

L'évaluation des acquis scolaires au Maroc : nouvelles approches

Introduction

Le Maroc fait partie des pays en développement dont le niveau des acquis des élèves reste relativement faible. En dépit des efforts engagés en vue d'améliorer en partie la qualité des apprentissages, les résultats des enquêtes internationales et nationales révèlent de faibles niveaux des acquis. Dès lors, l'objectif de cet article est de revenir sur les facteurs qui influencent les performances scolaires des élèves. Nous nous intéressons plus spécifiquement aux déterminants microéconomiques de la qualité de l'éducation à travers les performances des élèves.

Les études sur les déterminants des performances scolaires sont riches d'enseignement. Les premières contributions se sont focalisées sur le rôle de l'environnement familial dans l'explication de la réussite des élèves (Coleman, 1966, par exemple). D'autres, plus récemment, ont abordé les facteurs liés à l'établissement scolaire. Pour autant, les contributions récentes mettent en avant l'importance à la fois de l'environnement familial et de l'école (Hanushek, 2003). Des travaux plus récents soulignent également l'influence des pairs sur les performances scolaires. Le présent travail s'inscrit dans cette logique. Son originalité se situe à un double niveau. La première réside dans la mise en évidence de l'ensemble des facteurs explicatifs des performances des élèves et des inégalités scolaires. Malgré l'existence d'une littérature abondante sur le sujet, cette question n'a pas été abordée, à notre connaissance, dans le cas marocain. La seconde cherche à corriger les problèmes d'endogénéité dans les modèles multiniveaux. Enfin, la technique d'imputation adoptée permet de traiter de façon pertinente les valeurs manquantes dans les bases de données.

Cet article est structuré en trois parties. La première aborde la littérature empirique sur les déterminants de la réussite scolaire des élèves. La deuxième partie examine le modèle utilisé et décrit la base de données du Programme national d'évaluation des acquis (PNEA). Elle examine l'approche et la méthodologie utilisées. Enfin, la troisième partie traite des résultats obtenus et nous permet de formuler les principaux enseignements pouvant être tirés de nos résultats en matière de politiques publiques.

**Amina Benbiga
Saïd Hanchane
Nisrine Idir**

Instance nationale
d'évaluation du système
d'éducation et de
formation auprès
du Conseil supérieur de
l'enseignement, Rabat,
Maroc

Les déterminants des performances scolaires : un regard sur la littérature empirique

L'association entre origine sociale et réussite scolaire

Il existe un consensus dans les études sur le lien entre le contexte familial et la réussite scolaire des élèves. La plupart convergent vers un lien étroit entre l'environnement familial de l'élève et ses performances. En effet, la corrélation entre l'origine sociale et la réussite scolaire est une des relations les plus stables et les plus avérées en sciences sociales. Les contributions empiriques ont généralement tendance à estimer l'environnement familial par le statut socio-économique (SSE), mesuré par le niveau de la scolarité et la profession des parents ainsi que le revenu familial (Coleman *et al.*, 1966; Hakkinen *et al.* 2003; Heyneman et Loxley, 1983). Il ressort généralement de ces différentes études que le niveau de la scolarité des parents est le plus significatif. Il représente en effet une source importante de disparités dans les performances des élèves (Chevalier et Lanot, 2002; Fuchs et Wößmann, 2004; Purcel et Dufur, 2001; Schiller *et al.*, 2002; Willms et Somers, 2001; Yayan et Berberoglu, 2004). Les premières contributions considéraient séparément les variables liées au statut socio-économique des parents. Par exemple, l'éducation des parents n'a pas la même influence que le revenu de la famille sur la réussite des élèves. Yuang (2003) pense que le SSE est un concept multidimensionnel et une mesure hiérarchique qui fonctionne à différents niveaux. L'auteur suggère qu'il serait préférable de séparer les différentes composantes du SSE. Il trouve que si le capital culturel et éducatif influence les performances des élèves, la situation économique des parents n'a aucun impact sur leur réussite scolaire. D'autres travaux confirment le rôle de l'éducation des parents sur les performances scolaires des élèves. Fuchs et Wößmann (2004) concluent, à partir des données du Programme international pour le suivi des acquis des élèves (PISA), que les effets de l'éducation des parents sur la réussite des élèves de 15 ans en lecture sont plus significatifs que les impacts de celle-ci sur leur réussite en mathématiques et en sciences.

Notons également que d'autres contributions montrent que les antécédents familiaux exercent un effet important sur le rendement des élèves de 9 et 13 ans (Hanushek et Luque, 2003). Plus précisément, les auteurs trouvent que les étudiants issus de familles défavorisées ou dont le niveau d'éducation des parents est faible ont systématiquement de moins bons résultats. Suivant la même logique, Schiller *et al.* (2002) montrent que les parents instruits semblent plus aptes à fournir à leurs enfants un soutien pédagogique et social important pour la réussite scolaire, comparativement aux parents dont le niveau d'éducation est faible. Ceux dont le niveau d'éducation est plus élevé ont également un meilleur accès à une grande variété de ressources économiques et sociales (par exemple, la structure

familiale, l'environnement familial, l'interaction parent-enfant) pouvant être mises à profit pour aider leurs enfants à réussir à l'école (Wössmann, 2008 ; Schuetz, Ursprung & Wössmann, 2008 ; etc.). Par exemple, Yayan et Berberoglu (2004) remarquent que lorsque le niveau de scolarité des parents et le nombre de livres à la maison sont supérieurs, les performances des élèves de la 2^e année du collège augmentent en mathématiques.

De même, Thompson et Johnston (2006) trouvent que les élèves dont le niveau socio-économique est élevé (nombre de livres à la maison, etc.) sont plus avantagés dans les pays de l'OCDE. D'ailleurs, la littérature sociologique indique que le nombre de livres à la maison représente une indication relativement importante du niveau socioculturel de la famille.

Globalement, il existe un certain consensus sur l'effet significatif du milieu familial sur les performances des élèves. Il n'en demeure pas moins que l'impact du statut professionnel des parents et l'interaction parents-enfants n'est pas clair (Chevalier et Lanot, 2002 ; Ganzach, 2000). Dans certains travaux, l'effet de la famille peut être négatif ou non significatif. Ceci indique que le statut socio-économique des parents n'a pas d'impact sur les apprentissages des élèves et donc sur leur réussite scolaire. Iverson et Walberg (1982) ont montré que la performance académique est plus fortement liée à l'environnement psychosociologique et au degré de stimulation intellectuelle à la maison qu'au statut socio-économique de la famille (cité par Meuret et Morlaix, 2006).

Par ailleurs, plusieurs études montrent que le nombre croissant d'enfants au sein de la famille conduit à des résultats moins favorables pour les enfants. Ce résultat s'explique en partie par la dilution des ressources et du temps accordés à chaque enfant (Teachman *et al.*, 1996). Notons que ces résultats sont conformes aux données empiriques : les enfants de familles plus nombreuses bénéficient d'un environnement moins favorable et des niveaux inférieurs de facilités verbales (Parcel Menaghan, 1994), ainsi que des taux plus élevés de comportement à problèmes et des niveaux inférieurs de réussite scolaire (Downey, 1995).

D'autres travaux montrent que la relation entre l'origine sociale de l'enfant (l'éducation des parents, la structure familiale, etc.) et son rendement scolaire est plus forte dans les pays développés que dans les pays en développement (Coleman *et al.*, 1966 ; Heyneman et Loxley, 1983). En revanche, Simmons et Alexander (1978) ont conclu que les déterminants de la réussite des élèves semblent être essentiellement les mêmes dans les deux groupes de pays (pays industrialisés et pays en développement). De même, les récentes études constatent que les écarts dans les niveaux nationaux de développement économique n'ont aucun effet sur la relation entre le milieu social des enfants et leur réussite scolaire (Hanushek et Luque, 2003).

L'individu face à la réussite scolaire

Plusieurs contributions montrent que les caractéristiques individuelles des élèves, comme leur bien-être, la perception de l'environnement scolaire, la motivation, l'implication dans les activités scolaires, le genre, le travail, etc. ont tous des effets importants sur les performances scolaires. Par exemple, Konu et Rimpela (2002) définissent le bien-être à l'école comme un phénomène à quatre dimensions : les conditions de l'école, les relations sociales, les moyens d'épanouissement personnel et l'état de santé. Pour d'autres, il dépend de nombreux facteurs, y compris leurs opinions sur les règles de l'école et leurs relations avec les enseignants et les camarades de classe (Veenstra et Kuyper, 2004). Le bien-être des élèves peut également affecter d'autres caractéristiques comme le rendement, la motivation ou l'attitude à l'égard des apprentissages (Veenstra et Kuyper, 2004). D'ailleurs, certains auteurs montrent que ce phénomène a un impact non négligeable sur le comportement de l'élève et ses résultats (Hoy et Hannum, 1997).

Parallèlement, les activités scolaires et l'effort individuel sont également importants pour le rendement scolaire. Keith *et al.* (1986), par exemple, indiquent que les notes de l'élève s'améliorent lorsqu'il consacre plus de temps à faire ses devoirs. Dans une étude portant sur 23 pays, Postlethwaite et Wiley (1992) trouvent que le rendement en sciences est en moyenne élevé dans les pays où les élèves déclarent passer beaucoup de temps à faire leurs devoirs. Notons que le temps consacré aux devoirs est également lié à la motivation de l'élève qui souhaite s'approprier le sentiment de réussite (Steinberg *et al.*, 1992).

Par ailleurs, les différences de genre influencent également les résultats scolaires des élèves entre les niveaux et au sein des classes. Plusieurs contributions suggèrent que les filles et les garçons développent des stratégies d'apprentissage différentes (Murphy, 2000 ; etc.). Les filles semblent mieux préparées que les garçons, puisqu'elles développent des intérêts en accord avec les activités scolaires (Murphy, 2000). De plus, les différentes approches développées en matière d'apprentissage par les uns et les autres influencent la motivation, l'auto-perception et les aptitudes sociales qui ont un effet sur les performances. Murphy suggère que les comportements des filles et des garçons dépendent de leurs expériences et de leurs attentes respectives et influencent de ce fait leurs compétences relatives. Au sein de la société, les filles sont conditionnées dans un cadre par lequel on suppose qu'elles seraient plus performantes dans les domaines linguistiques. En revanche, les garçons sont censés être meilleurs dans les domaines quantitatifs comme les mathématiques et les sciences.

L'effet établissement : un rôle de plus en plus important

Les premiers travaux considéraient que les facteurs scolaires avaient peu d'impact sur les performances des élèves comparativement aux facteurs extrascolaires, notamment les caractéristiques individuelles de l'élève et son

environnement économique et social. Le rapport de Coleman (1966) a été à l'origine de cette première thèse. Motivant une grande enquête aux États-Unis sur les milieux sociaux et la réussite des élèves, l'auteur a montré que l'école a peu d'impact sur les rendements scolaires à travers l'introduction d'un ensemble d'indicateurs : ratio élève/maître, qualification des enseignants, type d'équipements, etc. Ce rapport a donné lieu par la suite à une littérature abondante qui cherchait à infirmer ou confirmer l'hypothèse selon laquelle l'établissement n'a pas d'influence sur les acquis et la réussite des élèves.

Aujourd'hui, les études sur la relation entre les ressources de l'école et les performances des élèves n'arrivent pas à un consensus. Certains travaux suggèrent que les ressources supplémentaires ne se traduisent pas forcément par un gain de performance pour les élèves (Hanushek, 1997 ; Hanushek et Luque, 2003, etc.). En revanche, d'autres contributions relèvent un impact positif des caractéristiques de l'école sur le rendement des élèves (Card et Kruger, 1996). Parcel et Dufur (2001) montrent que l'environnement physique au sein de l'école augmente les résultats en mathématiques. En somme, la relation entre l'école et les performances des élèves n'est pas clairement définie. Hanushek arrive à la conclusion qu'il n'existe pas de relations entre les ressources économiques et les résultats scolaires, à partir d'une méta-analyse qui englobe plusieurs contributions empiriques. De même, des études menées au niveau européen ont montré que les politiques et les ressources axées sur les élèves défavorisés n'ont ni amélioré le rendement des élèves, ni atténué les inégalités individuelles (Leuven et Oosterbeek, 2007). D'autres contributions considèrent qu'en plus des variables de l'école, d'autres variables de contrôle, comme l'influence de l'environnement externe à l'école (l'environnement familial) et les antécédents de l'élève lui-même, doivent être pris en compte. Lorsque ces variables ne sont pas prises en considération, les résultats obtenus peuvent être biaisés (Greenwald, Hedges et Laine, 1996). Des études menées dans les pays à plus faible revenu montrent que les ressources matérielles et humaines jouent un rôle important dans l'amélioration du rendement des élèves (Fuller et Clarke, 1994).

Outre les ressources de l'école, d'autres chercheurs s'intéressent plus spécifiquement à la problématique de la taille des classes. Ces travaux arrivent à des résultats souvent contrastés. Dans certaines contributions, l'augmentation de la taille des classes a un impact négatif sur les performances des élèves. On considère en effet qu'une classe de petite taille favorise les apprentissages et donc la réussite scolaire. Toutefois, la question de la taille des classes demeure problématique pour les décideurs politiques et les chercheurs. Si la taille apporte des résultats appréciables en matière de réussite scolaire, les décideurs seraient incités à agir dans ce sens, alors même que c'est une politique très coûteuse. Pour autant, les écrits sont très contrastés à ce sujet. Dans un travail réalisé dans trois de ses articles, Hanushek (1997, 2003 et 2006) montre, à partir d'une méta-analyse réunissant plusieurs articles sur

l'effet de la taille des classes, l'absence de toute relation constante et étroite entre la taille des classes et les performances des élèves.

L'étude plus récente d'Altinok *et al.* (2009) nuance néanmoins les conclusions d'Hanushek. Les auteurs utilisent une modélisation rigoureuse et une technique récente tenant compte des problèmes d'endogénéité liés à la taille des classes. De nombreux mécanismes peuvent conduire à une détermination conjointe et simultanée de la taille des classes et de la réussite des élèves, rendant la première endogène à la seconde. Par exemple, l'établissement peut réduire la taille des classes pour les élèves en difficulté. De leur côté, les décideurs peuvent accorder des mécanismes de financement compensatoire pour des établissements dont la proportion des élèves en difficulté est élevée. Pour corriger ce problème d'endogénéité, les auteurs effectuent une régression en différence en introduisant les effets fixes sur un échantillon de 33 pays, les menant à ce résultat : l'effet de la taille des classes n'est pas substantiel. L'impact est statistiquement significatif sur 16 des 33 pays. Il est négatif seulement pour 10 pays. Dès lors, les auteurs concluent que la réduction de la taille des classes n'est pas une stratégie rentable pour élever le niveau de rendement des élèves dans les pays industrialisés et les pays en développement.

Partant de la même logique, Wößmann et West (2006) estiment l'effet de la taille des classes sur les performances des élèves dans 11 pays, à partir des données TIMSS. Pour ce faire, les auteurs utilisent la méthode des variables instrumentales et corrigent ainsi le biais d'endogénéité. Ils trouvent un effet taille de classe variable selon les pays et qui dépend du système scolaire. Ils concluent également que des classes plus petites ont un effet observable sur les rendements des élèves seulement dans les pays où la capacité moyenne du corps enseignant est faible. L'effet est plus important en présence d'enseignants de qualité.

L'impact du préscolaire, du redoublement et des classes homogènes sur les rendements scolaires des élèves

Dans cette partie, nous nous focaliserons sur les études qui se rattachent à certains aspects de la vie de l'élève et qui sont directement liés aux politiques pédagogiques menées en matière d'éducation. Nous traiterons respectivement le préscolaire et le redoublement.

En ce qui concerne le préscolaire, le Maroc, comme d'autres pays en développement, cherche à se doter véritablement de moyens lui permettant de mettre en place les classes maternelles. Cet intérêt part de l'idée selon laquelle le préscolaire procure un avantage pour la suite de la scolarité, à la fois sur le plan des acquisitions et de la carrière scolaire. Il réduit la probabilité de redoubler. L'influence du préscolaire est d'autant plus positive que la scolarisation dans ce dernier est plus longue.

Cette intuition est derrière de nombreuses recherches menées sur ce sujet. Les travaux réalisés dans les années 90 révèlent un impact positif sur les acquisitions scolaires des élèves en mathématiques et en français. L'influence du préscolaire est perceptible également sur le redoublement. A titre d'exemple, l'étude de Caille (2001) montre que les enfants inscrits dans le préscolaire à l'âge de deux ans ont moins de risque de redoubler comparativement à des élèves scolarisés plus tardivement. Ces derniers présentent des capacités moindres que les autres enfants. Toutefois, l'écart enregistré au départ a tendance à se réduire avec le niveau de scolarisation. L'impact du préscolaire est plus perceptible dans les premières années qu'en fin du primaire ou encore au collège (Caille & Rosenwald, 2006).

Comme nous l'avons rappelé précédemment, le préscolaire réduit la probabilité de redoubler. Ce dernier a une influence mitigée sur la réussite scolaire des élèves. Dans les pays pratiquant le redoublement, il est perçu comme une solution pour remédier aux difficultés scolaires. Un élève en difficulté sera maintenu dans le même niveau pour mieux assimiler l'enseignement dispensé et combler le retard accumulé par rapport aux autres. Selon les tenants de cette thèse, le redoublement accorde aux élèves les moins performants l'occasion de suivre un enseignement à leur propre cadence. Toutefois, les chercheurs pédagogues et économistes de l'éducation considèrent que le redoublement est néfaste pour la réussite scolaire des élèves. Les recherches menées dans ces deux domaines conduisent à la même conclusion : le redoublement est préjudiciable pour l'élève puisqu'il impacte négativement ses progrès cognitifs, sa motivation envers l'école et son orientation.

Suivant les premières approches, le redoublement a une influence négative sur les performances des élèves. A partir d'une méta-analyse menée sur un échantillon de 40 articles de recherche, sélectionnés parmi 850 selon des critères scientifiques rigoureux, Holmes et Matthews (1984) montrent que les élèves qui redoublent progressent significativement moins que les élèves qui ne redoublent pas, tout en contrôlant les caractéristiques des élèves. Dans un papier plus récent, Holmes (1990) arrive aux mêmes conclusions. L'auteur réalise également une méta-analyse à partir de 63 articles. Il conclut que le redoublement est inefficace. L'originalité du travail d'Holmes tient au fait qu'il fait la distinction entre les études qui font des comparaisons entre âge constant et niveau constant, autrement dit, il considère l'effet à moyen et long termes du redoublement. Il trouve un écart important entre les redoublants et les promus en termes de performances. De plus, le développement social de l'élève et son comportement à l'école sont influencés négativement. Pour les études menées à âge constant, l'auteur observe un accroissement de l'écart de performance entre les redoublants et les promus, alors qu'à niveau constant, ils obtiennent des résultats équivalents.

De son côté, Seibel (1984) étudie l'évolution des scores d'un échantillon de 11 000 élèves à des épreuves standardisées de français et de mathématiques

entre juin et décembre 1983. Il trouve les mêmes résultats que les études précédentes. Grisay (1993) arrive également aux mêmes conclusions. L'auteur a fait passer des tests à un échantillon d'élèves de la 6^e année du primaire aux épreuves de français. De même, Manacorda (2007) trouve que le redoublement a un impact négatif sur le rendement des élèves, focalisant son étude en Uruguay. Ce pays applique en effet une loi particulière sur le redoublement. Au-delà de 25 jours d'absence, l'élève redouble sa classe. En revanche, il ne redouble pas si le nombre d'absence ne dépasse pas 24. Partant de là, l'auteur compare les élèves qui accumulent 26 jours d'absence à ceux qui ont comptabilisé 24 jours d'absence. Il conclut que les résultats des redoublants sont impactés négativement.

En revanche, Grenne et Winter (2004) trouvent que la politique de redoublement a un impact positif sur la réussite scolaire des élèves en Floride. Les auteurs évaluent la politique publique qui consiste à mettre en place le redoublement pour améliorer les capacités en lecture des élèves. Ils utilisent des données individuelles se rapportant au niveau des élèves en lecture. Ils examinent leurs performances un an après le redoublement effectif comparativement aux élèves faibles promus pour lesquels la politique de redoublement n'a pas été mise en place. Les auteurs concluent que les progrès réalisés en lecture par les élèves faibles redoublants sont largement supérieurs aux élèves faibles promus.

Globalement, les études sur le redoublement n'arrivent pas à un consensus. Deux courants existent : le premier prône les bienfaits du redoublement ou lui trouve un effet positif sur les performances scolaires, tandis que le second souligne son impact négatif sur la réussite scolaire des élèves.

L'influence des groupes de pairs sur les performances des élèves

Les premières approches donnaient une définition simpliste des inégalités scolaires. Elles étaient considérées comme l'impact du statut économique et social d'un individu sur ses performances scolaires. Toutefois, la réalité est beaucoup plus complexe. A titre d'exemple, les stratégies résidentielles des parents, compte tenu de leur niveau social, conduisent à une concentration des enfants issus de familles favorisées au sein d'une même école. De même, les familles de milieu défavorisé ont plus de chance de placer leurs enfants dans des établissements situés dans des zones défavorisées avec des enfants issus du même milieu. Dès lors, c'est la composition (les groupes de pairs) de la population de l'établissement qui impacte davantage la réussite des élèves. Ce n'est pas tant le statut social et économique de l'individu qui explique les inégalités dans les performances scolaires, mais ce sont les caractéristiques de la population qui compose le public de l'établissement, autrement dit les caractéristiques du groupe.

Il était déjà question de cet effet de groupes de pairs dans les travaux de Coleman en 1966. L'hypothèse sous-jacente à l'étude de Coleman était que

les inégalités des résultats dans les établissements étaient conditionnées par les écarts de ressources et de financement entre ces derniers (Dumay et Dupriez, 2004). Toutefois, l'auteur est arrivé à des conclusions différentes. Ce sont les caractéristiques familiales et les modes de composition de la population de l'établissement qui expliquent les inégalités dans les performances scolaires. À partir de là, une série de contributions scientifiques ont orienté les recherches vers d'autres directions. Les premiers cherchent à mettre en évidence le rôle de l'établissement en tant qu'entité organisationnelle et pédagogique dans l'explication des performances scolaires des élèves. Les seconds, en revanche, soulignent davantage l'influence des groupes de pairs. Pour ces derniers, «la capacité de l'établissement à faire acquérir des connaissances ne repose pas tant sur des processus organisationnels et pédagogiques internes, mais plutôt sur les caractéristiques agrégées de la population de l'établissement». Les structures organisationnelles et pédagogiques d'un établissement ne sont pas indépendantes du contexte dans lequel elles sont mises en place. Aussi, il est important de tenir compte à la fois des caractéristiques propres à l'établissement et de la composition de sa population.

Aujourd'hui, cette littérature reçoit une attention croissante de la part des chercheurs et des acteurs publics dans les pays industrialisés et en développement. Ces derniers s'intéressent de plus en plus aux inégalités scolaires. L'intérêt porté à l'influence des groupes de pairs concerne à la fois l'économie de l'éducation et la sociologie. Les travaux menés à ce sujet prennent en compte différentes formes d'effets de pairs : les effets de composition liés aux capacités individuelles, les effets de pairs sociaux, économiques, liés au genre ou aux motivations et l'intérêt porté aux apprentissages, etc.

Quel que soit l'effet considéré et en dépit de la richesse des travaux, les études sur les effets de pairs n'arrivent pas à un consensus. Cette absence d'accord est liée aux variations des résultats obtenus. Certaines contributions révèlent que les effets de pairs ont une influence sur l'explication des performances scolaires des élèves (Ammermueller et Pischke, 2009). Par ailleurs, les effets des pairs peuvent être de sources multiples. Par exemple, Figlio (2005) souligne les effets du comportement des pairs sur les résultats des élèves. Partant des données individuelles d'un grand district en Floride, l'auteur estime l'impact des pairs perturbateur sur le comportement de chaque élève et les résultats aux tests. Il constate que le comportement perturbateur des pairs est associé à une probabilité accrue de suspension d'un élève et une baisse des rendements des élèves. D'autres contributions trouvent que l'âge moyen des pairs peut avoir un effet négatif sur l'élève. Ces études considèrent que les élèves plus âgés fournissent moins d'efforts. Dès lors, lorsque la proportion des élèves plus âgés est élevée en classe, l'élève aura tendance à suivre le groupe et à fournir moins d'efforts. Ce comportement pourrait se traduire par une baisse des performances scolaires des élèves.

Cette mise en revue des études sur le sujet n'est pas exhaustive. Elle donne simplement un aperçu de l'influence de la composition de la population de l'établissement sur le rendement des élèves. Elle permet aussi de situer notre travail. Dans ce qui suit, nous exposerons les résultats des estimations sur les effets des groupes de pairs.

Spécification du modèle et présentation des données

Dans cette section, nous examinerons sous un angle différent les déterminants de la qualité interne de l'enseignement au Maroc. Celle-ci peut être appréhendée à travers l'évaluation des acquis des élèves qui varient en fonction du contexte dans lequel ils évoluent. Ils peuvent être influencés en partie par le milieu social, par l'établissement qu'ils fréquentent ou encore par les groupes de pairs. Avant de présenter les résultats de nos différentes estimations, nous reviendrons sur l'approche méthodologique et les données utilisées.

Une nouvelle approche de l'évaluation de la qualité des apprentissages : la résolution des problèmes d'endogénéité

Les acquis des élèves diffèrent selon la classe, l'établissement, la région ou encore le pays. Ces différences sont liées au contexte dans lequel l'élève évolue. Pour étudier l'influence de chaque contexte, nous estimons une fonction de production de l'éducation qui explique les performances scolaires des élèves par un ensemble de variables contextuelles qui portent sur plusieurs dimensions (l'élève, la classe et l'établissement) et qui ne peuvent être appréhendées par une simple régression linéaire. Celle-ci ne permet pas de prendre en considération l'information contenue à la fois dans des données microéconomiques et des données macroéconomiques. Ils ne permettent donc pas de donner des informations en même temps au sujet de l'individu et du groupe social ou de l'institution qui l'accueille ou l'influence.

Ce type de structure requiert une modélisation appropriée à laquelle ne peuvent répondre les modèles linéaires. Les modèles multiniveaux sont plus adaptés à ce type de données. Ils présentent plusieurs avantages. D'abord, ils permettent de combiner des variables de niveaux différents. En effet, ils traitent une information emboîtée sur plusieurs niveaux d'observations. L'exemple le plus classique est imputé à la structure même du système éducatif : les élèves appartiennent à des classes, qui elles-mêmes appartiennent à des écoles, qui sont à leur tour situées dans des villes ou des régions. De plus, ces modèles donnent une meilleure estimation de la variance.

Partant de là, nous utilisons le modèle multiniveaux pour estimer notre fonction de production de l'éducation dont l'*input* est le score moyen de l'élève et les *outputs* sont regroupés en catégories : les caractéristiques de

l'individu, celles de la famille, de l'établissement et le niveau de stratification. Ainsi le modèle s'écrit :

$$Y_{ij} = c + \beta_1 X_{ij} + \gamma_1 \bar{X}_{-j} + \gamma_2 K_j + V_j + \varepsilon_{ij}$$

avec :

Y_{ij} : la variable dépendante, elle représente dans notre cas les performances de l'élève i dans l'établissement j ;

c : un terme constant qui représente la moyenne générale des élèves;

X_{ij} : le vecteur des variables explicatives, il regroupe les caractéristiques de l'élève i dans l'établissement j ;

\bar{X}_{-j} : le vecteur des effets de pairs, ce sont les moyennes par école des caractéristiques des élèves;

K_j : le vecteur des variables caractérisant l'école j ;

V_j : un terme aléatoire qui constitue l'écart de l'école j par rapport à la moyenne générale c , il peut être considéré comme un terme qui contient les caractéristiques non observées de l'école, il est supposé de moyenne 0 et de variance τ^2 ;

ε_{ij} : le terme d'erreur qui suit une loi normale avec une moyenne de 0 et une variance constante σ^2 ($\varepsilon_{ij} \sim N(0, \sigma^2)$).

Ce modèle repose sur plusieurs hypothèses parmi lesquelles figurent l'hypothèse stipulant la non-corrélation entre les variables de niveau 1 et le terme d'erreur de niveau 2, c'est-à-dire $\text{cov}(X_{ij}, V_j) = 0$. Autrement dit, il ne doit pas y avoir une corrélation entre les caractéristiques observables des élèves et les caractéristiques non observables de l'établissement auquel ils appartiennent. La violation de cette hypothèse entraîne un problème d'endogénéité que l'on appelle endogénéité de niveau 2.

Pour résoudre le biais causé par cette endogénéité, plusieurs méthodes économétriques ont été développées, notamment sur des données de panel : la méthode des variables instrumentales et la méthode Mundlak. Nous optons pour la seconde dans la mesure où la première est fondée sur le choix d'instruments qui ne sont pas supposés être corrélés avec les variables indépendantes et non corrélés avec les résidus. Ce choix est généralement aléatoire, et il est relativement difficile de trouver de bons instruments.

Dès lors, nous adaptons l'approche Mundlak (annexe 1), utilisée initialement pour les estimations des données de panel, au modèle multiniveaux dont la structure est très proche. Selon l'auteur, le problème d'endogénéité peut être corrigé aisément par l'introduction, dans l'équation à estimer, des moyennes des variables individuelles pour chaque établissement. Ces moyennes peuvent représenter, en présence de données hiérarchiques sur l'éducation, différentes formes d'effets de pairs dans l'établissement (Hanchane et Mostafa, 2011).

Avant de présenter les résultats de nos estimations, nous exposerons dans ce qui suit les données utilisées et le traitement des données.

Les données utilisées

Dans ce travail, nous exploiterons les données issues du Programme national d'évaluation des acquis (PNEA). Lancé par le Conseil supérieur de l'enseignement (CSE) en 2008, il se veut un référentiel national d'évaluation des acquis, compétences et savoirs fondamentaux des élèves. Il est basé sur des tests disciplinaires (français, arabe, mathématiques, sciences et physique-chimie) et des questionnaires contextuels (questionnaires élève, parent, enseignant et établissement). Dès lors, il couvre un large éventail de variables. Les tests menés en 2008 ont porté sur une cohorte d'élèves des 4^e et 6^e années du primaire et des 2^e et 3^e années du collège. Les données proviennent d'un échantillon représentatif au niveau national : 230 établissements scolaires du primaire et 6 900 élèves par niveau. Elles couvrent également 212 collèges et 6 360 élèves par niveau.

Comme nous l'avons souligné précédemment, nous cherchons à estimer une fonction de production de l'éducation dont l'*output* est représenté par les résultats aux tests de compétences des élèves. La variable dépendante de notre étude est le score de compétences en mathématiques, en sciences et en physique.

Par ailleurs, les *inputs* de la fonction de production de l'éducation peuvent être regroupés en quatre catégories. La première regroupe des variables qui portent sur les caractéristiques intrinsèques de l'élève, ses antécédents et son environnement socio-économique. Ces données sont tirées des questionnaires élève et parent. La seconde réunit les variables de l'école. Celles-ci concernent l'environnement au sein de l'établissement scolaire et l'environnement pédagogique. Les questionnaires enseignant et établissement du PNEA réunissent des informations détaillées sur l'enseignant, la pédagogie en classe, la vie scolaire, etc. Dans le cas du Maroc, celles-ci peuvent être enrichies à partir de la base de données nationale et longitudinale sur les établissements scolaires (BLE). A cet effet, on a procédé à un appariement des données PNEA pour enrichir l'analyse.

On peut ainsi estimer (tableau 1) :

- les caractéristiques de l'élève : les principales variables retenues pour notre modèle empirique concernent le genre, la perception et le travail salarié. La première variable est une indicatrice qui prend la valeur 1 lorsqu'il s'agit d'un garçon et 0 ailleurs. En effet, plusieurs études montrent que les performances peuvent changer selon le sexe.

La seconde variable mesure la « passion » pour une matière donnée. Elle suppose que les écarts de performance entre les élèves peuvent être associés en partie aux penchants de l'élève pour telle ou telle matière. En effet, le fait de préférer une matière à une autre indique que l'enfant trouve plus d'aisance dans l'apprentissage de celle-ci et sera d'autant plus motivé pour l'évaluer. La motivation aura, à terme, un impact positif sur les performances de l'élève puisqu'il sera enclin à fournir un effort supplémentaire.

Enfin, le modèle introduit une variable mesurant l'influence d'une activité « salariale » sur les apprentissages et donc les performances scolaires des élèves. Cette variable est tirée du questionnaire parent. Outre ces variables, nous retenons également les antécédents des élèves matérialisés par le redoublement et le retard scolaire. Le redoublement est une politique pédagogique qui concerne les élèves en difficulté, en vue de favoriser leur rattrapage et leur maintien dans le système scolaire. En revanche, la politique de passage automatique se base sur l'idée selon laquelle le redoublement est source de démotivation de l'élève et qu'au contraire des politiques d'appui seraient plus favorables.

Nous pouvons également citer un autre exemple de politiques publiques pédagogiques qui concerne directement l'âge légal d'accès à l'école. L'entrée des élèves au primaire devrait se faire à partir d'un âge prédéfini. Or, le système éducatif marocain autorise l'inscription des enfants à un âge avancé et l'entrée tardive des élèves. Il en résulte un phénomène de retard scolaire lié, non pas au redoublement, mais à l'application sur le terrain de la politique concernant l'âge légal d'accès à l'école.

Tableau 1
Variables explicatives du modèle empirique

	Variables	Définition	PNEA
Caractéristiques de l'élève	Genre (garçon/fille)	Indicatrice égale à 1 lorsque l'élève est un garçon et 0 ailleurs	Questionnaire élève
	Préscolaire	Le nombre d'années passées dans le préscolaire	Questionnaire parent
	Travail de l'élève	Est-ce que l'élève travaille en contrepartie d'un salaire? Indicatrice égale à 1 lorsque l'élève travaille et 0 sinon	Questionnaire parent
	Redoublement	Indicatrice égale à 1 si l'élève a redoublé au moins une fois et 0 s'il n'a jamais redoublé	Questionnaire élève
	Retard scolaire de l'élève	Indicatrice égale à 1 si l'élève est en retard par rapport à l'âge légal d'accès à l'école et 0 sinon	Questionnaire élève
	Matière de préférence	Indicatrice égale à 1 lorsque l'élève a une préférence pour une matière spécifique et 0 ailleurs	Questionnaire élève
Caractéristiques de la famille	Conditions matérielles et culturelles	Indice composite (chambre individuelle, télévision, ordinateur, internet, bibliothèque, vélo, jeux éducatifs)	Questionnaire parent
	Éducation des parents	Nombre moyen d'années d'éducation des parents	Questionnaire parent
	Taille de la famille	Nombre des membres de la famille	Questionnaire élève

	Variables	Définition	PNEA
Caractéristiques de l'école	Problèmes de civisme (directeur)	Indice composite des problèmes de civisme au sein de l'école	Questionnaire école
	Absentéisme	Indicatrice égale à 1 en présence de problèmes d'absentéisme au sein de l'établissement	—
	Groupe favorisé	Indicatrice égale à 1 lorsque l'établissement regroupe une population d'élèves issus de milieux favorisés	—
	Milieu	Indicatrice égale à 1 pour le milieu urbain et 0 pour le milieu rural	Base BLE
	Infrastructure	Indice composite de l'infrastructure éducative disponible au sein de l'établissement	Base BLE
	Ratio élève/maitre	Ce ratio indique le nombre moyen d'élèves par enseignant dans l'établissement	Base BLE
	Taille des classes	La taille moyenne des classes est définie par des intervalles	—
	Retard scolaire	Proportion des élèves en retard scolaire dans l'établissement (retard d'accès)	—
	Redoublement	Proportion des élèves qui ont redoublé dans l'établissement	

Ces deux exemples présentent un intérêt dans l'examen des déterminants des performances scolaires des élèves, puisqu'elles représentent des politiques pédagogiques dont l'application effective est « anarchique ». Ils sont introduits dans le modèle à partir des questionnaires élèves et sont représentés par des variables indicatrices. Celles-ci prennent la valeur 1 lorsque l'élève accède tardivement à l'école par rapport à l'âge légal de scolarisation ou lorsqu'il a redoublé au moins une fois et 0 ailleurs.

- Les caractéristiques de la famille: les variables les plus couramment utilisées sont le niveau d'éducation et la profession des parents. Dans certains travaux, ces variables sont regroupées en un seul indicateur: le statut socio-économique des parents. Dans le présent travail, seul le niveau moyen d'éducation des parents est pris en considération.

Parallèlement, nous introduisons la taille de la famille. Nous supposons, en effet, qu'une famille plus importante mobilise les ressources et que l'allocation de celles-ci entre les enfants est moins efficiente. Par exemple, la répartition des ressources entre les membres d'une famille nombreuse implique moins de ressources consacrées à chaque enfant. En revanche, une famille avec un ou deux enfants seulement peut représenter un avantage pour ces derniers puisqu'ils vont bénéficier d'un taux plus important de ressources et de temps. Les données concernant cette variable sont tirées du questionnaire élève. Elles ne concernent pas la famille au sens strict du

terme, c'est-à-dire les parents et les enfants, mais intègrent les personnes qui vivent sous le même toit que l'élève, c'est-à-dire la famille au sens large.

Enfin, la dernière variable retenue dans le modèle concerne les conditions matérielles et culturelles. Cette variable est un indice composite regroupant l'addition des réponses obtenues à partir d'un bloc de questions sur la possession d'un ensemble d'éléments au sein du foyer familial. Préalablement à ce calcul, les réponses ont été recodées pour disposer d'une information cohérente.

- L'environnement scolaire de l'élève: récemment, des contributions empiriques ont souligné l'importance de l'environnement scolaire sur la réussite scolaire des élèves. Ainsi, il est souvent admis que la taille des classes a un impact sur les apprentissages et donc sur les performances des élèves. De même, la sécurité dans l'école et l'environnement général ont une influence sur les apprentissages des élèves.

Dans notre modèle, nous introduisons les variables utilisées couramment dans la littérature: la taille des classes ou le ratio élève/maître et les ressources de l'école. En plus de ces variables, nous introduisons deux variables additionnelles, à savoir les problèmes de civisme au sein de l'établissement et l'efficacité de celui-ci.

Notons que le ratio élève/maître est tiré de la base BLE. Il rapporte le nombre d'élèves de l'établissement au nombre d'enseignants. La variable infrastructure est également tirée de la base BLE. Elle prend la forme d'un indice composite, construit à partir de plusieurs variables qui reflètent la qualité des infrastructures au sein de l'école. Cet indice comprend des variables indicatrices relatives à la présence dans l'établissement d'une cantine ou d'un système de restauration, d'une bibliothèque, d'une salle informatique, d'un internat, d'une connexion internet et de latrines.

Concernant les variables liées aux problèmes de civisme au sein de l'école, nous construisons un indice composite. Celui-ci regroupe un ensemble de variables telles que l'agression des professeurs et celle des élèves par exemple.

Parallèlement, nous introduisons les scores d'efficacité afin d'établir une distinction entre établissements efficaces et établissements non efficaces. Ce score a été obtenu à partir des estimations réalisées par la méthode Data Envelopment Analysis (DEA). Celle-ci utilise le taux de réussite dans l'établissement comme variable à expliquer et le ratio élève/maître, l'ancienneté de l'enseignant, la proportion des élèves qui entrent à l'école à l'âge légal de scolarisation et le ratio élève/classe comme variables explicatives.

- Les effets de pairs ou de composition indiquent que l'élève évolue au sein de l'établissement avec une population qui l'influence et qui est influencée par lui. Nous introduisons les effets de pairs dans notre estimation de la fonction de production des performances scolaires pour mesurer le degré de stratification entre les établissements et les élèves et donc les inégalités scolaires. Ils sont introduits dans le modèle par les moyennes des caractéristiques individuelles et familiales de l'élève.

Le traitement des données

Nous disposions initialement de plusieurs bases de données. Le PNEA regroupe plusieurs bases (questionnaires) qui doivent être fusionnées. En plus de la réussite scolaire, matérialisée par des données sur les scores, le PNEA recueille des informations contextuelles sur les élèves, les enseignants et l'établissement. Les élèves répondent à des questions sur leur situation familiale et démographique, leur environnement scolaire et leurs activités extrascolaires. Ces informations sont réunies dans le questionnaire élève. Parallèlement, les enseignants en mathématiques et en sciences ont répondu à des questions qui se rattachent à leur formation, aux pratiques pédagogiques, à la taille des classes, etc. (questionnaire enseignant). Les directeurs des établissements ont aussi répondu à des questions relatives aux caractéristiques et ressources de l'établissement, au degré de centralisation du processus de décision, à la répartition des responsabilités au sein de l'école et d'autres sujets comme le degré d'implication des parents (questionnaire établissement). La fusion de ces informations a donné lieu à la construction de 14 bases de données, construites par matière et par niveau. En outre, nous avons procédé à un appariement des bases PNEA et BLE pour exploiter des données additionnelles concernant les établissements scolaires.

Les données issues de l'enquête sur l'évaluation des acquis, bien qu'elles soient riches en informations, présentent quelques difficultés liées à la présence de valeurs manquantes qui peuvent biaiser les estimations. Pour remédier à ce problème, plusieurs solutions sont proposées par certaines études. L'une des stratégies consiste à supprimer simplement les variables manquantes, ce qui revient à réduire la taille de l'échantillon. Cette méthode ne pose pas de problèmes majeurs lorsque les variables manquantes sont peu nombreuses (moins de 5 %). En revanche, en présence d'un nombre important de variables manquantes, le risque serait de perdre de l'information utile puisque les résultats obtenus peuvent être non représentatifs de la population sondée. De plus, le fait de supprimer une partie importante des données peut conduire à des résultats biaisés si le sondage qui en résulte n'est pas distribué de manière aléatoire sur la population totale de l'étude.

La seconde stratégie consiste à compléter les données manquantes pour faire une analyse à partir d'une base complète. De nombreuses méthodes ont été développées. La plus simple consiste à remplacer les données manquantes par la moyenne calculée sur les données réellement observées. Dans ce cas, on obtient une constante pour toutes les données manquantes. En dépit de sa simplicité, cette méthode présente quelques limites, notamment la sous-estimation de la variance. Une deuxième méthode simple consiste à imputer les données manquantes par des valeurs provenant d'un individu similaire pour lequel toute l'information est observée.

D'autres méthodes plus complexes ont été utilisées, notamment pour réduire les problèmes de la sous-estimation de la variance. L'une de ces

approches consiste à remplacer les valeurs manquantes par des valeurs prédites selon un modèle de régression basé sur l'imputation simple. Même si elle réduit les problèmes de biais, cette méthode d'imputation présente quelques limites. En effet, elle ne prend pas en considération toute l'incertitude liée à la variable manquante, traitant les données manquantes comme parfaitement observables.

Enfin, des techniques plus récentes ont été employées pour pallier aux problèmes des autres méthodes, notamment en matière de prise en compte de l'incertitude. Celle de l'imputation multiple introduite par Rubin (1987) est l'une des plus usitées dans les contributions récentes. Elle correspond davantage à la nature de nos données. La méthode d'imputation multiple peut être décrite en trois étapes. La première consiste à créer $m \geq 2$ valeurs plausibles pour les données manquantes. Chacun de ces ensembles sert à remplir les données manquantes et créer ainsi m bases de données. Ensuite, les bases complètes sont analysées avec les méthodes utilisées traditionnellement (différentes méthodes économétriques d'estimation). Enfin, les résultats obtenus des analyses réalisées à partir des m bases complètes sont combinées selon une procédure spécifique (rapport INE, 2010) dans le but d'obtenir des estimateurs non biaisés.

Une fois les données imputées, nous estimons notre fonction de production de l'éducation à l'aide d'un modèle multiniveaux et de la méthode Mundlak (1978), cette dernière permettant de corriger le biais d'endogénéité.

Les résultats du modèle

Le rôle de la famille dans la réussite scolaire des élèves : une influence significative quels que soient le niveau et la matière

Les résultats des estimations concernant le contexte familial sont présentés dans le tableau 2. Il regroupe les résultats de deux spécifications : modèle sans l'introduction des effets de groupes des pairs (moyennes des variables individuelles) et modèle avec effets des groupes de pairs.

En ce qui concerne l'éducation des parents, il ressort des différentes régressions que celle-ci influence positivement et significativement les performances scolaires des élèves et ce, quels que soient la matière et le niveau enseignés. Ce résultat est en adéquation avec la littérature empirique sur le sujet. Des parents dont le niveau d'éducation est élevé disposent d'une grande variété de moyens sociaux et culturels pour accompagner l'apprentissage de leurs enfants. Toutefois, l'ampleur de l'effet diffère selon les matières et les niveaux. Notons également que l'ampleur de l'effet de l'éducation des parents sur la réussite scolaire des élèves diffère selon que l'on se situe au primaire ou au collège. Il est plus important au primaire, suggérant que l'élève a davantage besoin de ses parents pour son apprentissage scolaire (analyse des résultats des données du PNEA). Notons également que l'effet

éducation des parents reste faible comparativement à d'autres variables de contrôle qui lui sont liées. Les conditions matérielles et culturelles au sein du ménage impactent plus les performances scolaires des élèves.

Outre les conditions matérielles et culturelles, la taille de la famille influence négativement les rendements des élèves. En effet, le mode de gestion de la famille diffère selon sa taille. L'allocation des ressources par enfant est moins importante lorsque le nombre des membres de la famille augmente. Une famille avec un ou deux enfant(s) a tendance à concentrer ses ressources matérielles et son temps à son ou ses deux enfant(s). En revanche, une famille nombreuse n'est pas en mesure d'accorder suffisamment de ressources ni de temps à chaque enfant, notamment en ce qui concerne les activités scolaires et parascolaires. Dès lors, l'allocation des ressources et du temps est différente selon la taille de la famille.

S'agissant des mathématiques, des sciences et de la physique-chimie, les mêmes conclusions s'appliquent pour l'impact du contexte familial sur les performances scolaires des élèves. L'éducation des parents joue un rôle positif et significatif dans l'explication des rendements scolaires des élèves au primaire et au collège. L'effet est plus important au primaire qu'au collège. En ce qui concerne le coefficient associé à la taille de la famille, il est négatif et significatif pour les performances scolaires des élèves au primaire en mathématiques et en sciences. De même, les coefficients se rattachant aux conditions matérielles et culturelles au sein du ménage ont le signe attendu. Ils impactent positivement et significativement la réussite des élèves au primaire. En revanche, l'influence de la taille de la famille et des conditions matérielles et culturelles sur les performances des élèves au collège n'est pas significative. Ce résultat confirme l'hypothèse, avancée précédemment, selon laquelle les apprentissages des élèves sont davantage impactés par l'éducation des parents et la composition de la population de l'établissement.

Tableau 2

Influence des caractéristiques familiales sur les performances des élèves en mathématiques et en sciences

Variables	Mathématiques		Sciences		Physique-chimie	
			4 ^e année du primaire			
	Spécification 1	Spécification 2	Spécification 1	Spécification 2	Spécification 1	Spécification 2
	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
	(t-statistique)	(t-statistique)	(t-statistique)	(t-statistique)	(t-statistique)	(t-statistique)
Constante	40,03"	50,22	46,57"	51,52"	—	—
	(2,31)	(0,24)	(16,15)	(11,83)		
Taille de la famille	-0,17"	-0,17"	-0,20"	-0,20"	—	—
	(4,45)	(4,45)	(12,47)	(12,47)		
Éducation des parents	0,43"	0,43"	0,35"	0,35"	—	—
	(13,62)	(13,62)	(45,03)	(45,03)		
Conditions matérielles et culturelles	2,81"	2,81"	2,20"	2,20"	—	—
	(2,76)	(2,76)	(2,45)	(2,45)		
2 ^e année du collège						
Constante	10,61"	8,76"	12,72"	13,32"	17,96"	18,52"
	(5,27)	(3,31)	(6,29)	(4,83)	(6,64)	(5,83)
Taille de la famille	ns	ns	ns	ns	ns	ns
Education des parents	0,15"	0,15"	0,13'	0,13'	0,25"	0,25"
	(2,22)	(2,22)	(2,05)	(2,05)	(3,52)	(3,52)
Conditions matérielles et culturelles	ns	3,04"	ns	ns	ns	1,681* (1,91)

Note : Spécification 1 regroupe les variables liées aux caractéristiques individuelles, de la famille et de l'école. Spécification 2 intègre, en plus des variables citées précédemment, les effets de pairs. *, **, *** indiquent un seuil de significativité respectivement de 1, 5 et 10%. ns : non significativement différent de zéro.

Les résultats des deux spécifications révèlent que l'éducation des parents a toujours un impact positif et significatif sur les performances des élèves en mathématiques et en sciences. Le coefficient de cette variable est relativement stable en ce qui concerne les mathématiques. Toutefois, il tend à baisser pour les sciences, suggérant que l'éducation des parents influence peu la réussite des élèves dans cette matière. Ce résultat peut s'expliquer en partie par le manque de connaissances des parents en sciences. Notons toutefois que l'influence de l'éducation des parents reste limitée comparativement à l'impact de la taille de la famille.

Le rôle des caractéristiques individuelles dans l'explication des performances scolaires

Le tableau 4 présente les coefficients des variables individuelles estimés. La première variable a trait à la motivation de l'élève mesurée par la matière de préférence. L'ensemble de nos différentes estimations aboutit à un consensus. La matière de préférence joue un rôle positif dans la réussite scolaire des élèves en 4^e année du primaire et en 2^e année du collège.

Les résultats au collège confirment l'assiduité des filles comparativement aux garçons. Les filles réalisent de meilleures performances que les garçons pour l'ensemble des matières, à l'exception des mathématiques, pour lesquelles la variable n'est pas significative.

En ce qui concerne l'exercice d'une activité salariée par l'élève, les résultats du tableau 3 suggèrent que cette variable n'a pas d'influence sur la réussite scolaire des élèves marocains quels que soient le niveau et la matière.

La variable, qui mesure le nombre d'années que l'élève passe dans le préscolaire, est positive et significative pour les élèves de la 4^e année du primaire et les spécifications 1 et 2 qui résolvent le biais d'endogénéité souligné précédemment. En effet, le préscolaire joue un rôle important puisqu'il prépare les enfants à l'apprentissage scolaire. Notons également que la durée de fréquentation du préscolaire est corrélée avec les apprentissages : plus la durée du préscolaire est longue et plus les bénéfices que peuvent en tirer les élèves sont meilleurs. Ce résultat est conforme aux intuitions de Campbell et Ramey (1994), qui considèrent que plus l'intervention du préscolaire dans la vie de l'enfant dure et débute tôt, plus les bénéfices cognitifs sont importants. Suivant la même logique, Reynolds (1995) suggère que les enfants qui participent à un programme préscolaire pendant deux ans présentent un niveau de préparation scolaire supérieur à ceux qui le font pendant un an.

Tableau 3

Impact des caractéristiques individuelles sur les performances des élèves

Variables	Mathématiques		Sciences		Physique-chimie	
4 ^e année du primaire						
	Spéc. 1	Spéc. 2	Spéc. 1	Spéc. 2	Spéc.1	Spéc. 2
	Coefficient (t-statistique)	là.	là.	là.	là.	là.
Genre	ns	ns	ns	ns	—	—
Matière de préférence	ns	ns	1,23*** (2,51)	1,23*** (2,51)	—	—
Travail	ns	ns	ns	ns	—	—
2 ^e année du collège						
Genre	ns	ns	−2,86*** (6,18)	−2,86*** (6,18)	−2,27*** (3,78)	−2,27*** (3,78)
Matière de préférence	4,23*** (7,69)	4,23*** (7,69)	2,29*** (4,76)	2,29*** (4,76)	3,69*** (6,19)	3,69*** (6,19)
Travail	ns	ns	ns	ns	ns	ns

Note: Spécification 1 regroupe les variables liées aux caractéristiques individuelles, de la famille et de l'école. Spécification 2 intègre, en plus des variables citées précédemment, les effets de paires. *, **, *** indiquent un seuil de significativité respectivement de 1, 5 et 10%. ns: non significativement différent de zéro.

Toutefois, l'ampleur de l'impact change selon les matières. L'effet du préscolaire semble plus élevé en mathématiques. Ce résultat indique en particulier que le nombre d'années passées dans le préscolaire est utile pour les apprentissages des mathématiques. Horton (1996) souligne que les expériences préscolaires affectent les apprentissages en mathématiques. Les compétences en cette matière sont acquises très tôt, même en dehors de l'école, et sont renforcées par le nombre d'années passées dans le préscolaire. En revanche, l'impact du préscolaire sur les acquis des élèves en sciences est plus faible comparativement aux mathématiques. Globalement, nos résultats suggèrent qu'il est important d'investir dans le préscolaire, de façon appropriée, pour promouvoir l'adaptation sociale et scolaire des élèves.

La deuxième variable concerne le redoublement. Les résultats des estimations renvoient à des effets significatifs du redoublement sur la réussite scolaire des élèves. La nature de l'effet est de signe attendu pour l'ensemble des estimations, des niveaux et des matières. Comme indiqué dans le tableau 4, le redoublement dans le primaire impacte négativement les performances scolaires des élèves en sciences. Toutefois, il est non significatif en ce qui concerne les mathématiques.

Nos résultats suggèrent que le redoublement est inefficace au primaire. Les élèves qui redoublent perdent confiance en leurs capacités et sont dès

lors moins motivés. Même s'il est conçu comme un moyen de remédier aux difficultés de l'élève, le redoublement n'améliore pas les performances de ce dernier pour autant. Plusieurs études montrent d'ailleurs que les élèves qui obtiennent les meilleurs résultats sont issus de pays qui ne pratiquent pas ou peu le redoublement (résultats des enquêtes internationales de l'évaluation des acquis).

Ces conclusions s'appliquent également à la 2^e année du collège. Les résultats du tableau 4 indiquent que le redoublement influence négativement la réussite scolaire des collégiens. L'ampleur diffère selon les matières comme dans le primaire. Le redoublement a un impact négatif sur la réussite des élèves en français, en arabe et en physique-chimie, comparativement aux autres matières.

Globalement, l'effet du redoublement est négatif au primaire et au collège, indiquant qu'une année supplémentaire dans le même niveau n'apporte pas les bénéfices escomptés en termes d'acquis scolaires. Ce résultat est conforme à ceux obtenus dans les différentes contributions de la pédagogie et de l'économie de l'éducation sur le redoublement. Toutefois, la mise en pratique du passage automatique ne produit pas toujours les effets attendus. Certains pays, comme la Belgique, la France et le Québec, ont mis en place des politiques de lutte contre le redoublement qui ont eu des effets mitigés (Draeloants, 2006). Parallèlement, les pays qui pratiquent des politiques de passage automatique depuis longtemps obtiennent de meilleurs résultats concernant les acquis des élèves. Ceci suggère que toute politique de passage automatique (dans les pays qui pratiquent traditionnellement le redoublement) devrait tenir compte des caractéristiques propres à chaque système éducatif et de l'ancrage de la vision positive du redoublement autour de l'environnement de l'élève. Notons également que cette politique devrait être accompagnée d'un dispositif d'accompagnement pour les élèves en difficultés.

La troisième variable individuelle, et qui donne une indication sur les choix de politiques publiques, est liée à la problématique de l'homogénéité dans l'accès à l'éducation. Elle porte sur l'entrée tardive ou non dans le système scolaire, par rapport à l'âge légal.

Les résultats des régressions reproduits dans le tableau 4 montrent que le retard scolaire influence les performances scolaires des élèves. La nature et l'ampleur de cet effet au primaire sont variables selon les niveaux. L'entrée tardive des élèves à l'école influence négativement leurs performances scolaires en sciences et ce, quelle que soit la méthode d'estimation utilisée. L'effet est similaire au collège, quoique d'une ampleur plus importante.

Notons également que la corrélation entre l'âge d'entrée à l'école et la réussite scolaire n'est pas significative en ce qui concerne les mathématiques et la physique-chimie. Globalement, les résultats relatifs au retard scolaire indiquent que les élèves qui accèdent à l'école à un âge tardif, c'est-à-dire au-delà de huit ans, sont relativement moins performants que les élèves inscrits précocement ou à l'âge convenu à l'école.

Tableau 4
Influence des antécédents de l'élève sur ses performances scolaires

Variables	Mathématiques		Sciences		Physique-chimie	
4 ^e année du primaire						
	Spéc. 1	Spéc. 2	Spéc. 1	Spéc. 2	Spéc. 1	Spéc. 2
	Coefficient (t-statistique)	là.	là.	là.	là.	là.
Redoublement	ns	ns	− 3,52*** (5,32)	− 3,52*** (5,32)	—	—
Retard scolaire	ns	ns	− 0,92** (2,06)	− 0,92** (2,05)	—	—
Préscolaire	1,09** (2,68)	1,09** (2,68)	0,82*** (4,40)	0,82*** (4,40)	—	—
2 ^e année du collège						
Redoublement	− 1,80** (2,58)	− 1,80** (2,58)	− 2,49*** (3,91)	− 2,49*** (3,91)	− 4,20*** (5,21)	− 4,20*** (5,21)
Retard scolaire	ns	ns	− 1,19** (2,63)	− 1,19** (2,63)	ns	ns
Préscolaire	—	—	—	—	—	—

Note: Spécification 1 regroupe les variables liées aux caractéristiques individuelles, de la famille et de l'école. Spécification 2 intègre, en plus des variables citées précédemment, les effets de paires. *, **, *** indiquent un seuil de significativité respectivement de 1, 5 et 10 %. ns: non significativement différent de zéro.

L'école et la réussite scolaire des élèves

Les contributions récentes montrent que l'école joue également un rôle important dans la réussite scolaire des élèves grâce à l'environnement et au climat qu'elle leur offre ainsi qu'au cadre administratif dont elle dispose. Les établissements scolaires, même s'ils s'inscrivent dans un cadre administratif national, développent un environnement spécifique lié au cadre social dans lequel ils évoluent et définissent l'identité de l'établissement. Aussi intéressons-nous à la fois aux aspects « administratifs » et au climat qui caractérisent les établissements. Dès lors, quatre groupes de variables sont introduits dans les deux spécifications estimées : le climat au sein de l'école, l'environnement, les ressources de l'école et l'encadrement dans l'établissement.

Le climat au sein de l'école est estimé à partir des problèmes de civisme. Nous avons opté pour ce type de variables au lieu de la perception de l'élève, des parents ou encore de l'enseignant pour construire un indice de climat à l'école. Ce choix est à notre sens plus cohérent et moins subjectif, puisqu'il se base sur des faits perceptibles au sein de l'établissement. De ce fait, un établissement qui regroupe des comportements de délinquance scolaire ne peut pas fonctionner de manière efficace et offrir aux élèves un

cadre d'apprentissage favorable. Dès lors, il produit des effets négatifs sur les performances des élèves.

Les problèmes de civisme augmentent le sentiment d'insécurité au sein de l'établissement et peuvent se répercuter de manière négative sur les apprentissages et donc sur la réussite scolaire. Ainsi, il ressort de notre modèle que les problèmes de civisme ont une influence négative sur la réussite scolaire des élèves du primaire et du collège. L'ampleur et la nature de l'effet diffèrent toutefois selon les matières. Pour ce qui est du primaire, nos résultats suggèrent qu'un climat défavorable impacte négativement les mathématiques et les sciences. Ceci indique clairement qu'un climat défavorable est le signe d'une structure sociale fragile ne réunissant pas les conditions de coopération entre les différents acteurs (élèves, enseignants, administration et parents). Au collège, les problèmes de civisme ne produisent pas d'effets positifs sur les performances des élèves.

Tableau 5
Impact des variables de l'école sur les performances des élèves

Variables	Mathématiques		Sciences		Physique-chimie	
4 ^e année du primaire						
	Spéc. 1	Spéc. 2	Spéc. 1	Spéc. 2	Spéc. 1	Spéc. 2
	Coefficient (t-statistique)	Id.	Id.	Id.	Id.	Id.
Problèmes de civisme	−0,42*** (7,48)	−0,34*** (9,38)	−0,39*** (6,43)	−0,37*** (4,79)		
Ratio élèves/maître	−0,26*** (108,92)	−0,21*** (70,96)	−0,26*** (180,25)	−0,22*** (153,94)		
Taille des classes	—	—		—		
Infrastructure	0,69*** (0,52)	ns	0,22*** (5,31)	0,31*** (7,15)		
Milieu	5,33*** (11,21)	3,60*** (6,69)	3,50*** (17,82)	3,18*** (11,64)		
2 ^e année du collège						
Problèmes de civisme	−0,25*** (4,68)	−0,31*** (4,56)	−0,29*** (5,64)	−0,27*** (6,06)	−0,38*** (6,00)	−0,27*** (3,12)
Ratio élèves/maître	0,10*** (6,15)	0,09*** (5,44)	0,04** (2,53)	0,04** (2,52)	0,04** (2,10)	0,05*** (2,22)
Infrastructure	0,40** (2,58)	ns	0,73*** (4,68)	−0,70*** (3,09)	1,41*** (7,05)	1,35*** (6,41)
Milieu	3,45*** (12,91)	2,09*** (5,24)	2,52*** (9,12)	1,63*** (3,98)	3,09*** (9,18)	1,66*** (4,00)
Score d'efficience	9,18*** (4,12)	10,74*** (5,06)	10,58*** (4,61)	13,91*** (5,35)	13,58*** (4,22)	16,22*** (4,54)

Note: Spécification 1 regroupe les variables liées aux caractéristiques individuelles, de la famille et de l'école. Spécification 2 intègre, en plus des variables citées précédemment, les effets de pairs. *, **, *** indiquent un seuil de significativité respectivement de 1, 5 et 10%. ns: non significativement différent de zéro.

Le second groupe réunit les variables d'environnement comme le milieu (rural ou urbain). Cette variable concerne les établissements situés dans le milieu urbain. Ces derniers obtiennent de meilleurs résultats que ceux du milieu rural, au primaire et au collège. En effet, ils évoluent dans un contexte économique et social favorable. Le troisième groupe a trait aux ressources de l'école matérialisées par l'infrastructure disponible au sein de l'établissement. Le coefficient lié à cette variable est positif et significatif. Au primaire, l'impact des ressources matérielles disponibles au sein de l'établissement sur la réussite des élèves n'est pas toujours significatif selon les spécifications, et il est différent au collège. Il est significatif et de signe contre-intuitif pour les sciences au collège. En revanche, il est significatif et positif pour la physique-chimie. Ces résultats pourraient être liés aux effets captés par le milieu et l'établissement d'efficience dans l'hypothèse où l'on considère que l'efficience implique la concentration de ressources importantes. Toutefois, la matrice de corrélation des variables explicatives du modèle ne révèle pas de relation de corrélation entre les variables. De plus, l'efficience des établissements est construite à partir d'indicateurs qui ne sont pas liés aux ressources matérielles. Ils sont davantage rattachés aux aspects qualitatifs.

Enfin, le dernier groupe porte sur deux types de variables liées à l'encadrement : le ratio élèves/maître et la taille des classes. L'introduction du premier dans le modèle a été possible grâce à l'appariement des bases PNEA et BLE. Les résultats des estimations indiquent que le ratio moyen élèves/maître de l'établissement a une influence négative et significative sur la réussite des élèves au primaire. Toutefois, la portée de cet indicateur est limitée comparativement aux autres variables de l'établissement : le coefficient γ est très faible. Au primaire, le coefficient lié au taux d'encadrement moyen de l'établissement est faible. Ceci suggère que les besoins en encadrement restent importants.

Ces coefficients sont encore plus faibles au collège et pour l'ensemble des matières. Certes, le ratio élèves/maître est significatif, indiquant le rôle important de l'encadrement de l'enseignant dans la réussite scolaire des élèves ; il n'en demeure pas moins qu'il a un signe positif et contre-intuitif au collège.

Inégalités des performances scolaires : les effets des groupes de pairs

Les précédentes sous-sections ont mis en évidence le rôle combiné des facteurs individuels, socio-économiques et liés à l'école dans la réussite scolaire des élèves. Nous prolongerons l'analyse en intégrant les effets de pairs, c'est-à-dire les effets liés aux caractéristiques agrégées de la population de l'établissement, pour appréhender les inégalités scolaires.

La littérature sur les effets de pairs est abondante et arrive à des résultats contrastés. Dans ce qui suit, nous analyserons les résultats de nos propres

estimations concernant le Maroc pour la 4^e année du primaire et la 2^e année du collège. Nous distinguerons différentes formes d'effets de composition : ceux liés au sexe, à la taille de la famille, à la motivation, au redoublement, au retard d'accès, à l'éducation des parents, au travail et au confort.

Le tableau ci-dessous résume les résultats obtenus. En ce qui concerne les effets de pairs liés au genre, nos résultats suggèrent globalement que la nature et l'ampleur des effets de groupe de pairs diffèrent en 4^e année du primaire. Pour ce qui est des mathématiques, les différences de sexe n'ont aucune influence sur les performances des élèves. Il est donc évident que, dans ce cas, les effets de composition n'ont aucun impact dans l'explication des inégalités scolaires entre les sexes. Au collège, une classe composée principalement de garçons a un impact sur les rendements des élèves en mathématiques et en sciences.

Tableau 6

Impact des variables liées aux effets de pairs sur les performances scolaires

	Mathématiques	Sciences	Physique-chimie
	Coefficient (t-statistique)	Coefficient (t-statistique)	Coefficient (t-statistique)
Moyenne genre	ns	ns	—
Moyenne taille de la famille	ns	-0,27*** (3,05)	—
Moyenne matière de préférence	ns	-1,23*** (522,82)	—
Moyenne redoublement	ns	ns	—
Moyenne retard	ns	-6,18*** (2,84)	—
Moyenne travail	ns	ns	—
Moyenne éducation des parents	ns	-0,33*** (9,82)	—
Moyenne conditions matérielles et culturelles	ns	ns	—
2 ^e année du collège			
Moyenne genre (garçon)	3,06*** (4,37)	5,05*** (6,33)	ns
Moyenne taille de la famille	ns	ns	-0,48*** (2,28)
Moyenne matière de préférence	ns	ns	5,79*** (3,54)
Moyenne redoublement	ns	-4,66*** (4,47)	ns
Moyenne retard	ns	ns	-2,53** (2,11)
Moyenne travail	ns	-6,76*** (5,22)	-5,58** (2,87)
Moyenne éducation des parents	0,35*** (3,48)	ns	ns
Moyenne conditions matérielles et culturelles	3,04*** (5,13)	2,50*** (3,59)	ns

Note: *, **, *** indiquent un seuil de significativité respectivement de 1, 5 et 10 %. ns: non significativement différent de zéro.

Les effets de groupe liés au redoublement et au retard d'accès sont de même nature. Ils ont un impact négatif sur les rendements scolaires des élèves. En moyenne, une classe avec une proportion élevée de redoublants impacte négativement les performances scolaires des élèves en sciences au collège. Les effets de composition liés au redoublement ne sont pas significatifs au primaire, suggérant que la proportion des élèves redoublants n'y est pas élevée. En revanche, le retard d'accès à l'école a un impact négatif et significatif sur les performances des élèves en sciences au primaire. Des classes hétérogènes en termes d'âge ont une influence négative et significative sur les performances des élèves du primaire en sciences.

Au collège, les effets de pairs liés au redoublement ont un impact négatif et significatif sur les performances des élèves en sciences. Ce résultat suggère que les redoublants dans une classe peuvent influencer les rendements des autres élèves en perturbant l'organisation pédagogique de l'établissement. L'hétérogénéité des individus au sein de la classe requiert de la part de l'enseignant un effort d'adaptation des modes d'enseignement. De plus, les redoublants sont dans une situation de manque de motivation et de rejet de l'école. Ils peuvent, dès lors, perturber le déroulement des cours en classe. Les sciences paraissent les plus touchées par ces effets de composition. En revanche, les mathématiques sont moins impactées par les effets de pairs liés au redoublement.

Parallèlement aux effets de compositions cités précédemment, notre modèle intègre également des effets de pairs liés à la composition sociale et économique de l'individu. Trois catégories de variables complémentaires sont intégrées dans le modèle. La première concerne l'éducation des parents. Au primaire, l'effet groupe de pairs n'est pas significatif pour les mathématiques. Il est en revanche significatif et négatif pour les sciences. Ce résultat laisse suggérer que le niveau d'éducation des parents ne leur permet pas d'opérer des choix stratégiques quant à l'établissement dans lequel ils inscrivent leurs enfants. Des parents dont le niveau d'éducation est élevé peuvent inscrire leurs enfants dans des écoles proches de leur lieu de résidence, celui-ci n'étant pas forcément situé dans des zones favorisées.

Au collège, les résultats sont quelque peu différents. Les effets de compositions ne sont pas significatifs dans le cas des sciences et de la physique-chimie. Toutefois, une classe regroupant des élèves dont le niveau d'éducation des parents est élevé impacte positivement les performances scolaires des élèves en mathématiques. Cet impact est à relier avec une variable de contrôle complémentaire qui concerne les conditions matérielles et culturelles de la famille.

La seconde variable est liée aux conditions matérielles et culturelles. Comme nous l'avons précisé précédemment, cette variable est un indice composite intégrant les éléments indispensables à l'épanouissement des enfants. Cet indice représente une proxy du niveau de vie des parents. Au primaire, nos estimations révèlent que le confort au sein du foyer familial n'a

pas d'impact sur les apprentissages et donc sur la réussite scolaire des élèves. Par ailleurs, nos résultats confirment l'hypothèse avancée précédemment selon laquelle les collégiens sont davantage influencés par le groupe de pairs dans lequel ils évoluent que par l'environnement familial.

Les effets de pairs liés aux conditions matérielles et culturelles impactent positivement et significativement les performances des élèves au collège en mathématiques et en sciences.

Ce résultat rejoint ceux des contributions empiriques qui intègrent la moyenne du statut socio-économique des parents, traduisant le degré de stratification et d'inégalités sociales. L'éducation des parents et les conditions matérielles et culturelles déterminent dans quel établissement l'enfant est inscrit et avec quel groupe d'élèves il évolue. Il en résulte des interactions sociales qui peuvent avoir une influence positive sur la réussite scolaire de l'élève. En revanche, l'effet de composition lié à cette variable est absent au primaire.

La troisième variable concerne les effets des groupes de pairs liés à la taille de la famille. Cette variable a, comme attendu, un signe négatif et significatif pour les sciences au primaire et la physique-chimie au collège. Les performances scolaires des élèves sont corrélées négativement avec la moyenne de la taille de la famille. L'effet des groupes de pairs taille de la famille donne une indication indirecte sur le statut socio-économique des parents. Généralement, les ménages défavorisés ont un nombre important d'enfants. Cette variable est complémentaire des autres variables liées au statut de la famille. Les écoles réunissant un groupe important d'élèves issus de familles favorisées sont plus efficaces dans la mesure où les parents opèrent des choix stratégiques en matière de scolarisation de leurs enfants. De plus, ils ont tendance à être plus exigeants vis-à-vis de l'établissement qui scolarise leurs enfants (organisation pédagogique, gestion, etc.).

Conclusion

Ce travail a pour objectif d'analyser les facteurs explicatifs de la qualité des apprentissages des élèves marocains. Il exploite les données issues de l'enquête nationale sur l'évaluation des acquis des élèves marocains du PNEA. Les niveaux concernés sont la 4^e année du primaire et la 2^e année du collège et les disciplines scientifiques. Il se base sur une méthodologie rigoureuse qui prend en considération la variété des facteurs explicatifs de la réussite scolaire des élèves. Outre le traitement préalable des données, les estimations ont été réalisées de sorte à corriger les biais liés aux problèmes d'endogénéité soulevés dans les modèles multiniveaux.

Les résultats des différentes estimations révèlent que les performances des élèves ne sont pas uniquement liées aux caractéristiques familiales et individuelles. S'agissant de l'environnement familial, ce sont les conditions matérielles et culturelles qui contribuent davantage à la réussite scolaire des

élèves comparativement à l'éducation des parents. En ce qui concerne les antécédents des élèves, les résultats montrent, par exemple, que le nombre d'années passées dans le préscolaire augmente les chances de réussite des élèves dans le primaire. Ces derniers sont en effet mieux préparés et présentent moins de risque de redoublement. De même, le retard scolaire en termes d'accès et le redoublement ont une influence négative sur les rendements scolaires des élèves. Ceci est d'autant plus vrai que lorsque la proportion des élèves en retard scolaire ou qui redoublent est élevée dans une classe, elle influence négativement les performances des autres élèves. En effet, il ressort de ces résultats que les classes homogènes produiraient de meilleurs résultats sur les acquis des élèves comparativement aux classes hétérogènes en termes de retard (âge d'accès) et de redoublement.

D'autres facteurs concourent également à l'explication des performances scolaires des élèves, en particulier l'environnement scolaire, l'établissement et les effets de pairs. Par exemple, la taille des classes, les problèmes de civisme et d'absentéisme influencent significativement les performances des élèves. Notons aussi que les effets de composition ont une influence significative sur les performances des élèves marocains. Cette influence est plus perceptible au collège qu'au primaire. Par ailleurs, l'influence des groupes de pairs est plus marquée pour ce qui est du redoublement. En effet, une classe qui regroupe une proportion élevée de redoublants ou encore qui réunit des élèves hétérogènes en termes d'âge aura une influence négative sur les rendements scolaires des élèves.

Globalement, plusieurs composantes participent à l'explication des performances des élèves marocains. Outre les conditions matérielles et la taille de la famille, les antécédents des élèves jouent un rôle déterminant sur leur réussite scolaire. Il en est de même de l'environnement scolaire dans lequel ils évoluent. Pour autant, il ne faut pas négliger les facteurs pédagogiques autres que l'encadrement. Ces derniers peuvent contribuer de façon significative à l'explication des rendements scolaires des élèves. Toutefois, la prise en compte de ce type de composante requiert des analyses plus qualitatives.

Annexe 1

Le modèle multi-niveaux non corrélé

L'article fondateur de Mundlak (1978) a donné naissance à une série de travaux qui ont réussi à résoudre les problèmes d'endogénéité spécifiques aux données de panel. Mundlak a plus précisément évoqué la notion de « différence imaginaire » entre le modèle à effets fixes et le modèle à effets aléatoires, dès lors que l'on s'écarte de l'hypothèse d'exogénéité faible indispensable pour une estimation convergente et efficace de ce dernier.

C'est toute la génération des estimateurs IV en panel et ceux plus récents, GMM, qui adopte et élargit l'approche de Mundlak.

Les problèmes d'endogénéité évoqués dans l'encadré 2 peuvent être résolus par une adaptation de l'estimateur de Mundlak aux données multi-niveaux.

Si on considère le modèle suivant :

$$Y_{it} = \beta X_{it} + \alpha_i + \varepsilon_{it}$$

avec $i = 1, 2, \dots, N$ et $t = 1, 2, \dots, T$

l'indice i indique l'individu i et l'indice t indique l'instant t

$$\varepsilon_{it} \sim N(0, \delta^2)$$

Le terme α_i peut être traité de deux manières selon que l'on est dans un modèle à effets fixes ou dans un modèle à effets aléatoires. Dans le premier cas, le problème d'endogénéité ne se pose pas puisque α_i est traité comme une constante que l'on doit estimer pour chaque individu. Dans l'autre cas de figure, α_i est traité comme une variable aléatoire, ce qui peut entraîner un problème d'endogénéité. Cette endogénéité est manifestée par la violation de l'hypothèse suivante :

$$\text{cov}(X_{it}, \alpha_i) = 0$$

L'estimateur des α_i pour le modèle à effets fixes est donné par :

$$\hat{\beta} = W_{XX}^{-1} W_{XY}$$

L'estimateur des α_i pour le modèle à effets aléatoires est l'estimateur des Moindres Carrés Généralisés qui est donné par :

$$\hat{\beta}_{MCG} = (W_{XX} + \theta \beta_{XX})^{-1} (W_{XY} + \theta \beta_{XY})$$

$$\text{avec } W_{XX} = \sum (X_{it} - \bar{X}_i)^2;$$

$$W_{XY} = \sum (X_{it} - \bar{X}_i)(Y_{it} - \bar{Y}_i);$$

$$\beta_{XX} = \sum (\bar{X}_i - \bar{X})^2;$$

$$\beta_{XY} = \sum (\bar{X}_i - \bar{X})(\bar{Y}_i - \bar{Y});$$

$$\theta = \frac{\sigma^2}{\sigma^2 + \tau \sigma_\alpha^2}.$$

L'utilisation de la transformation de Fuller and Battese (1973) permet d'obtenir, avec un OLS, des estimateurs pour les α_i qui sont similaires à ceux obtenus par les MCG. Cette transformation consiste à régresser $Y_{it} - \lambda \bar{Y}_i$ sur $X_{it} - \lambda \bar{X}_i$ avec $\lambda = 1 - \sqrt{\theta}$.

Lorsque $\theta = 1$ le modèle est identique à un modèle à effets fixes.

Le modèle à effets aléatoires est plus approprié que le modèle à effets fixes parce que selon Maddala (1987) :

1. quand on est en présence d'un grand nombre d'observations, le modèle à effets aléatoires permet d'estimer seulement la moyenne et la variance des α_i au lieu d'estimer N valeurs de α_i dans le cas du modèle à effets fixes ;
2. si on veut inférer les résultats obtenus à partir d'un échantillon sur la population à partir de laquelle on a tiré cet échantillon, il est nécessaire de traiter α_i comme étant un terme aléatoire et non pas fixe ;
3. lorsqu'on a des variables constantes dans le temps, Z_i , le modèle deviendra donc :

$$Y_{it} = \beta X_{it} + \gamma Z_i + \alpha_i + \varepsilon_{it}$$

Dans de telles situations, le modèle à effets fixes ne permet pas d'estimer les α_i ainsi le modèle à effets aléatoires est mieux adapté dans ce cas.

Mundlak (1978) montre que le modèle à effets fixes et le modèle à effets aléatoires donnent les mêmes estimations si on suppose que les α_i dépendent des moyennes des X_{it} :

$$\alpha_i = \pi \bar{X}_i + w_i$$

avec cette hypothèse le modèle devient alors :

$$Y_{it} = \pi \bar{X}_i + \beta X_{it} + \gamma Z_i + w_i + \varepsilon_{it} \quad (*)$$

Avec $\text{cov}(w_i, X_{it}) = 0$

La transformation de Fuller and Battese permet d'obtenir les estimations de α_i par un OLS appliqué à l'équation suivante :

$$Y_{it} - \lambda \bar{Y}_i = \pi(\bar{X}_i - \lambda \bar{X}_i) + \beta(X_{it} - \lambda \bar{X}_i) + \gamma(Z_i - \lambda Z_i) + v_{it}$$

où $v_{it} = w_i + \varepsilon_{it}$

En ajoutant et retranchant $\beta \bar{X}_i$ l'équation devient :

$$Y_{it} - \lambda \bar{Y}_i = \pi \bar{X}_i - \lambda \pi \bar{X}_i + \beta X_{it} - \lambda \beta \bar{X}_i + \beta \bar{X}_i - \beta \bar{X}_i + \gamma(Z_i - \lambda Z_i) + v_{it}$$

On a donc :

$$Y_{it} - \lambda \bar{Y}_i = (\pi + \beta) \bar{X}_i - \lambda(\pi + \beta) \bar{X}_i + \beta(X_{it} - \bar{X}_i) + \gamma(1 - \lambda)Z_i + v_{it}$$

$$Y_{it} - \lambda \bar{Y}_i = (1 - \lambda)(\pi + \beta) \bar{X}_i + \beta(X_{it} - \bar{X}_i) + \gamma(1 - \lambda)Z_i + v_{it}$$

On note $\delta = (1 - \lambda)(\pi + \beta)$ et $\lambda = 1 - \sqrt{\theta}$ avec $\theta = \frac{\sigma^2}{\sigma^2 + \tau\sigma_\alpha^2}$

On peut estimer les coefficients α_i , β et γ indépendamment par un OLS car $(X_{it} - \bar{X}_i)$ et \bar{X}_i sont indépendants.

L'estimateur des β est l'estimateur intra-individuel : $\hat{\beta} = W_{xx}^{-1}W_{xy}$.

L'estimateur des δ est donné par $\hat{\delta} = (\sum \bar{X}_i \bar{X}_i)^{-1} (\sum \bar{Y}_i \bar{X}_i) (1 - \lambda)$

L'estimateur des π n'est que la différence entre les estimateurs inter- et intra-individuels des β :

$$\hat{\pi} = (\sum \bar{X}_i \bar{X}_i)^{-1} (\sum \bar{Y}_i \bar{X}_i) - \hat{\beta}$$

Les $\hat{\gamma}$ sont obtenus par la régression des variables constantes dans le temps Z_i sur la moyenne des Y_{it} dans le temps, c'est-à-dire sur les \bar{Y}_i .

La méthode de Mundlak qui apporte une solution au problème d'endogénéité des effets individuels pour le cas des données de panel peut être adaptée aux données multi-niveaux comme on va le montrer.

En effet, si avec les données de panel on a deux niveaux d'observation, l'un est formé par les individus et l'autre par les instants d'observation ; dans les données PNEA et TIMSS on a aussi deux niveaux, le premier est constitué par les élèves et le deuxième par les écoles.

L'endogénéité à laquelle nous nous intéressons est l'endogénéité de niveau 2, elle est causée par la corrélation entre les caractéristiques non observées de l'école (niveau 2) et les caractéristiques observées des élèves.

Si on note les caractéristiques non observées de l'école V_j , le modèle s'écrit alors :

$$Y_{ij} = \beta X_{ij} + V_j + \varepsilon_{ij}$$

L'endogénéité dans ce cas est présentée par la violation de l'hypothèse : $\text{cov}(X_{ij}, V_j) = 0$.

Comme pour les données de panel, l'approche de Mundlak apporte une solution à ce problème en introduisant les moyennes des variables explicatives \bar{X}_j . Ces moyennes constituent, pour notre cas, les effets de pairs. L'approche de Mundlak permet en outre d'estimer les variables qui sont propres à l'école que l'on note K_j . Le modèle devient ainsi :

$$Y_{ij} = c + \beta X_{ij} + \gamma_1 \bar{X}_j + \gamma_2 K_j + V_j + \varepsilon_{ij}$$

A partir de ce modèle, on peut dégager trois spécifications :

Modèle 1

Dans la première spécification du modèle, les effets de pairs sont supprimés de l'équation, ainsi le modèle devient :

$$Y_{ij} = c + \beta X_{ij} + \gamma_2 K_j + V_j + \varepsilon_{ij}$$

Avant d'estimer ce modèle, on utilise le test de Hausman qui permet de tester l'existence d'endogénéité de niveau 2. Ce test permet de comparer le modèle à effets fixes, contenant les X_{ij} seulement, au modèle à effets aléatoires, contenant les X_{ij} et les K_j .

Rappelons que pour le modèle à effets fixes les V_j sont traités comme des paramètres fixes, ainsi le problème d'endogénéité de niveau 2 ne se pose pas, contrairement au modèle à effets aléatoires où V_j est une variable aléatoire.

En utilisant la transformation de Fuller et Battese (1973), le modèle devient :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_2(K_j - \lambda K_j) + w_{ij}$$

avec $w_{ij} = V_j + \varepsilon_{ij}$

Les estimateurs de \square sont obtenus par un OLS.

Ajoutons et retranchons $\beta \bar{X}_{.j}$:

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_2(K_j - \lambda K_j) + \beta \bar{X}_{.j} - \beta \bar{X}_{.j} + w_{ij}$$

En développant l'équation on aura :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta X_{ij} - \lambda \beta \bar{X}_{.j} + \gamma_2 K_j - \lambda \gamma_2 K_j + \beta \bar{X}_{.j} - \beta \bar{X}_{.j} + w_{ij}$$

Finalement l'équation devient :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \bar{X}_{.j}) + (1 - \lambda)\beta \bar{X}_{.j} + (1 - \lambda)\gamma_2 K_j + w_{ij}$$

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \bar{X}_{.j}) + (1 - \lambda)\beta \bar{X}_{.j} + \delta_2 K_j + w_{ij}$$

$$\text{Avec } \delta_2 = (1 - \lambda)\gamma_2 \text{ et } \lambda = \bar{\lambda}_j \text{ avec } \lambda_j = 1 - \frac{\sigma_w^2}{\sigma_w^2 + n_j \sigma_b^2}$$

σ_w^2 : est la variance intra-école.

σ_b^2 : est la variance inter-école.

n_j : est le nombre d'élèves dans l'école j .

Identification et estimation

On suppose que $(X_{ij} - \bar{X}_{.j})$ et K_j sont indépendants, ainsi leurs effets peuvent être estimés séparément.

1. On régresse $(Y_{ij} - \bar{Y}_{.j})$ sur $(X_{ij} - \bar{X}_{.j})$. On obtient les $\hat{\beta}$.

2. On régresse $\bar{Y}_{.j}$ sur $\bar{X}_{.j}$ et K_j . On obtient $\hat{\gamma}_2$.

3. On calcule λ_j avec la formule $\lambda_j = 1 - \frac{\sigma_w^2}{\sigma_w^2 + n_j \sigma_b^2}$

Pour obtenir λ on calcule la moyenne des λ_j .

4. On multiplie $\hat{\gamma}_2$ par $(1 - \lambda)$ pour obtenir $\widehat{\delta}_2$.

Modèle 2

Dans la deuxième spécification du modèle, on supprime les caractéristiques de l'école K_j . Le modèle devient donc :

$$Y_{ij} = c + \beta X_{ij} + \gamma_1 \bar{X}_{.j} + V_j + \varepsilon_{ij}$$

Comme dans le modèle 1, le test de Hausman est utilisé pour comparer le modèle à effets fixes, contenant seulement les caractéristiques de l'élève X_{ij} , au modèle à effets aléatoires, contenant les X_{ij} et les $\bar{X}_{.j}$.

En utilisant la transformation de Fuller et Battese (1973), le modèle devient :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_1(\bar{X}_{.j} - \lambda \bar{X}_{.j}) + w_{ij}$$

Avec $w_{ij} = V_j + \varepsilon_{ij}$

Les estimateurs de β sont obtenus par un OLS.

Ajoutons et retranchons $\beta \bar{X}_{.j}$:

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_1(\bar{X}_{.j} - \lambda \bar{X}_{.j}) + \beta \bar{X}_{.j} - \beta \bar{X}_{.j} + w_{ij}$$

En développant l'équation on aura :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta X_{ij} - \lambda \beta \bar{X}_{.j} + \gamma_1 \bar{X}_{.j} - \lambda \gamma_1 \bar{X}_{.j} + \beta \bar{X}_{.j} - \beta \bar{X}_{.j} + w_{ij}$$

Finalement l'équation devient :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \bar{X}_{.j}) + (\gamma_1 - \lambda \gamma_1 - \lambda \beta + \beta) \bar{X}_{.j} + w_{ij}$$

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \bar{X}_{.j}) + (1 - \lambda)(\gamma_1 + \beta) \bar{X}_{.j} + w_{ij}$$

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \bar{X}_{.j}) + \delta_1 \bar{X}_{.j} + w_{ij}$$

$$\text{Avec } \delta_1 = (1 - \lambda)(\gamma_1 + \beta) \text{ et } \lambda = \bar{\lambda}_j \text{ avec } \lambda_j = 1 - \frac{\sigma_w}{\sqrt{\sigma_w^2 + n_j \sigma_b^2}}$$

σ_w^2 : est la variance intra école.

σ_b^2 : est la variance inter école.

n_j : est le nombre d'élèves dans l'école j.

Identification et estimation

$(X_{ij} - \bar{X}_{.j})$ et $\bar{X}_{.j}$ sont orthogonaux, ceci permet d'estimer leurs coefficients séparément.

1. On régresse $(Y_{ij} - \bar{Y}_{.j})$ sur $(X_{ij} - \bar{X}_{.j})$. On obtient les $\hat{\beta}$.

2. On régresse $\bar{Y}_{.j}$ sur $\bar{X}_{.j}$. On obtient $(\widehat{\gamma_1 + \beta})$.

3. On calcule λ_j avec la formule $\lambda_j = 1 - \frac{\sigma_w}{\sqrt{\sigma_w^2 + n_j \sigma_b^2}}$

Pour obtenir λ on calcule la moyenne des λ_j .

4. On multiplie $(\widehat{\gamma_1 + \beta})$ par $(1 - \lambda)$ pour obtenir $\hat{\delta}_1 = (1 - \lambda)(\widehat{\gamma_1 + \beta})$.

On a $\hat{\beta}$, $(\widehat{\gamma_1 + \beta})$, λ et $\hat{\delta}_1$, on peut calculer $\hat{\gamma}_1$.

Modèle 3

Dans la troisième spécification on estime le modèle complet suivant :

$$Y_{ij} = c + \beta X_{ij} + \gamma_1 \bar{X}_{.j} + \gamma_2 K_j + V_j + \varepsilon_{ij}$$

On utilise le test de Hausman pour comparer le modèle à effets fixes contenant les caractéristiques de l'élève au modèle à effets aléatoires contenant les caractéristiques de l'élève X_{ij} , les effets de pairs $\bar{X}_{.j}$ et les caractéristiques de l'école K_j .

Pour estimer ce modèle, on utilise la transformation de Fuller et Battese (1973) :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_1(\bar{X}_{.j} - \lambda \bar{X}_{.j}) + \gamma_2(K_j - \lambda K_j) + w_{ij}$$

$$\text{Avec } w_{ij} = V_j + \varepsilon_{ij}$$

En ajoutant et retranchant $\beta \bar{X}_{.j}$ l'équation devient :

$$Y_{ij} - \lambda \bar{Y}_{.j} = c + \beta(X_{ij} - \lambda \bar{X}_{.j}) + \gamma_1(\bar{X}_{.j} - \lambda \bar{X}_{.j}) + \gamma_2(K_j - \lambda K_j) + \beta \bar{X}_{.j} - \beta \bar{X}_{.j} + w_{ij}$$

En développant l'équation on aura :

$$Y_{ij} - \lambda \bar{Y}_j = c + \beta X_{ij} - \lambda \beta \bar{X}_j + \gamma_1 \bar{X}_j - \lambda \gamma_1 \bar{X}_j + \gamma_2 K_j - \lambda \gamma_2 K_j + \beta \bar{X}_j - \beta \bar{X}_j + w_{ij}$$

Finalement l'équation devient :

$$Y_{ij} - \lambda \bar{Y}_j = c + \beta(X_{ij} - \bar{X}_j) + (\gamma_1 - \lambda \gamma_1 - \lambda \beta + \beta) \bar{X}_j + (1 - \lambda) \gamma_2 K_j + w_{ij}$$

$$Y_{ij} - \lambda \bar{Y}_j = c + \beta(X_{ij} - \bar{X}_j) + (1 - \lambda)(\gamma_1 + \beta) \bar{X}_j + (1 - \lambda) \gamma_2 K_j + w_{ij}$$

$$Y_{ij} - \lambda \bar{Y}_j = c + \beta(X_{ij} - \bar{X}_j) + \delta_1 \bar{X}_j + \delta_2 K_j + w_{ij}$$

$$\text{Avec } \delta_1 = (1 - \lambda)(\gamma_1 + \beta), \delta_2 = (1 - \lambda)\gamma_2 \text{ et } \lambda = \bar{\lambda}_j \text{ avec } \bar{\lambda}_j = 1 - \frac{\sigma_w}{\sqrt{\sigma_w^2 + n_j \sigma_b^2}}$$

σ_w^2 : est la variance intra-école.

σ_b^2 : est la variance inter-école.

n_j : est le nombre d'élèves dans l'école j .

Identification et estimation

$(X_{ij} - \bar{X}_j)$ et \bar{X}_j sont orthogonaux et on suppose que $(X_{ij} - \bar{X}_j)$ et K_j sont indépendants, ainsi les effets des différentes composantes peuvent être estimés séparément.

1. On régresse $(Y_{ij} - \bar{Y}_j)$ sur $(X_{ij} - \bar{X}_j)$, on obtient les $\hat{\beta}$.

2. On régresse \bar{Y}_j sur \bar{X}_j et K_j , on obtient $(\widehat{\gamma_{1+\beta}})$ et $\hat{\gamma}_2$.

3. On calcule λ_j avec la formule $\lambda_j = 1 - \frac{\sigma_w}{\sqrt{\sigma_w^2 + n_j \sigma_b^2}}$

Pour obtenir \square on calcule la moyenne des λ_j .

4. On multiplie $(\widehat{\gamma_{1+\beta}})$ et $\hat{\gamma}_2$ par $(1 - \lambda)$ pour obtenir $\hat{\delta}_1 = (1 - \lambda)(\widehat{\gamma_{1+\beta}})$ et $\hat{\delta}_2 = (1 - \lambda)\hat{\gamma}_2$.

On a $\hat{\beta}$, $(\widehat{\gamma_{1+\beta}})$, λ et $\hat{\delta}_1$, on peut calculer $\hat{\gamma}_1$.

Références

- ALTINOK N. et KINGDON G. (2009), « New Evidence on Class Size Effects: A pupil Fixed Effects Approach », *CSAE Workings Papers*, n° 2009-16.
- AMMERMUELLER A. et PISCHKE J.S. (2009), « Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study », *Journal of Labor economics*, volume 27, issue 3, p. 315-448.
- CAILLE J.P. (2001), « Scolarisation à 2 ans et réussite de la carrière scolaire au début de l'école élémentaire », *Education et Formation*, n° 60, p. 7-18.
- CAILLE J.P. et ROSENWALD F. (2006), « Les inégalités de réussite à l'école élémentaire: construction et évolution », *La documentation française*, INSEE, France.
- CARD D. et KRUEGER A. (1996), « School Resources and Student Outcomes: an Overview of the Literature and New Evidence from North and South Carolina », *Journal of Economic Perspectives*, vol. 10, p. 31-40.
- CHEVALIER A. et LANOT G. (2002), « The Relative Effect of Family Characteristics and Financial Situation on Educational Achievement », *Education Economics*, vol. 10, issue 2, p. 165-182.
- COLEMAN J.S., CAMPBELL E.Q., HOBSON C.J., MCPARTLAND J., MOOD A.M. et WEINFELD F.D. (1966), *Equality of Educational Opportunity*, National Center for Educational Statistics, Washington, DC.
- DOWNEY D.B. (1995), « Bigger is not Better: Family Size, Parental Resources, and Children's Educational Performances », *American Sociological Review*, vol. 60, p. 746-761.
- DUMAY X. et DUPRIEZ V. (2004), « Effet établissement: effet de processus et/ou effet de composition? », *les Cahiers de recherche en éducation et formation*, n° 36, décembre, p. 3-20.
- FIGLIO D. (2005), « Boys Named Sue: Disruptive Children and Theirs Peers », *NBER Working Paper*, n° 11277.
- FUCHS T. & WOBMANN L. (2004), « What Accounts for International Differences in Student Performance? A Re-examination Using PISA Data », *CESIFO Working Paper Series*, n° 1235.
- FULLER B. et CLARKE P. (1994), « Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules and Pedagogy », *Review of Educational Research*, vol. 64, p. 122-131.
- GANZACH Y. (2000), « Parent's Education, Cognitive Ability, Educational Expectation and Educational Attainment », *British Journal of educational Psychology*, vol. 70, p. 419-441.
- GREENWALD R., HEDGES L.V. et LAINE R. (1996), « The Effect of School Resources on Student Achievement », *Review of Educational Research*, vol. 66, p. 361-396.
- GRISAY A. (1993), « Le fonctionnement des collèges et ses effets sur les élèves de sixième et de cinquième », *Les dossiers d'Éducation et Formation*, MEN, Direction de l'évaluation et de la prospective, n° 32, novembre.
- HAKKINEN I., KIRJAVAINEN T. et UUSITALO R. (2003), « School Resources and Student Achievement Revisited: New Evidence from Panel Data », *Economics of Education Review*, vol. 22, p. 329-335.
- HANUSHEK E.A. (1997), « Assessing the Effects of School Resources on Student Performance: an Update », *Educational Evaluation and Policy Analysis*, vol. 9, issue 2, p. 141-164.
- HANUSHEK et LUQUE J. (2003), « Efficiency and Equity in Schools around the World », *Economics of Education Review*, vol. 22, issue 5, p. 481-502.

- HANUSHEK (2003), « The Failure of Input-based Schooling Policies », *The Economic Journal*, vol. 11, p. 64-98.
- HANUSHEK (2006), « School Resources », in *Handbook of the Economics of Education*, Hanushek et F. Welch, Elsevier Edition, chap. 4, p. 865-908.
- HANCHANE S. et MOSTAFA T. (2011), « Solving Endogeneity Problems in Multilevel Estimation: An Example Using Education Production Functions », *Journal of Applied Statistics*.
- HEYNEMAN S.P. et 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*, vol. 88, issue 6, p. 1162-1194.
- HOLMES C.T. et MATTHEWS K.M., « The Effects of Nonpromotion on Elementary and Junior High School Pupils: a Meta-analysis », *Review of Educational Research*, vol. 54, 1984, p. 225-236.
- HOLMES (1990), « Grade Level Retention Effects: A Meta-analysis of Research Studies », in Shepard L. & Smith Mary Lee (Eds), *Flunking Grades. Research on Retention*, Falmer Press, p. 16-33.
- HOY W.K. et HANNUM J.W. (1997), « Middle School Climate: an Empirical Assessment of Organizational Health and Student Achievement », *Educational Administration Quarterly*, vol. 33, issue 3, p. 290-311.
- IVERSON B.K. et WALBERG H.J. (1982), « Home Environment and Learning: a Quantitative synthesis », *Journal of Experimental Education*, vol. 50, p. 144-51.
- KEITH T.Z., REIMERS T.M., FEHRMANN P.G., POTTEBAUM S.M. et AUBEY L.W. (1986), « Parental Involvement, Homework, and TV Time: Direct and Indirect Effects on High School Achievement », *Journal of Educational Psychology*, vol. 5, p. 373-380.
- KONU A. et RIMPELA M. (2002), « Well-Being in Schools: a Conceptual Model Health, Health Promotion International », *Oxford Journals*, vol. 17, issue 1, p. 79-87.
- LEUVEN E., LINDHAL M., OOSTERBEEK H. et WEBBINK D. (2007), « The Effect of Extra Funding for Disadvantaged Pupils on Achievement », *The Review of Economics and Statistics*, vol. 89, issue 4, p. 721-736.
- MANACORDA M. (2008), « The Cost of Grade Retention », *CEP Discussion Paper*, n° 878.
- MEURETET D., MORLAIX S. (2006), « L'influence de l'origine sociale sur les performances scolaires: par où passe-t-elle? », *Revue française de sociologie*, vol. 47, issue 1, p. 49-79.
- PARCEL T.L. et DUFUR M.J. (2001), « Capital at Home and at School: Effects on Child Social Adjustment », *Journal of Marriage and the Family*, vol. 63, n° 1, p. 32-47.
- PARCEL T.L. et MENAGHAN E.G. (1994), « Early Parental Work, Family Social Capital, and Early Childhood Outcomes », *American Journal of Sociology*, vol. 99, p. 972-1009.
- POSTLETHWAITE T.N. et WILEY D.E. (1992), « The IEA Study of Science II: Science Achievement in Twenty-three Countries », Pergamon Press, Oxford.
- RUBIN D.B. (1987), « Multiple Imputation for Non-response in Surveys », *Wiley Series in Probability and Statistics*.
- SCHILLER K.S., KHMELKOV V.T. et WANG X.Q. (1978), « Economic Development and the Effect of Family Characteristics on Mathematics Achievement », *Journal of Marriage and Family*, vol. 64, 2002, p. 730-742.
- SEIBEL C. (1984), « Genèse et conséquences de l'échec scolaire: vers une politique de prévention », *Revue française de pédagogie*, n° 67, p. 7-28.
- SIMMONS J. et ALEXANDER L., « The Determinants of School Achievement in Developing Countries: a Review of the Research », *Economic Development and Cultural Change*, vol. 6, issue 2, p. 341-357.
- STEINBERG L., LAMLBORN S.D., DORNBUSH S.M. et DARLING N. (1992), « Impact of Parenting

- Practices on Adolescent Achievement: Authoritative Parenting School Involvement, and Encouragement to Succeed», *Child Development*, vol. 63, p. 1266-1281.
- TEACHMAN J.-D., PAASCH K. et CARVER K. (1996), «Social Capital and Dropping out of School Early», *Journal of Marriage and the Family*, volume 58, p. 773-783.
- THOMPSON G.H. et JOHNSTON J.S. (2006), «Variation in the Relationship between Nonschool Factors and Student Achievement on International Assessments», National Center for Education Statistics, Institute of Education Sciences, NCES.
- VEENSTRA R. et KUYER H. (2004), «Effective Students and Families: the Importance of Individual Characteristics for Achievement in High School», *Educational Research and Evaluation*, vol. 10, issue 1, p. 41-70.
- WEST M.R. et WÖHMANN L. (2006), «Class-size Effects in School Systems around the World: Evidence from Between-grade Variation in TIMSS», *European Economic Review*, n° 50, p. 695-736.
- WILLMSET J.D., SOMERS M.A. (2001), «Family, Classroom and School Effects on Children's Educational Outcomes in Latin America», in *School Effectiveness and School Improvement* 12 (4), p. 409-445.
- WÖSSMANN L. et SCHÜTZ G. (2006), Efficiency and Equity in European Education and Training Systems, Analytical Report for the European Commission, SWP 060426, 43.
- WÖSSMANN L. (2008), «Efficiency and Equity of European Education and Training Policies», *International Tax and Public Finance*, vol. 15, n° 2, p. 199-230.
- YAYAN B. et BERBEROLU G. (2004), «A Re-Analysis of the TIMSS 1999 Mathematics Assessment Data of the Turkish Students», *Studies in Educational Evaluation*, vol. 30, p. 87-104.