d'abord d'effectuer des comparaisons des niveaux de compétence des élèves sur le plan international. Ils permettent également aux décideurs de se fixer des objectifs d'amélioration, comme par exemple celui d'atteindre les scores moyens d'autres pays ou de parvenir à un degré d'équité plus élevé en termes de perspectives et de résultats éducatifs. Enfin, ces résultats permettent de comprendre les points forts et les points faibles des différents systèmes éducatifs. Ils représentent ainsi une source interminable d'information pour les chercheurs intéressés par l'étude et la comparaison des performances scolaires et des rendements des systèmes éducatifs à l'international.

L'objectif principal de cette thèse a été d'analyser les déterminants de la performance scolaire (chapitre 1) en accordant une attention toute particulière à deux facteurs déterminants de la réussite scolaire à savoir la taille de la classe (**chapitre 2**) et la mixité sociale (**chapitre 3**). Le chapitre 1 a tenté de mettre en évidence les différents facteurs déterminant les performances scolaires. L'objectif a été d'étudier non seulement l'influence des caractéristiques individuelles de l'élève, mais aussi de révéler l'effet de la taille des classes et la mixité sociale sur l'amélioration de la performance scolaire des élèves, en développant un modèle économétrique robuste au problème de l'endogénéité. Dans le chapitre 2, nous nous sommes focalisés sur la détermination de la taille optimale de la classe à partir de laquelle la performance scolaire se détériore de manière significative en tenant compte de contextes divers. Enfin dans le chapitre 3, notre analyse a porté sur le degré de la composition sociale et économique à l'école qui assure une performance scolaire généralisée de tous les élèves.

Ces analyses ont été réalisées en mobilisant différents outils méthodologiques selon l'objectif considéré : une modélisation multiniveau avec la correction du biais d'endogénéité dans le chapitre 1, un modèle multiniveau semi-paramétrique dans le chapitre 2 et un modèle multiniveau semi-paramétrique conditionnel au contexte dans le chapitre 3.

Dans cette conclusion générale, il y a quatre points essentiels qui s'imposent, et que nous ne pouvons pas les négliger : tout d'abord les résultats, les apports de la thèse, quelques implications de la politique éducative, et enfin les limites et les prolongements possibles de la thèse.

Les résultats du **chapitre 1** montrent d'abord que la performance scolaire des élèves marocains est influencée par plusieurs facteurs, dont les caractéristiques intrinsèques à l'élève, et aussi bien le rôle déterminant de l'école pour une réussite scolaire des élèves. Ce résultat n'est pas nouveau, car plusieurs travaux empiriques (Hanchane, Benbiga, Idir et Mostafa 2012) ont abouti aux mêmes résultats. Notre étude basée sur l'estimation du modèle multiniveaux avec la correction du biais d'endogénéité à l'aide de la méthode des variables instrumentales met en lumière le résultat essentiel suivant : la taille de la classe et la composition sociale et économique de l'école provoquent des différences de performance scolaire assez prononcées et très significatives entre les élèves marocains. De plus, et à partir des résultats du premier chapitre, il apparaît que ces derniers facteurs liés à l'établissement scolaire sont eux qui révèlent les impacts les plus élevés par rapport aux variables intrinsèques à l'élève. Cette conclusion du premier chapitre nous a motivé à approfondir davantage notre analyse afin d'identifier la taille optimale de la classe qui assure une égalité des résultats entre les élèves marocains.

En effet, à partir l'estimation d'un modèle semi-paramétrique sous une spécification hiérarchique, le résultat important du **chapitre 2** réside dans le fait que la taille optimale de la classe qui assure une performance scolaire élevé varie d'un contexte à un autre, mais cette variation est limitée entre 20 à 27 élèves. Cela dit, quel que soit le contexte où l'élève est assigné, la taille optimale qui permet une performance scolaire juste et généralisée entre l'ensemble des élèves marocains ne doit pas dépasser les 27 élèves par classe. Cela a été démontré à partir d'ensemble de combinaisons entre les différentes variables pertinentes disponibles dans les deux programmes TIMSS et PIRLS. Autrement dit, si la taille des classes dépasse les 27 élèves, la performance scolaire baisse au fur et à mesure de l'augmentation de la taille de la classe.

Ce résultat permet de souligner la pertinence d'une politique publique visant la réduction de la taille de la classe afin de créer les conditions d'un environnement pédagogique favorable pour l'apprentissage des élèves et pour le métier d'enseignement, et par conséquent assurer une performance scolaire comparable à l'échelle internationale.

Par la suite, le modèle multiniveau semi-paramétrique catégoriel proposé dans **le chapitre 3** montre que la mixité sociale à l'école est bénéfique pour favoriser la réussite, et plus largement, la future intégration des élèves dans la société. Remédier aux problèmes de la ségrégation scolaire passe naturellement par un assouplissement de la carte scolaire, voire même une suppression de cette dernière dans des contextes ciblés. Autrement dit, un système scolaire ségrégatif engendre inévitablement des différences très prononcées au niveau des acquis des élèves. Aussi, cette thèse défend un modèle d'école fondé sur deux principes essentiels :

- La mixité sociale apparaît comme un instrument favorable à la réussite des élèves
- Une taille de classe réduite, adéquate aux méthodes pédagogiques et qui apparaît comme optimale pour fournir un environnement efficace la scolarisation de nos élèves

Plusieurs travaux empiriques traitant la problématique de l'inégalité des résultats, montrent que la démocratisation de l'école n'a pas été réalisée. Ils montrent également que l'amélioration de la qualité du système éducatif marocain suppose une parfaite synergie et une forte implication de l'ensemble des intervenants qui s'activent dans le processus éducationnel de l'élève , en occurrence : - Les parents, à travers la réactivation des associations des tuteurs au sein des écoles, - Les enseignants, qui doivent bien assimiler les objectifs pédagogiques de l'école et réussir la mise en œuvre des curricula, - Les directeurs des établissements scolaires, pierre angulaire de l'édifice éducationnel et garants de l'accomplissement du succès académique des élèves.

Dans cette thèse, on démontre que tous les facteurs essentiels pour l'amélioration des acquis scolaire des élèves passent en premier lieu par une offre scolaire équitable, caractérisée par des tailles de classes maîtrisées et par plus de mixité sociale à l'école.

En effet, les résultats concluants dans cette thèse permettront d'éclairer les politiques publiques sur les choix et les priorités à faire valoir, pour améliorer la qualité du système éducatif et instaurer une justice sociale inclusive en matière d'éducation. De même, cela incite les pouvoirs publics à instaurer un débat public profond sur la carte scolaire et les mécanismes à travers lesquels un assouplissement de cette dernière s'avère nécessaire. Ainsi, des analyses en termes d'expérimentation sont nécessaires pour appuyer les effets positifs des conclusions défendues dans cette thèse à savoir la réduction de la taille de la classe et l'assouplissement de la carte scolaire pour garantir plus de liberté aux parents pour le choix des écoles de leurs enfants. Des

expérimentations contrôlées dans des régions ou provinces ciblées méritent d'être instaurées dans le but d'appliquer et tester d'abord la pertinence de ces choix de politiques éducatives pour ensuite envisager de les généraliser sur l'ensemble du système éducatif. En effet, ces expérimentations renvoient aux travaux du prix Nobel d'économie (Duflo, 2019), et sur la généralisation ou non d'une telle politique publique.

Si la démocratisation scolaire est une notion complexe à définir, cette thèse essaye en quelque sorte de mettre en relief les politiques éducatives et les mécanismes utiles à travers desquels une relecture critique de l'article 31 de la constitution de 2011 s'avère nécessaire pour concrétiser réellement le droit d'accès des citoyens à une éducation moderne, accessible et surtout de qualité.

Pour conclure, nous pouvons dire que cette thèse en plus d'être à travers lequel la fonction de production du capital humain a été revisitée sur le plan méthodologique a permis de dégager les facteurs incontournables capables de garantir l'égalité de résultats, et de contribuer de manière directe et effective à la qualité de l'éducation.

Bibliographie

Bibliographie

- Alan Krueger and Diane Whitmore, The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR, NBER Working Paper No. 7656 ,Issued in April (2000).
- Alan Krueger and Diane Whitmore, The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR, NBER Working Paper No. 7656 ,Issued in April (2000).
- Alain Mingat , Expliquer la variété des acquisitions au cours préparatoire : les rôles de l'enfant, la famille et l'école, Revue française de pédagogie, 1991
- Atila Abdulkadiroğlu and Tayfun Sönmez, School Choice: A Mechanism Design Approach, vol. 93, no. 3, June (2003), (pp. 729-747)
- Akee, R. K., Copeland, W. E., Keeler, G., Angold, A., & Costello, E. J. (2010). Parents' incomes and children's outcomes: a quasi-experiment using transfer payments from casino profits. *American Economic Journal: Applied Economics*, 86-115.
- Arnaud Chevalier and Gauthier Lanot , (2002), The Relative Effect of Family Characteristics and Financial Situation on Educational Achievement ,*Education Economics*
- Aslan, D., & Aktaş Arnas, Y. (2015). The immediate impacts of preschool attendance on Turkish children's mathematics achievement. *Educational Studies*, 41(3), 231-243.
- Al-Samarrai, S. (2002) Achieving education for all: how much does money matter? Working paper series, 175. Brighton: IDS.
- Alan B. Krueger,(1999), Experimental Estimates of Education Production Functions, *The Quarterly Journal of Economics*, Volume 114, Issue 2, May, Pages 497–532
- Alan B. Krueger,2003, Economic Considerations and Class Size,*The economic journal*
- Sandoval-Hernandez, A., Taniguchi, K., & Aghakasis, P. (2013, Décembre). Is participation in preschool education associated with higher student achievement? (IEA, Éd.) IEA's Policy Brief Series (2), pp. 1-8.
- Aomar Ibourk ,2012, Les déterminants des acquis scolaires au Maroc, analyse micro économétrique appliquée aux enquêtes internationales TIMSS et PIRLS, ERF
- AitKin & Longford (1986) Statistical Modelling Issues in School Effectiveness Studies, *Journal of the Royal Statistical Society Series A (General)*
- Angrist and Lavy (1999),Using Maimonides' Rule To Estimate the Effect of Class Size on Scholastic Achievement,*MIT Economics*
- Atila Abdulkadiroğlu Tayfun Sönmez, (2003), School Choice: A Mechanism Design Approach,*American Economic Review*
- Benbiga, A ; Hanchane, S ; Idir, N ; Mostafa, T ; (2012) Les déterminants des performances scolaires des élèves marocains, *Maghreb Machrek*
- Bressoux, Kramarz et Prost (2009) ,Teachers' Training, Class Size and Students' Outcomes: Learning from Administrative Forecasting Mistakes,*The Economic Journal*
- Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92, (5-6), 1416-1440.
- Bouhlila, D. S., & Sellaouti, F. (2013). Multiple imputation using chained equations for missing data in TIMSS: a case study. *Largescale Assessments in Education*, 1(1), 1-33.
- Baltagi, B. H., Bresson, G., & Pirotte, A. (2012). A Robust Hausman–Taylor Estimator. *Advances in Econometrics* (29), 175-214.
- Baltagi, B. H., Bressonb, G., & Pirot, A. (2003). F ixed effects, random effects or Hausman–Taylor? A pretest estimator. *Economics Letters* (79), 361-369.
- Bussière. & Gluszynski T. (2004) Les incidences de l'utilisation de l'ordinateur sur la capacité de lecture des jeunes de 15 ans, Direction générale de la politique sur l'apprentissage Politique stratégique et planification Ressources humaines et Développement des compétences Canada
- Baxter, J. (2002). How much does parental education explain educational attainment of males and females in Australia? *Negotiating the Life Course*.

- Black S.E., Devereux P.J., Salvian K.G. (2008), «Too Young to Leave the Nest: The Effects of School Starting Age». NBER Working Paper Series (No. 13969).
- Bruce Sacerdote ,May (2001), Peer Effects with Random Assignment: Results for Dartmouth Roommates, *The Quarterly Journal of Economics*, Volume 116, Issue 2, , Pages 681–704,
- Chen K, Jin Z. Local polynomial regression analysis for clustered data. *Biometrika*. (2005); 92:59–74.
- Chang-Tai Hsieh and Miguel Urquiola , The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program, *Journal of Public Economics*, (2006), vol. 90, issue 8-9, 1477-1503
- Crahay M. & Monseur C. (2006). « Différences individuelles et effets d'agrégation en ce qui concerne les performances en lecture. Analyse secondaire des données PISA 2000 ». In C. Houssemand, R. Martin & P. Dickes, *Perspectives de psychologie différentielle*. Rennes : Presses universitaires de Rennes, p. 23-34.
- Christian Monseur and Marcel Crahay, *Composition académique et sociale des établissements, efficacité et inégalités scolaires : une comparaison internationale*, *Revue française de pédagogie* ,2008
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). *Equality of educational opportunity*. Washington, dc, 1066-5684.
- Chevalier, A., & Lanot, G. (2002). The Relative Effect of Family Characteristics and Financial Situation on Educational Achievement. *Education Economics*, 10 (Issue 2), 165-181.
- Cooper, H., & Valentine, J. C. (2001). Using research to answer practical questions about homework. *Educational Psychologist*, 36(3), 143–153.
- Crede, J., Wirthwein, L., McElvany, N., & Steinmayr, R. (2015). Adolescents' academic achievement and life satisfaction: the role of parents' education. *Frontiers in psychology*, 6, 52.
- Cooper, H., & Valentine, J. C. (2001). Using research to answer practical questions about homework. *Educational Psychologist*, 36(3), 143-153.
- Cortázar, A. (2015). Long-term effects of public early childhood education on academic achievement in Chile. *Early Childhood Research Quarterly*, 32, 13-22.
- Cobb-Clark, D. A., & Jha, N. (2016). Educational achievement and the allocation of school resources. *Australian Economic Review*, 49(3), 251-271.
- Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2013). The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, 2(1), 8.
- Chiu, J., Economos, J., Markson, C., Raicovi, V., Howell, C., Morote, E. S., & Inserra, A. (2016). Which matters most? Perceptions of family income or parental education on academic achievement. *New York Journal of Student Affairs*, 16(2), 3-16.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2010). How does your kindergarten classroom affect your earnings? Evidence from Project STAR (No. w16381). National Bureau of Economic Research.
- Conseil Supérieur de l'Enseignement. (2009). *Rapport analytique du Programme Nationale d'Evaluation des acquis (P.N.E.A)*.
- Cascio E.U., Lewis E.G. (2006), « Schooling and the Armed Forces Qualifying: Evidence from School-Entry Laws». *Journal of Human Resources*, XLI (2), 294-318.
- Chapman, J. W., & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, 87(1), 154–167.
- Caroline Hoxby,(2000), *Peer Effects in the Classroom: Learning from Gender and Race Variation*, NBER ,Labor Studies, Public Economics
- Chetty Raj, Friedman John, Hilger Nathaniel, Saez Emmanuel, Schanzenbach Diane, Yagan Danny. (2011), *How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR*. *Quarterly Journal of Economics*. ;126(4) :1593–660.
- Caille, J.-P. (2001). *Scolarisation à 2 ans et réussite de la carrière scolaire au début de l'école élémentaire*, *Education et Formations* (60).
- Denis Meuret,(2001), *La justice du système éducatif*, *Revue française de pédagogie*

- Daniel J. Henderson, Raymond J. Carrollb, Qi Li, (2008), Nonparametric estimation and testing of fixed effects panel data models, *Journal of econometrics*
- Diane F. Halpern, Camilla P. Benbow, David C. Geary, Ruben C. Gur, Janet Shibley Hyde, Morton Ann Gernsbacher, (2007), *The Science of Sex Differences in Science and Mathematics, Psychological Science in the Public Interest*
- Davis-Kean, P. E. (2005) The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. *Journal of Family Psychology*, 19 (2), 294-304.
- Dickson, M., Gregg, P., & Robinson, H. (2016). Early, late or never? When does parental education impact child outcomes? *The Economic Journal*, 126(596).
- Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: Mediation by family interactions, child aggression, and teenage aspirations. *Merrill-Palmer quarterly* (Wayne State University. Press), 55(3), 224.
- Dufur, M. J., & Parcel, T. L. (2001). Capital at home and at school: Effects on child social adjustment. *Journal of Marriage and the family*, volume 63 (1), 32-47.
- Elbers, C., Lanjouw, J.O. and Lanjouw, P. (2003). Micro-Level Estimation of Poverty and Inequality *Econometrica* 71(1): 355-64.
- Else-Quest, Hyde, & Linn, (2010) , Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis *Psychological Bulletin* 136(1):103-27
- Emler, N. (2001). *Self-esteem: The costs and causes of low self-worth*. York Publishing Services.
- Erika Hoff, (2003), *Child Development, The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech*, Volume 74, Number 5, Pages 1368–1378,
- Eskil Heinesen ,*Estimating Class-size Effects using Within-school Variation in Subject-specific Classes* , *The economic journal*, 01 June (2010)
- Fack, G. & Grenet, J. (2009) "Sectorisation des collèges et prix des logements à Paris", *Actes de la Recherche en Sciences Sociales*, no 180, pp. 44-63, December
- Fuchs, T., & Wößmann, L. (2008). What accounts for international differences in student performance? A reexamination using PISA data (pp. 209-240). *Physica-Verlag HD*.
- Fan, H., Xu, J., Cai, Z., He, J., & Fan, X. (2017). Homework and students' achievement in math and science: A 30-year metaanalysis, 1986–2015. *Educational Research Review*, 20, 35-54.
- Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-term effects of class size. *The Quarterly Journal of Economics*, 128(1), 249-285.
- Flashman, J. (2012, August). Academic Achievement and Its Impact on Friend Dynamics. *Sociology of Education*, 85 (1), pp. 61-80.
- Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. *Review of educational research*, 64(1), 119-157.
- Fusaro, J. A. (1997). The Effect of Full-Day Kindergarten on Student Achievement: A Meta-Analysis. *Child study journal*, 27(4), 269-77.
- Frans van der Slik, Geert Driessen & Kees de Bot, Ethnic and socioeconomic class composition and language proficiency: A longitudinal multilevel examination in Dutch elementary schools, *European Sociological Review* Volume 22 Number 3 July (2006) 293–308 293
- Fehrman, P. G., Reimers, T. M., Keith, T. Z., Aubey, L. W., & Pottebaum, S. M. (1986). Parental involvement, homework, and TV time: Direct and indirect effects on high school achievement. *Journal of Educational Psychology*, 78 (5), 373-380.
- Fouzi Mourji, Abdelilah Abbaia ,(2013), Les déterminants du rendement scolaire en mathématiques chez les élèves de l'enseignement secondaire collégial au Maroc : une analyse multiniveaux, *Revue d'économie du développement*

- Goldstein et Blatchford , Research on class size effects: A critique of methods and a way forward , International Journal of Educational Research, November (1998)
- Gilotte O., Girard P., (2005). – « La sectorisation, l’affectation et l’évitement scolaire dans les classes de sixième à Paris en 2003 », Éducation et formations, 71, pp. 137-149.
- Grisay A. (1993) Le fonctionnement des collèges et ses effets sur les élèves de 6e et 5e. Les Dossiers d’Éducation et formations, n° 32.
- Gilotte Olivier & Girard Pierre (2005). « Sectorisation, l’affectation et l’évitement scolaire dans les classes de sixième à Paris en 2003 ». Éducation et Formations, n° 71, p. 137-149.
- Gabrielle Fack, Julien Grenet, (2010) , Que peut-on attendre de la réforme de la sectorisation en France ? Quelques enseignements des politiques de choix scolaire ,Revue d’économie politique
- Greenwald, R., & Hedges, L. V. (1996). The Effect of School Resources on Student Achievement. Review of Educational Research, 66 (3), 361-396.
- Gibbons, S., & Telhaj, S. (2016). Peer effects: Evidence from secondary school transition in England. [63] Mary A. Burke and Tim R. Sass, "Classroom Peer Effects and Student Achievement," Journal of Labor Economics 31, no. 1 (January 2013): 51-82.
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. Review of Educational Research, 66(3), 361–396.
- Ganzach, Y. (2000). Parents’ education, cognitive ability, educational expectations and educational attainment: Interactive effects. British Journal of Educational Psychology, 70 (Issue 3), 419-441.
- George Y. Wong and William M. Mason, (1985), The Hierarchical Logistic Regression Model for Multilevel Analysis,Journal of the American Statistical Association
- Grégoire, D.A., Shepherd, D.A., & Lambert, L.S. (2010). Measuring opportunity-recognition beliefs. Organizational Research Methods, 13: 114-145.
- Hulin Wu, Jin-Ting Zhang, Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches, Wiley series in probability and statistics Apr (2006)
- Hair, N. L., Hanson, J. L., Wolfe, B. L., & Pollak, S. D. (2015). Association of child poverty, brain development, and academic achievement. JAMA pediatrics, 169(9), 822-829.
- Hanchane et Mostafa,(2012), Solving endogeneity problems in multilevel estimation: An example using education production functions, Journal of Applied Statistics
- Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. Child Development, 74 (5), 1368–1378.
- Hwang, M. H., Choi, H. C., Lee, A., Culver, J. D., & Hutchison, B. (2016). The Relationship Between Self-Efficacy and Academic Achievement: A 5-Year Panel Analysis. The Asia-Pacific Education Researcher, 25(1), 89-98.
- Hong, K., & Zimmer, R. (2016). Does Investing in School Capital Infrastructure Improve Student Achievement? Economics of Education Review, 53, 143-158.
- Häkkinen, I., Kirjavainen, T., & Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. Economics of Education Review (22), 329-335.
- Michaelowa, K. (2001). Primary education quality in francophone Sub-Saharan Africa: Determinants of learning achievement and efficiency considerations. World Development, 29(10), 1699-1716.
- Heyneman, S. P., & Loxley, W. A. (1983). The effect of primary school quality on academic achievement across twenty-nine high and low-income countries. American Journal of sociology, 88(6), 1162-1194.
- H. Goldstein, (1986), Multilevel mixed linear model analysis using iterative generalized least squares, Biometrika
- Hanushek, E. A. (1997). Assessing the effects 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. (2006). School resources. *Handbook of the Economics of Education*, 2, 865-908.
- Hanushek, E. A., & Woessmann, L. (2017). School Resources and Student Achievement: A Review of Cross-Country Economic Research. In *Cognitive Abilities and Educational Outcomes* (pp. 149-171). Springer International Publishing.
- Hoxby, C. M. (2000, August). Peer Effect in the Class Room: Learning From Gender and From Race Variation. NBER Working Paper (7867), pp. 1-54.
- Hiller et Tercieux ;(2014),School choice in france: An assessment of the Affelnet procedure,Ecole d'économie de paris, Working paper
- Hopkins, J. (1995). The Cooperative Elementary School: Effects on Students' Achievement, Attitudes, and Social Relations. *American Educational Research*, 32 (2), 321-351.
- Haveman, R., Wolfe, B., & Spaulding, J. (1991). Childhood events and circumstances influencing high school completion. *Demography*, 28(1), 133-157.
- Hanchane, S., Idiri, N., & Benbiga, A. (2012). Pourquoi les acquis des élèves marocains sont-ils faibles ? *Critiques Economiques* .
- Häkkinen, Kirjavainen, & Uusitalo (2003) , School Resources and Student Achievement Revisited: New Evidence from Panel Data,*Economics of Education Review*
- Hugh Lauder , D. Hughes and S. Watson, The introduction of educational markets in New Zealand: Questions and consequences, *New Zealand Journal of Educational Studies* , 34(1):86-98 · January (1999)
- Iverson, B. K., & Walberg, H. J. (1982). Home environment and learning: A quantitative synthesis. *Journal of Experimental Education*, 50, 144-151.
- Jeremy D. Finn and Charles M. Achilles , *Tennessee's Class Size Study: Findings, Implications, Misconceptions, Educational Evaluation and Policy Analysis* , June 1, (1999)
- Jerry A. Hausman and William E. Taylor , Panel Data and Unobservable Individual Effects *Econometrica*,Vol. 49, No. 6 (Nov., 1981), pp. 1377-1398 (22 pages), The Econometric Society
- Juan Rodriguez-Poo ,Alexandra Soberon , 2017, Nonparametric and Semiparametric Panel Data Models: Recent Developments, *Journal of Economic Surveys*, Vol. 31, Issue 4, pp. 923-960, (2017)
- Joel L.Horowitz and SokbaeLee,(2017), Nonparametric estimation and inference under shape restrictions,*Journal of econometrics*
- Jiming Jiang and P. Lahiri , Mixed model prediction and small area estimation, *An Official Journal of the Spanish Society of Statistics and Operations Research*, (2006)
- Joshua D. Angrist and Kevin Lang, Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program, *American Economic Review* , vol. 94, no. 5, December (2004)
- Jeffrey C. Valentine,David L. DuBois &Harris Cooper ,(2010), The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review, *Journal Educational Psychologist*
- Jennifer Flashman ,(2012), Academic Achievement and Its Impact on Friend Dynamics,*Sociology of education*
- Karen Akerhielm, Does class size matter? , *Economics of Education Review*, (1995), vol. 14, issue 3, 229-241
- Krueger, A. B. (2003). Economic Considerations and Class Size. *The Economic Journal*, 113 (485), F34–F63.
- Keeves, J. P. & Saha, L. J. (1992) “Home Background Factors and Educational Outcomes.” In J. P. Keeves (ed.), *The IEA Study of Science III: Changes in Science Education and Achievement: 1970 to 1984*, Oxford: Pergamon Press, 165-186
- Kingdon, G., & Altinok, N. (2012). New Evidence on Class Size Effects: A Pupil Fixed Effects Approach. *Oxford Bulletin of Economics and Statistics*, 74 (2), 203–234.
- Karen Akerhielm,(1995), Does class size matter?, *Economics of Education Review*,
- K. Chen, Z. Jin, Local polynomial regression analysis of clustered data, *Biometrika* 92 (2005) 59–74.

- Leuven, E., Oosterbeek, H., and Roenning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics*, 110(4):663–693.
- Li G, Wang QH. Empirical likelihood regression analysis for right censored data. *Statistica Sinica*. (2003);13(1):51–68.
- Lochner, L., & Dahl, G. B. (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, 102 (5), 1927-1956.
- Liu, T., Holmes, K., & Albright, J. (2015). Predictors of mathematics achievement of migrant children in Chinese urban schools: A comparative study. *International Journal of Educational Development*, 42, 35-42.
- Lijian Yang, Byeong U Park, Lan Xue & Wolfgang Härdle, Estimation and Testing for Varying Coefficients in Additive Models With Marginal Integration? *Journal of the American Statistical Association*, (2006)
- Lauder H., Kounali D., Robinson T., Goldstein H., Thrupp M., (2007). – « Social class, pupil composition, pupil progress and school performance: an analysis of primary schools », working paper
- LI, Q., ULLAH, A. (1998). Estimating partially linear panel data models with one-way error components. *Econometric Reviews* 17 145–166.
- Lin, X. and Carroll, R. J. (2006). Semiparametric estimation in general repeated measures problems. *Journal of the Royal Statistical Society B* 68, 69-88.
- Le Bastard Landrier, S. (2005). L'expérience subjective des élèves de seconde : influence sur les résultats scolaires et les voeux d'orientation. *L'Orientation Scolaire et Professionnelle*, 34 (2), pp. 143-164.
- Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross– section of countries. *Economica*, 68(272), 465-488.
- Lavy, V., & Angrist, J. D. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics*, 114 (2), 533-575.
- Lee, V. E., & Burkam, D. T. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. Economic Policy Institute, 1660 L Street, NW, Suite 1200, Washington, DC 20036.
- Le, A., & Miller, P. (2002). Educational attainment in Australia: a cohort analysis.
- Lance Lochner Gordon B. Dahl, (2012), The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit, *American economic review*
- Learning Stephen W. Raudenbush, Anthony S. Bryk, (1988), *Methodological Advances in Analyzing the Effects of Schools and Classrooms on Student Review of Research in Education*
- Miguel Urquiola, Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia, *Review of Economics and Statistics*, March 29, (2006)
- Maciej Jakubowski and Pawel Sakowski, Quasi-Experimental Estimates of Class Size Effect in Primary Schools in Poland, *Warsaw University - Faculty of Economic Sciences; CASE - Center for Social and Economic Research, University of Warsaw*, (2006)
- Martin Browning and Eskil Heinesen, Class Size, Teacher Hours and Educational Attainment, *The Scandinavian journal of economics*, (2007)
- María José Lombardía and Stefan Sperlich, Semiparametric inference in generalized mixed effects models, *Journal of the Royal Statistical Society*, (2008)
- MR Chernick, W González-Manteiga, RM Crujeiras, EB Barrios, (2012), Bootstrap methods *International encyclopedia of statistical science*, 169-174
- Mario Francisco-Fernández and Juan M. Vilar-Fernández, Bandwidth selection for the local polynomial estimator under dependence: A simulation study, *Computational Statistics*, December (2005)
- Maria Jose Lombardía and Stefan Sperlich, A new class of Semi-Mixed Effects Models and its Application to Tourist Expenditures in Small Areas, April 17, (2009)
- Mundlak, Y. (1978): On the pooling of time series and cross section data. *Econometrica* 46 :69-85.

- Marie Duru-Bellat, Magali Danner, Séverine Le Bastard-Landrier et Céline Piquée, Les effets de la composition scolaire et sociale du public d'élèves sur leur réussite et leurs attitudes : évaluation externe et explorations qualitatives, Institut de Recherche sur l'Éducation Sociologie et Economie de l'Éducation, IREDU, CNRS-Université de Bourgogne Janvier (2004)
- Marie Christine Opdenakker et Jan Van Damme , Relationship between School Composition and Characteristics of School Process and their Effect on Mathematics Achievement First ,British Educational Research journal , 02 January (2013)
- M Thrupp, Hugh Lauder, Tony Robinson ,(2002), School composition and peer effects, University of Bath
- Marie DURU-BELLAT ; Nathalie MONS ; Bruno SUCHAUT , Organisation scolaire et inégalités sociales de performance : les enseignements de l'enquête PISA, Education et formations (n° 70, décembre 2004)
- Moussa W.S. (2012, November), The Impacts of School Entry Age on Student Achievement: Evidence from New York City Public Schools.
- Mingat, A., (1991), « Expliquer les acquisitions au Cours Préparatoire : les rôles de l'enfant, de la famille et de l'école », Revue Française de Pédagogie, n° 95, pp. 47-63
- Murphy, P. (2000). Equity, assessment, and gender. Gender, policy and educational change: Shifting agenda in the UK and Europe, 134-152.
- Muhlenbruck, L., Cooper, H., Nye, B., & Lindsay, J. J. (1999). Homework and achievement: Explaining the different strengths of relation at the elementary and secondary school levels. Social Psychology of Education, 3 (4), pp. 295-317.
- Memon S, Goswami P, Memon IA. Effects of Schooling and Parental Education on Student's Achievement in Medical Entrance Test. J Liaquat Uni Med Health Sci. (2016); 15(03):116-20.
- Murphy, P. (2000). Equity, assessment, and gender. Gender, policy and educational change: Shifting agenda in the UK and Europe, 134-152.
- Maltese, Tai, & Fan,(2012), When is Homework Worth the Time?: Evaluating the Association Between Homework and Achievement in High School Science and MathThe High School Journal 96(1):52-72
- Michael P. Keane ,Kenneth I. Wolpin,(2001), The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment,International economic review
- Martin Browning ,Eskil Heinesen, (2007), Class Size, Teacher Hours and Educational Attainment, The Scandinavian journal of economics
- Nye, B., Hedges, L.V., & Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. Educational Evaluation and Policy Analysis, 21, 127-142
- Nathaniel Baum-Snow and Byron F. Lutz , School Desegregation, School Choice, and Changes in Residential Location Patterns by Race, American Economic Review , vol. 101, no. 7, December (2011)
- Nadaraya, E.A. (1964) On estimating regression. Theory of Probability and its Applications 10, 186-190
- Nosek, et al. (2009) , National differences in gender-science stereotypes predict national sex differences in science and math achievement, Proceedings of the National Academy of Sciences 106(26):10593-7
- Nechyba, T. J., McEwan, P. J., & Older-Aguilar, D. (1999). The impact of family and community resources on student outcomes: An assessment of the international literature with implications for New Zealand. Ministry of Education.
- Neiss, M. B., Sedikides, C., & Stevenson, J. (2002). Self-esteem: a behavioural genetic perspective. European Journal of Personality, 16(5), 351-367.
- Núñez, J. C., Suárez, N., Rosário, P., Vallejo, G., Valle, A., & Epstein, J. L. (2015). Relationships between perceived parental involvement in homework, student homework behaviors, and academic achievement: differences among elementary, junior high, and high school students. Metacognition and learning, 10(3), 375-406.

- Nadir Altinok Geeta Kingdon ,(2011), New Evidence on Class Size Effects: A Pupil Fixed Effects Approach,Oxford Bulletin of economics and statistics
- Oosterbeek et Roenning (2008) ,Quasi-experimental Estimates of the Effect of Class Size on Achievement in NorwayScandinavian Journal of Economics
- Oliver Linton, Estimation of semiparametric models when the criterion function is not smooth, London School of Economics and Political Science, London School of Economics and Political Science
- Ozkan Eren Daniel J. Henderson, (2008), The impact of homework on student achievement, Econometrics journals
- Pascal Bressoux, Francis Kramarz, Corinne Prost, (2008). « Teachers’ Training, Class Size and Students’ Outcomes: Learning from Administrative Forecasting Mistakes », IZA Discussion Papers 3871, Institute for the Study of Labor (IZA)
- Pascal Bressoux, 2007, L’apport des modèles multiniveaux à la recherche en éducationWhat Multilevel Models Bring to Educational Research, Education didactique
- Peter Diggle ,Patrick Heagerty , Kung-Yee Liang and Scott Zeger, Analysis of Longitudinal Data ,Oxford Statistical Science Series),(2013)
- Peter Hall and Xiao-Hua Zhou ,(2011), Nonparametric estimation of component distributions in a multivariate mixture, The Annals of Statistics 2003, Vol. 31, No. 1, 201–224
- Pelletier, R., & Normore, A. H. (2013). The predictive power of homework assignments on student achievement in mathematics.
- Postlethwaite, T.N. & Wiley, D.E. (1992). The IEA Study of Science 11: Science Achievement in Twenty-Three Countries. Oxford: Pergamon.
- Pelletier, R., & Normore, A. H. (2007). The predictive power of homework assignments on student achievement in mathematics. Proceedings of the Sixth Annual College of Education Research Conference: Urban and International Education, Florida International University.
- Qi LiThanasis Stengos , (1996), Semiparametric estimation of partially linear panel data models ,Journal of econometrics
- Robert Gary-Bobo, Mohamed Mahjoub, (2006). « Estimation of Class-Size Effects, Using ‘Maimonides’ Rule’: The Case of French Junior High Schools », CEPR Discussion Papers 5754, C.E.P.R. Discussion Papers.
- Richard Harker &Peter Tymms , (2010), The Effects of Student Composition on School Outcomes, Journal School Effectiveness and School Improvement
- Russell W. Rumberger et Gregory Palardy,(2005), Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School, Teachers College Record
- Rong Su , James Rounds, Patrick Ian Armstrong, (2009),Men and Things, Women and People: A Meta-Analysis of Sex Differences in InterestsPsychological Bulletin 135(6):859-84
- Robert GARY-BOBO and Mohamed-Badrane MAHJOUB, 2013, “Estimation of Class-Size Effects, Using “Maimonides’ Rule” and Other Instruments: The Case of French Junior High Schools”, Annals of Economics and Statistics, no 111-112, p 193-225.
- Rasmusson, M. A. (2016, February 3). A multilevel analysis of Swedish and Norwegian students’ overall and digital reading performance with a focus on equity aspects of education. Largescale Assessments in Education.
- Roschelle, J., Feng, M., Murphy, R. F., & Mason, C. A. (2016). Online Mathematics Homework Increases Student Achievement. AERA Open, 2(4), 2332858416673968.
- Rubin, D. B. (1976, December). Trust Inference and Missing Data. Biometrika, 63 (3), pp. 581-592. [70] Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple Imputation of Missing Blood Pressure Covariate in Survival Analysis. Statistics in medicine, 18 (6), 681-694.
- Robert J. Barro and Jong-Wha, (2000), International Data on Educational Attainment: Updates and Implications Lee CID Working Paper No. 42
- Royston, P. (2009). Multiple imputation of missing values: further update of ice, with an emphasis on categorical variables. Stata Journal, 9(3), 466. [68] Royston, P., & White, I. R. (2011). Multiple imputation by chained equations (MICE): implementation in Stata. Journal of Statistical Software, 45(4), 1-20.

- Söderström, Martin, (2006). "School choice and student achievement – new evidence on open-enrolment," Working Paper Series 2006:16, IFAU - Institute for Evaluation of Labour Market and Education Policy.
- Suchaut, B. et Le Bastard, S., (2000), Lecture-écriture au cycle II, Evaluation d'une démarche innovante, Cahiers de l'IREDU, Dijon, N°61.
- Svoboda, R. C., Rozek, C. S., Hyde, J. S., Harackiewicz, J. M., & Destin, M. (2016). Understanding the Relationship Between Parental Education and STEM Course Taking Through IdentityBased and Expectancy-Value Theories of Motivation. *AERA Open*, 2(3), 2332858416664875.
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests.
- Suárez, N., Regueiro, B., Epstein, J. L., Piñeiro, I., Díaz, S. M., & Valle, A. (2016). Homework Involvement and Academic Achievement of Native and Immigrant Students. *Frontiers in psychology*, 7.
- Schiller, K. S., Khmelkov, V. T., & Wang, X.-Q. (2002). Economic development and the effects of family characteristics on mathematics achievement. *Journal of Marriage and Family*, 64, 730-742.
- Sandoval-Hernandez, A., Taniguchi, K., & Aghakasir, P. (2013, Décembre). Is participation in preschool education associated with higher student achievement? (IEA, Éd.) IEA's Policy Brief Series (2), pp. 1-8.
- Schütz, G. (2009). Does the quality of pre-primary education pay off in secondary school? An international comparison using PISA 2003 (No. 68). Ifo Working Paper.
- Sacerdote, B. (2001). Peer Effects with Random Assignment: Results for Dartmouth Roommates. *Quarterly Journal of Economics*, 116 (2), 681-704.
- Schafer, J. L., & Olsen, O. K. (1998). Multiple Imputation for multivariate Missing-Data Problems: A Data Analyst's Perspective. *Multivariate Behavioral Research*, 33 (4), 545-571.
- Spencer, N. H., & Fielding, A. (2000). An instrumental variable consistent estimation procedure to overcome the problem of endogenous variables in multilevel models. *Multilevel Modelling Newsletter*, 12(1), 4-7.
- Sammons, P., Elliot, K., Sylva, K., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (2004). The Impact of Pre-school on Young Children's Cognitive Attainments at Entry to Reception. *British Educational Research Journal*, 30 (5), 691-712.
- Séverine Le Bastard-Landrier, (2005), L'expérience subjective des élèves de seconde : influence sur les résultats scolaires et les vœux d'orientation, *L'orientation scolaire et professionnelle*
- Shultz et al ,(2009) , Assessment of Social Information Processing in Early Childhood: Development and Initial Validation of the Schultz Test of Emotion Processing—Preliminary Version,*Journal of Abnormal Child Psychology*
- Spencer, N. H., & Fielding, A. (2000). An instrumental variable consistent estimation procedure to overcome the problem of endogenous variables in multilevel models. *Multilevel Modelling Newsletter*, 12(1), 4-7.
- Thomas Fuchs & Ludger Woessmann, February (2004) , "Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School," CESifo Working Paper Series 1321, CESifo.
- Topçu, S., & Leana-Taşçılar, M. Z. (2016). The role of motivation and self-esteem in the academic achievement of Turkish gifted students. *Gifted Education International*.
- Tse, S. K., & Xiao, X. Y. (2014). Differential influences of affective factors and contextual factors on high-proficiency readers and low-proficiency readers: a multilevel analysis of PIRLS data from Hong Kong. *Large-scale Assessments in Education*, 2(1), 6.
- Verbeke Geert and Molenberghs Geert , *Linear Mixed Models for Longitudinal Data* , Springer Series in Statistics , (2000)
- Valentine, J. C., Dubois, D. L., & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta Analytic Review. *Educational Psychologist*, 39 (2), 111-133.

- Victor Lavy,(2012),The Friends Factor: How Students' Social Networks Affect Their Academic Achievement and Well-Being?,NBER Working Paper No. w18430
- Vandenberghe V. (2001) Educational Equity: Social interactions may matter, in W. Hutchmacher et al. (eds), In Pursuit of Equity. Dordrecht: Kluwer Academic Publishers.
- Wenceslao González-Manteiga,Extending induced ROC methodology to the functional context , Revue Biostatistics , (2012)
- WANGLI XU and LIXING ZHU, (2009), Kernel-based Generalized Cross-validation in Non-parametric Mixed-effect Models,Scandinavian Journal of Statistics
- W. Xu, L. Zhu, Kernel-based generalized cross-validation in non-parametric mixed-effect models, Scandinavian Journal of Statistics 36 (2009) 229–247.
- William S. Cleveland , (1979) , Robust Locally Weighted Regression and Smoothing Scatterplots,
Journal of the American Statistical Association
- William S. Cleveland & Susan J. Devlin,(1988), Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting, Journal of the American Statistical Association
- Weijing Wang,(2003), Nonparametric estimation of the sojourn time distributions for a multipath model,Journal of the royal Statistical society
- Woessmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. European Economic Review, 50 (3), 695-736.
- Wail Benaabdelaali, (2017), L'effet des inégalités d'éducation sur le développement économique : un essai d'évaluation, HAL
- Xihong Lin and Raymond J. Carroll, Nonparametric Function Estimation for Clustered Data When the Predictor is Measured Without/With Error, (2000)
- Yoav Ganzach,(2010), Parents' education, cognitive ability, educational expectations and educational attainment: Interactive effects,Educational psychology
- Zhang D., Lin X., Raz J., and Sowers M. (1998). Semiparametric stochastic mixed models for longitudinal data, Journal of the American Statistical Association, 93, 710-719.

Table

Index des tableaux

Tableau 1. Répartition des élèves selon le genre.....	31
Tableau 2. Répartition des élèves selon l'âge.....	32
Tableau 3. Répartition des élèves de la 4ème année primaire selon le nombre d'années passées dans le préscolaire.....	32
Tableau 4. Répartition des élèves de la 4ème année primaire selon la langue parlée à la maison.....	33
Tableau 5. Répartition des élèves selon le plus haut niveau d'instruction des parents ou tuteurs.....	33
Tableau 6. Répartition des élèves de la 4ème année primaire selon la situation d'emploi des parents ou tuteurs.....	34
Tableau 7. Répartition des élèves selon leurs conditions de vie familiales.....	34
Tableau 8. Répartition des enseignants selon le genre.....	35
Tableau 9. Répartition des enseignants selon le nombre d'année d'expérience.....	35
Tableau 10. Répartition des enseignants selon le niveau d'instruction.....	36
Tableau 11. Répartition des enseignants selon la participation à la formation continue durant les deux dernières années.....	36
Tableau 12. Milieu de l'établissement.....	37
Tableau 13. Répartition des élèves selon la composition des établissements.....	37
Tableau 14. Répartition des élèves selon les ressources de l'école.....	38
Tableau 15. Répartition des élèves selon les conditions de travail des enseignants.....	38
Tableau 16. Pourcentage des écoles avec des problèmes de discipline et de sécurité.....	39
Tableau 17. Résultats internationaux des élèves de la quatrième année primaire.....	43
Tableau 18. Répartition des élèves marocains selon les niveaux de compétences.....	44
Tableau 19. Scores moyens des élèves de la quatrième année primaire par nombre d'années passées au préscolaire.....	46
Tableau 20. Scores moyens selon les ressources de l'établissement.....	53
Tableau 21. Description préliminaire de la base de données utilisée.....	55
Tableau 22. Liste des variables du premier niveau « Elève ».....	56
Tableau 23. Liste des variables du deuxième niveau « Ecole ».....	56
Tableau 24. Résultats du modèle multiniveaux en lecture avec résolution du biais d'endogénéité.....	66
Tableau 25. Résultats du modèle multiniveaux en matières scientifiques avec résolution du biais d'endogénéité- 2 ^e année collégiale.....	67
Tableau 26. Résultats du modèle multiniveau en matières scientifiques avec résolution du biais d'endogénéité- 2 ^e année collégiale (Suite).....	69
Tableau 27. Catégorie des tailles de la classe.....	93
Tableau 28. La répartition de la taille des classes.....	94
Tableau 29. Le score moyen en lecture selon le retard scolaire et la taille des classes.....	95
Tableau 30. Le score moyen en mathématiques et en sciences selon la taille des classes.....	99
Tableau 31. Le score moyen en mathématiques et en sciences selon le genre des élèves et la taille des classes.....	100
Tableau 32. Le score moyen en mathématiques et en sciences selon le niveau d'instruction des parents et la taille des classes.....	100
Tableau 33. Le score moyen en mathématiques et en sciences selon le milieu de l'établissement et la taille des classes.....	102
Tableau 34. Le score moyen en mathématiques et en sciences selon le degré de mixité sociale.....	103
Tableau 35. Le score moyen en mathématiques et en sciences selon l'âge de l'enseignant.....	104
Tableau 36. Descriptions des variables sélectionnées du programme PIRLS 2011.....	134
Tableau 37. Descriptions des variables sélectionnées du programme TIMSS 2011.....	138
Tableau 38. Les résultats du modèle de la performance scolaire en lecture.....	143
Tableau 39. Les résultats du modèle de la performance en mathématiques.....	145
Tableau 40. Les résultats du modèle de la performance scolaire en sciences.....	147
Tableau 41. Répartition des écoles selon leur composition sociale.....	169
Tableau 42. Score en lecture et composition sociale et économique de l'école.....	170

Tableau 43. Scores en mathématique/science et composition sociale et économique de l'école.....	171
Tableau 44. Les résultats des modèles économétriques en lecture.....	189
Tableau 45. Les résultats des modèles économétriques en mathématique.....	191
Tableau 46. Les résultats des modèles économétriques en science.....	193
Tableau 47. Descriptions des variables du programme PIRLS 2011.....	195
Tableau 48. Descriptions des variables du programme TIMSS 2011.....	198

Index des graphiques

Graphique 1. Evolution des scores en mathématiques.....	40
Graphique 2 . Evolution des scores en sciences.....	40
Graphique 3. Evolution des scores des élèves de la quatrième année primaire en lecture.....	41
Graphique 4. Résultats internationaux des élèves de la deuxième année collégiale.....	42
Graphique 5 . Proportion d'élèves n'ayant pas atteint le niveau inférieur des compétences.....	45
Graphique 6 . Scores moyens par genre.....	46
Graphique 7 . Scores moyens selon le retard scolaire.....	47
Graphique 8 . Scores moyens en quatrième année primaire selon la langue parlée à la maison.....	47
Graphique 9 . Scores moyens en deuxième année collégiale selon la fréquence avec laquelle les élèves parlent la langue arabe à la maison.....	48
Graphique 10 . Scores moyens selon le niveau d'éducation des parents.....	48
Graphique 11 .Scores moyens en quatrième année primaire selon la situation d'emploi des parents ou tuteurs.....	49
Graphique 12 . Scores moyens selon le genre des enseignants.....	49
Graphique 13 . Scores moyens selon le nombre d'années d'expérience des enseignants.....	50
Graphique 14. Scores moyens selon le niveau d'éducation des enseignants.....	51
Graphique 15 .Scores moyens en mathématiques et en sciences selon la participation des enseignants aux sessions de développement professionnel.....	51
Graphique 16 . Scores moyens selon le milieu de l'établissement.....	52
Graphique 17 . Scores moyens selon la composition socioéconomique de l'établissement.....	52
Graphique 18 . Scores moyens selon les conditions de travail des enseignants.....	54
Graphique 19 .scores moyens selon les problèmes de discipline et de sécurité rencontrés dans les établissements scolaires.....	54
Graphique 20. Le score moyen en lecture selon la taille des classes.....	94
Graphique 21. Le score moyen en lecture selon le genre et la taille des classes.....	95
Graphique 22. Le score moyen en lecture selon la préscolarisation et la taille des classes.....	96
Graphique 23. Le score moyen en lecture selon l'expérience des enseignants et la taille des classes..	96
Graphique 24. Le score moyen en lecture selon le genre des enseignants et la taille des classes.....	97
Graphique 25. Le score moyen en lecture selon le degré du problème de la charge horaire et la taille des classes.....	97
Graphique 26. Le score moyen en lecture selon le milieu de l'établissement et la taille des classes....	98
Graphique 27 . Le score moyen en mathématiques et en sciences selon l'expérience des enseignants et la taille des classes.....	101
Graphique 28 .Le score moyen en mathématiques et en sciences selon le genre des enseignants et la taille des classes.....	102
Graphique 29 .Le score moyen en sciences selon la présence d'un laboratoire et la taille des classes	103
Graphique 30 .Les courbes du modèle semi-paramétrique expliquant la performance scolaire en lecture.....	122
Graphique 31 .Les courbes du modèle semi-paramétrique expliquant la performance en mathématiques.....	126
Graphique 32 .Les courbes du modèle semi-paramétrique expliquant le score en science.....	128
Graphique 33. Les courbes simulées du modèle multiniveaux semi-paramétrique conditionnel aux contextes en lecture.....	176
Graphique 34. Les courbes simulées du modèle multiniveaux semi-paramétrique en mathématique	180
Graphique 35. Les courbes du modèle multiniveaux semi-paramétrique en science.....	184

|| Résumé

Résumé

L'objectif de ce travail est d'évaluer la qualité du système éducatif marocain à partir des données du programme TIMSS et PIRLS 2011. Le travail s'articule autour de trois chapitres. Il s'agit d'étudier, dans le premier chapitre, l'influence des caractéristiques individuelles de l'élève et de l'école sur les performances scolaires, ainsi que le rôle important de l'environnement scolaire (effet taille et composition sociale). Dans le deuxième chapitre, nous cherchons à estimer la taille de classe optimale qui assure une réussite généralisée de tous les élèves des deux niveaux à savoir, la 4^e année primaire et la 2^e année collégiale. Le troisième chapitre propose d'étudier la relation existante entre la composition sociale et économique de l'établissement et la performance scolaire, tout en démontrant le rôle de la mixité sociale dans la réussite des élèves. Pour ce faire, nous avons utilisé différentes approches économétriques, à savoir une modélisation multiniveau avec correction du problème de l'endogénéité (chapitre 1), un modèle semi-paramétrique hiérarchique dans le (chapitre 2) et un modèle semi paramétrique hiérarchique contextuel (chapitre 3). Les résultats montrent que la performance scolaire est déterminée par plusieurs facteurs intrinsèques à l'élève et également contextuels. En effet, une taille de classe moins chargée et une école à composition sociale mixte sont les deux éléments essentiels pour un environnement favorable et un apprentissage assuré pour l'ensemble des élèves. Selon nos résultats, les pouvoirs publics devraient accorder la priorité à la réduction de la taille des classes en la limitant à 27 élèves au maximum. De plus, il est nécessaire d'envisager un assouplissement de la carte scolaire afin de favoriser la mixité sociale à l'école. Les résultats obtenus permettent une meilleure compréhension du système scolaire marocain, dans son aspect qualitatif et la justification des politiques éducatives pertinentes pour améliorer la qualité du système éducatif marocain.

Mots clés : Performance scolaire, taille de la classe, mixité sociale, Modèle multiniveaux, modèle de Hausman et Taylor, modèle non paramétrique et semi-paramétrique

Abstract

Abstract

This thesis objective is to examine the quality of the Moroccan education system exploiting the data of the programs TIMSS and PIRLS 2011. The thesis is structured around three chapters. The first chapter examines the influence of individual student and school characteristics on school performance, as well as the important role of the school environment (effect of size and social composition). In the second chapter, we seek to estimate the optimal class size that ensures widespread success for all students at both levels, namely, the fourth year of primary school and the second year of college. The third chapter proposes to study the relationship between the social and economic composition of the school and academic performance, while demonstrating the role of social mix in student success. In order to study this relationship, we mobilize different econometric approaches, by applying a multilevel model with correction for the problem of endogeneity (chapter 1), a hierarchical semi-parametric model (chapter 2) and a contextual hierarchical semi-parametric model (chapter 3). The results show that academic performance is determined by several factors that are intrinsic to the student and also contextual. Indeed, a smaller class size and a school with a mixed social composition are the two essential elements for a favourable environment and assured learning for all students. According to our results, governments should give priority to reducing class size by limiting it to a maximum of 27 students. In addition, it is necessary to consider making the school map more flexible in order to promote social mixing at school. The results obtained allow a better understanding of the Moroccan school system, in its qualitative aspect and the justification of relevant educational policies to improve the quality of the Moroccan education system.

Key words: School performance, class size, social mix, multilevel model, Hausman Taylor model, non-parametric and semiparametric model