

10. MODÈLE SOMMAIRE, EFFETS PERSISTANTS ET APPROFONDISSEMENT DE LA RECHERCHE

Les résultats présentés dans les chapitres précédents renforcent les conclusions d'autres études à grande échelle, à savoir qu'un grand nombre de variables sont associées aux résultats en lecture, de manière positive et négative. Cependant, les modèles montrent aussi clairement que bon nombre de ces variables sont corrélées entre elles de façon complexe, à tel point que beaucoup d'effets bivariés simples sont atténués quand d'autres variables sont neutralisées.

L'objectif ici est d'identifier les effets pouvant être considérés comme « persistants », c'est-à-dire qui restent significatifs même après neutralisation de nombreuses autres variables.

La méthode suivie pour l'analyse multivariée des chapitres précédents a consisté à neutraliser les variables démographiques, à partir de l'hypothèse qu'il s'agit de conditions antécédentes, sur lesquelles ni les écoles ni le personnel enseignant n'a de contrôle, de sorte qu'il faut en tenir compte avant d'examiner les effets des variables qui sont d'intérêt plus direct dans l'optique d'une politique d'éducation. Cette méthode peut être élargie à un modèle plus exhaustif, dans lequel les variables de toutes les catégories examinées précédemment sont neutralisées. Les coefficients d'un tel modèle peuvent être considérés comme uniques ou comme

des effets résiduels de variables individuelles, après neutralisation de tout le reste.

L'analyse de tous les liens possibles entre les variables dépasse le cadre du présent rapport. En fait, il serait pratiquement impossible de la faire sans examiner aussi tous les rapports théoriquement possibles entre les variables et renvoyer à d'autres recherches qui peuvent nous aiguiller dans la direction voulue. Par exemple, l'effet de l'effectif de la classe constaté dans la présente étude (à savoir que les élèves des classes ayant un effectif plus nombreux ont un meilleur rendement en lecture), est contraire aux attentes et s'éloigne des résultats d'autres recherches ainsi que des orientations stratégiques suivies dans bon nombre d'instances. Il est toutefois vraisemblable que l'effectif de la classe soit confondu avec de nombreuses autres variables démographiques telles que l'instance, la taille et l'emplacement de l'école et d'autres effets plus subtils comme la tendance à former de plus petites classes à l'intention des élèves ayant de moins grandes compétences ou des différences de stratégies pédagogiques utilisées selon un petit ou un grand effectif de la classe. Il n'est évidemment pas possible de sonder toutes les corrélations possibles entre l'effectif de la classe et le rendement en lecture dans la présente étude. Toutefois, des recherches axées plus explicitement sur cette variable pourraient éclairer les raisons de ces effets contraires aux attentes. Avant de tirer toute conclusion définitive des nombreux effets constatés ici sur le plan stratégique, il faut sonder plus avant ces raisons et établir des liens avec d'autres recherches sur le même sujet.

Étant donné la tendance générale à une réduction des divers effets dans le modèle multivarié par rapport au modèle bivarié, il paraît justifié de construire un modèle plus exhaustif à partir de ces variables qui présentent des effets statistiquement significatifs dans les modèles bivariés. Le présent chapitre montre les résultats des premières tentatives de construction de ce modèle, conçu pour isoler les effets des variables choisies (principalement celles qui produisent des effets statistiquement significatifs dans les modèles bivariés). L'objectif ici est d'identifier les effets pouvant être considérés comme « persistants », c'est-à-dire qui restent significatifs même après neutralisation de nombreuses autres variables. À cause de leurs effets persistants, ils peuvent être considérés comme d'intérêt direct pour l'élaboration d'une politique. Simultanément, on peut considérer que ces variables qui donnent lieu à des différences

significatives méritent une étude plus approfondie des sources de ces changements. Dans la perspective d'une modélisation statistique, il faut penser à une démarche par étape, qui soumettrait les variables d'intérêt direct à une neutralisation progressive.

Modèle sommaire

Le modèle présenté ici inclut 30 variables au niveau des élèves et 10 au niveau des écoles, choisies à cause de leur importance statistique dans les modèles bivariés. Il inclut aussi quelques variables qui ne sont pas statistiquement significatives dans les modèles bivariés, mais qui présentent un intérêt dans l'optique de l'élaboration d'une politique. C'est le cas entre autres du nombre de minutes par semaine consacrées aux arts du langage, là où les résultats, comme ceux de l'effectif de la classe, sont contraires aux attentes et divergent des résultats d'autres recherches.

Observation statistique : données manquantes

Les données manquantes représentent un problème important dans les modèles multivariés comportant un grand nombre de variables. En effet, pour peu que des données soient manquantes pour toute variable d'un cas particulier, ce dernier est supprimé avant la compilation du modèle. Ce problème n'est pas important si les données manquantes le sont de manière aléatoire et si l'on dispose d'un grand nombre de cas. Mais si les réponses manquantes sont systématiquement liées à d'autres réponses et particulièrement au résultat (p. ex., si les élèves dont le rendement est faible ne répondent pas à certains items), les données manquantes peuvent avoir une incidence significative sur les résultats. Il existe des solutions techniques (dont l'imputation), mais elles n'ont pas été appliquées ici faute de temps. Le modèle présenté est donc préliminaire et l'interprétation des résultats appelle à la prudence. Il est recommandé d'approfondir les recherches à partir de la banque de données du PPCE en appliquant les méthodes d'imputation appropriées pour réduire l'incidence des données manquantes.

Il convient à ce point de rappeler certaines des principales caractéristiques des modèles utilisés.²⁸ Ceci peut être résumé ci-dessous :

- La méthode est une variante de l'analyse de régression multiple, spécifiquement conçue pour traiter des données de nature hiérarchique. Dans ce cas, la hiérarchie est déterminée par une méthode d'échantillonnage en deux étapes, selon laquelle les écoles puis les élèves ont été choisis.
- L'existence de données provenant des questionnaires de l'élève, du personnel enseignant et de l'école produit effectivement des niveaux de modélisation. Toutefois, comme il n'y avait en général qu'un ou deux enseignants ou enseignantes par école, les données produites par le questionnaire du personnel enseignant ont été agrégées aux données du niveau de l'école pour les besoins de la modélisation.
- L'analyse de régression multiple est essentiellement un modèle prédictif. Ce modèle produit des coefficients qui peuvent être interprétés comme l'effet, sur le résultat (c'est-à-dire les scores à l'évaluation en lecture), d'une variation d'une unité dans la valeur d'une variable d'intérêt.
- Les modèles initiaux sont de simples modèles bivariés, dans lesquels est calculé l'effet d'un seul

prédicteur. Les modèles multivariés subséquents incluent d'autres variables, ou covariables, toutes les variables du modèle servant de témoins pour toutes les autres. Toute variation du coefficient correspondant à une variable donnée entre le modèle bivarié et le modèle multivarié peut être interprétée comme une mesure de l'incidence des intercorrélations entre les prédicteurs sur l'effet résiduel d'un prédicteur particulier en présence de tous les autres prédicteurs.

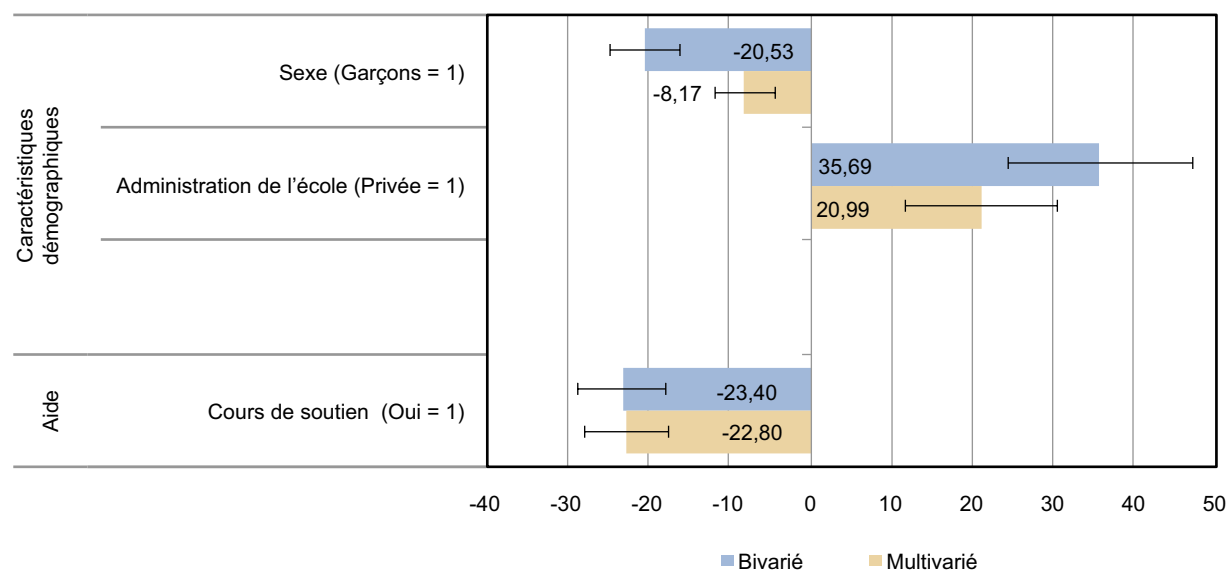
- L'effet du modèle multivarié consiste à « isoler » le plus possible, compte tenu des données disponibles, les effets d'une seule variable de l'effet de toutes les autres.
- Les coefficients de régression calculés dans ces modèles ne sont pas uniformes. Chacun doit être interprété en fonction de l'échelle de la variable d'intérêt. Dans le présent chapitre, les coefficients sont regroupés en fonction de l'échelle utilisée de sorte que ceux qui figurent dans un graphique donné soient au moins approximativement comparables. Le tableau A.1 (Annexe A) donne l'échelle de chaque variable.

Les graphiques 10.1 à 10.4 montrent les coefficients bivariés et multivariés du modèle sommaire.²⁹ Chaque graphique est suivi de brefs commentaires sur les effets observés.

²⁸ Pour une description plus détaillée du processus de modélisation, prière de consulter l'observation statistique vers la fin du chapitre 3.

²⁹ Il peut y avoir de petites différences entre les coefficients bivariés présentés ici et ceux des chapitres précédents. Ceci est dû aux différences de traitement des données manquantes dans les deux cas et à un changement mineur d'échelle pour certaines variables. Aucune de ces différences n'a d'effet significatif sur les résultats.

Graphique 10.1 Coefficients de régression portant sur les variables à échelle dichotomique

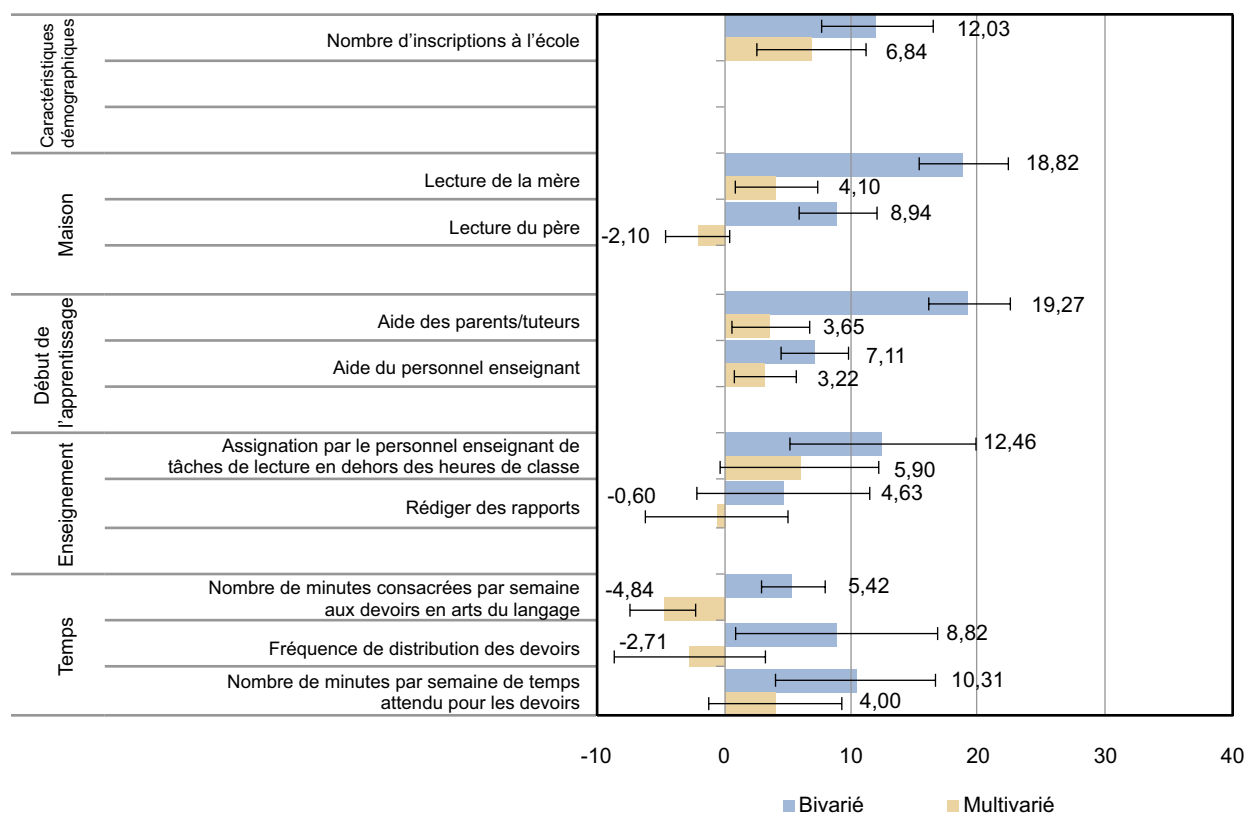


- L'effet du sexe semble indiquer que les scores moyens en lecture des garçons sont inférieurs. Cet effet est significativement atténué dans le modèle multivarié, ce qui suggère que des facteurs autres que le sexe contribuent à l'effet bivarié.
- L'administration de l'école est très favorable aux écoles privées. Cet effet n'est pas

significativement atténué dans le modèle multivarié, ce qui indique qu'il n'est pas tellement influencé par les autres variables du modèle.

- L'effet des cours de soutien est négatif dans les deux modèles, et il n'y a pas de changement significatif entre le modèle bivarié et le modèle multivarié.

Graphique 10.2 Coefficients de régression portant sur les variables à l'échelle ordinale³⁰ 1-3 et 1-4



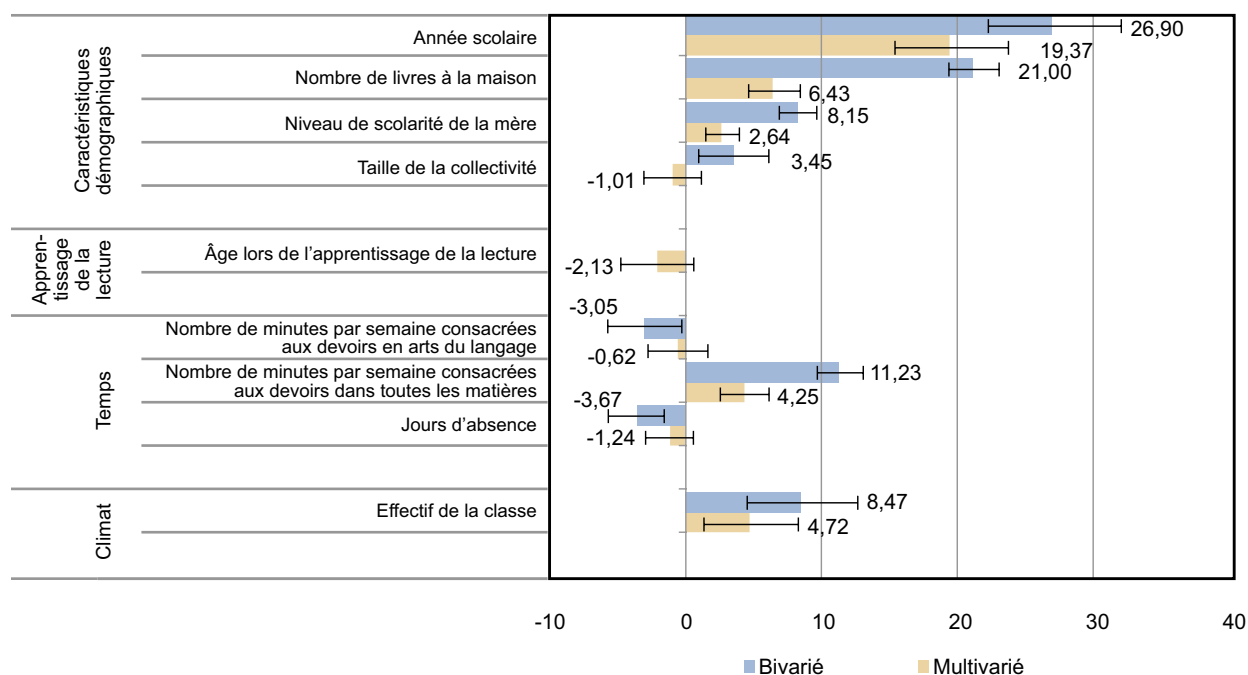
- Les écoles de plus grandes tailles sont associées à des rendements en lecture élevés dans les deux modèles, ce qui indique que cet effet n'est pas significativement influencé par d'autres variables du modèle.
- Le fait que la mère lise à la maison a un effet positif dans les deux modèles. Le fait que le père lise à la maison a un effet positif dans le modèle bivarié et un effet non significatif dans le modèle multivarié. Ceci indique que le fait que l'élève voit son père lire à la maison est contrebalancé par d'autres facteurs, par exemple ses propres comportements de lecture.
- L'aide des parents/tuteurs et du personnel enseignant au début de l'apprentissage de la lecture a des effets positifs dans le modèle bivarié, qui sont atténués quelque peu par la présence des autres variables dans le modèle multivarié.

L'effet des parents/tuteurs est significativement plus grand que celui du personnel enseignant, au niveau bivarié, mais il est aussi très influencé par les autres variables, comme le montre le modèle multivarié.

- Dans le modèle bivarié, le fait que le personnel enseignant assigne des tâches de lecture en dehors des heures de classe a un effet positif et la rédaction de rapports a un effet positif marginal. Dans le modèle multivarié, ces effets ne sont pas significativement différents du zéro.
- Toutes les variables relatives aux devoirs ont un effet positif sur le rendement en lecture dans le modèle bivarié. Le nombre de minutes chaque semaine consacrées aux devoirs en arts du langage a toutefois un effet négatif dans le modèle multivarié et les effets des autres variables liées aux devoirs sont non significatifs.

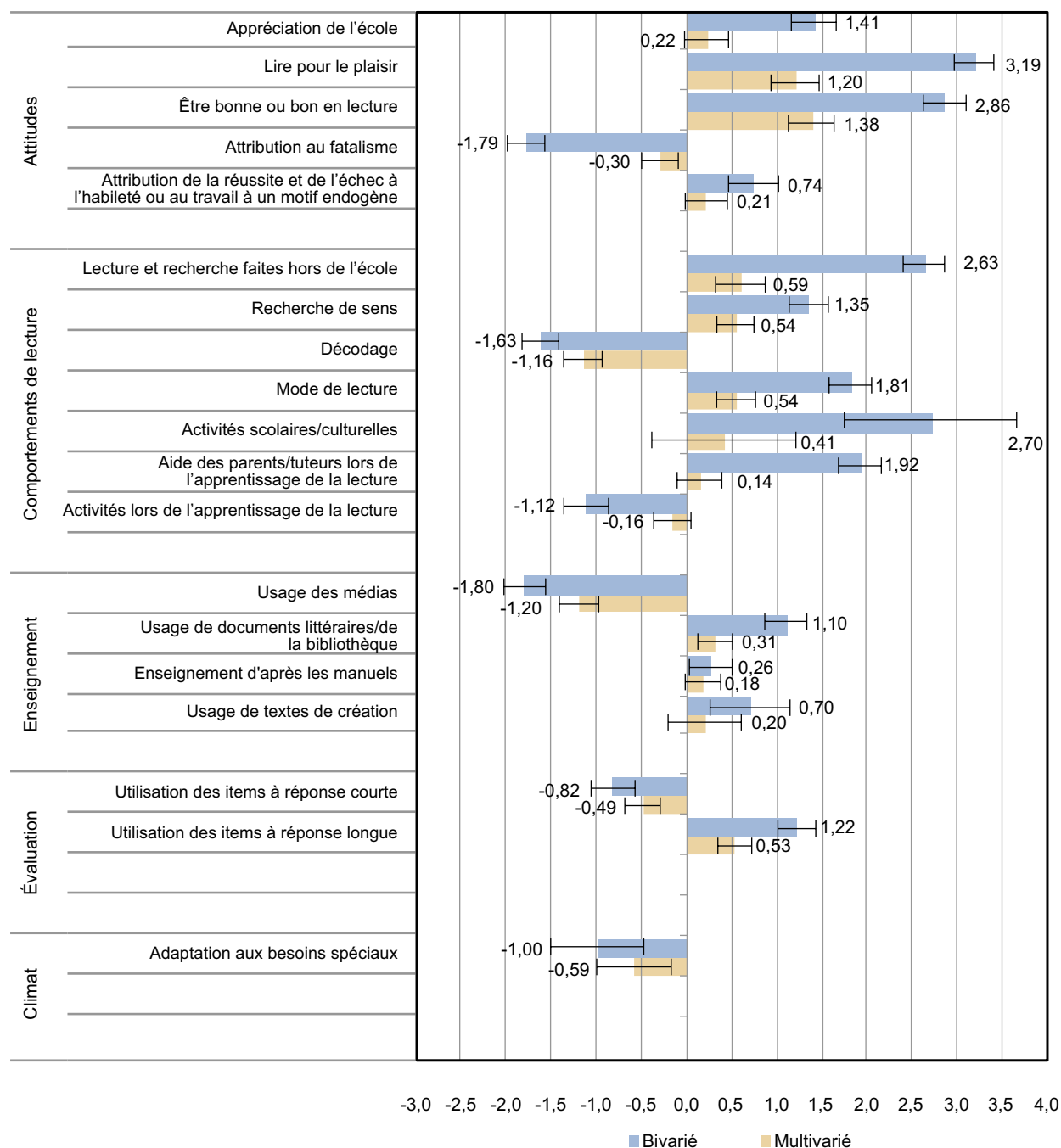
³⁰ Une échelle ordinale est une échelle fondée sur un classement, mais dont les intervalles ne sont pas égaux entre les valeurs d'échelle. Par exemple, la taille de l'école est présentée sur une échelle de 1 à 4, où 1 = 100 ou moins; 2 = 101 à 500; 3 = 501 à 1000; 4 = plus grand que 1000.

Graphique 10.3 Coefficients de régression portant sur les variables à l'échelle ordinale 1-5 et 1-6



- L'année scolaire dans laquelle se trouve l'élève au moment du test a une forte corrélation positive avec le rendement en lecture dans les deux modèles, mais l'effet est significativement plus petit dans le modèle multivarié.
- Les deux variables d'ordre socio-économique, soit le nombre de livres à la maison et le niveau de scolarité de la mère sont aussi étroitement reliées au rendement en lecture dans les deux modèles. Ces deux effets sont significativement atténués dans le modèle multivarié, ce qui laisse à penser que d'autres variables peuvent contrebalancer dans une certaine mesure les effets du statut socio-économique.
- Les élèves dont l'école est située dans une grande collectivité ont un meilleur rendement que ceux qui fréquentent une école située dans une plus petite collectivité. Cependant, une fois les autres variables neutralisées, l'effet de la taille de la collectivité est essentiellement réduit à zéro.
- Les élèves qui apprennent à lire à un plus jeune âge ont un meilleur rendement en lecture que ceux qui apprennent à lire plus tardivement. Encore une fois, cet effet est contrebalancé par d'autres variables dans le modèle multivarié.
- Le temps consacré aux devoirs dans toutes les matières et le temps consacré aux devoirs en arts du langage ont des effets positifs sur le rendement en lecture. L'effet des devoirs en arts du langage est inverse dans le modèle multivarié. Ceci est probablement dû en partie au fait que les devoirs en arts du langage sont inclus dans le total. Mais l'inversion suggère aussi l'incidence possible d'autres facteurs.
- Les jours d'absence ont un effet négatif sur le rendement, mais cet effet est relativement petit par rapport aux autres effets de cette série. Il n'est toutefois pas significativement atténué par d'autres variables.
- Le nombre de minutes consacrées chaque semaine aux arts du langage n'a pas d'effet significatif sur le rendement en lecture.
- Finalement, l'effectif de la classe a un effet significatif dans les deux modèles et semble indiquer que les élèves des classes ayant un effectif plus nombreux ont un meilleur rendement que ceux des plus petites classes, même après neutralisation des facteurs de confusion comme la taille de l'école ou de la collectivité.

Graphique 10.4 Coefficients de régression portant sur les variables regroupées en scores factoriels³¹



³¹ Comme on l'a vu précédemment, tous les facteurs ont été ramenés à une moyenne de 50 et à un écart-type de 10. Une différence de un point d'un score factoriel représente donc un changement d'unité beaucoup plus petit que dans le cas d'autres variables. C'est pourquoi l'échelle utilisée dans ce graphique est beaucoup plus petite. La lectrice ou le lecteur voudra peut-être examiner ces résultats sur la base d'unités d'écart-type. Pour calculer la différence de score en lecture associée à une variation égale à la valeur d'un écart-type, il faut multiplier le coefficient par 10.

- L'attitude à l'égard de l'école et de la lecture a un effet généralement positif dans les deux modèles. L'exception est l'appréciation de l'école, dont l'effet devient marginal dans le modèle multivarié. Tous ces effets sont en outre significativement réduits dans le modèle multivarié, ce qui indique que d'autres variables sont liées à l'attitude.
- L'effet négatif de l'attribution au fatalisme et l'effet positif de l'attribution du succès et de l'échec à un motif endogène sont atténués dans le modèle multivarié, bien que l'effet de l'attribution au fatalisme reste significativement négatif.
- Les effets de la lecture et de la recherche hors de l'école, de la lecture pour le sens et du mode de lecture sont significativement positifs dans les deux modèles, mais réduits par les effets d'autres variables dans le modèle multivarié.
- La lecture par décodage a un effet négatif dans les deux modèles.
- Les effets positifs de l'aide des parents/tuteurs lors de l'apprentissage de la lecture constatés dans le modèle bivarié sont contrebalancés par les effets d'autres variables dans le modèle multivarié.
- L'apprentissage de la lecture dès le plus jeune âge est associé à un rendement en lecture plus élevé dans les deux modèles. (Cet effet semble négatif dans le graphique parce que la valeur numérique de l'échelle de l'apprentissage dès le plus jeune âge est inférieure à celle des débuts plus tardifs.)

Toutefois, dans le modèle bivarié, d'autres activités de lecture dès le plus jeune âge ont un effet négatif qui est réduit à près de zéro dans le modèle multivarié.

- Les effets négatifs de l'usage des médias pour lire et les effets positifs de la lecture de documents littéraires restent similaires dans les deux modèles. Toutefois, les effets positifs relativement faibles de la lecture de manuels et de textes de création sont réduits à près de zéro dans le modèle bivarié.
- L'utilisation d'items à réponse longue pour évaluer la lecture a un effet positif tandis que l'utilisation d'items à réponse courte a un effet négatif et ce, dans les deux modèles.
- L'effet composite de l'adaptation pour certains élèves ayant des besoins spéciaux est lui aussi négatif dans les deux modèles, mais relativement plus faible que celui de plusieurs des autres facteurs du modèle.

Effets persistants

Parmi les nombreuses variables qui semblent liées à la réussite quand elles sont prises individuellement, seules quelques-unes conservent un effet statistiquement significatif quand un grand nombre des autres variables sont neutralisées. Dans le cadre de cette étude, les variables qui sont statistiquement significatives pour les deux modèles peuvent être considérées comme suffisamment persistantes pour influencer directement sur la politique et la pratique. Ces effets persistants sont montrés dans le tableau 10.1.

Tableau 10.1 Effets positifs et négatifs persistants

Positifs	Négatifs
Sexe féminin	Sexe masculin
École privée	Cours de soutien
Assignation par le personnel enseignant de tâches de lecture en dehors des heures de classe	Nombreux jours d'absence de l'école
Grande taille de l'école	Lecture par décodage
Année scolaire supérieure	Lecture dans les médias
Grand nombre de livres à la maison	Utilisation d'items à réponses courtes pour l'évaluation de la lecture
Niveau de scolarité de la mère	Adaptation pour les élèves ayant des besoins spéciaux
Mère lisant à la maison	
Aide des parents/tuteurs lors de l'apprentissage de la lecture	
Aide du personnel enseignant lors de l'apprentissage de la lecture	
Être dans une classe ayant un grand effectif	
Appréciation de la lecture par l'élève	
Perception de l'élève d'être bon ou bonne en lecture	
Lecture pour le sens	
Modes de lecture	
Lecture hors de la classe	
Lecture d'œuvres littéraires/de la bibliothèque	
Utilisation d'items à réponses longues pour l'évaluation de la lecture	

Approfondissement de la recherche

La nature générale du présent rapport empêche d'explorer plus en détail les intercorrélations entre les nombreuses variables qui semblent influencer sur le rendement en lecture. Elle empêche aussi un examen détaillé des liens entre ces résultats et ceux d'autres travaux de recherche d'une part et de l'application de ces résultats pour l'orientation des politiques d'autre part.

La conception du PPCE permet une phase de recherche postérieure à la publication du rapport public et du rapport contextuel. Cette phase

permettra de sonder plus avant certains points qui émanent des données et qui correspondent au programme de recherche du CMEC et des instances membres. L'aperçu des facteurs associés au rendement présenté dans ce rapport devrait paver la voie à cette phase de recherche. Le choix de certains des effets les plus intéressants et leur examen approfondi à l'aide de modèles qui palient à certaines des limites du présent rapport pourront donner lieu à un programme de recherches commandées ou à la transmission de la banque de données à des chercheuses et chercheurs indépendants.

Tableau A.1. Coefficients des modèles bivariés et multivariés de prédicteurs de rendement en lecture

				Modèle bivarié				Modèle multivarié			
Variables ³²	Catégorie	Niveau	Échelle	Coeff	ET	DI	Prob	Coeff	ET	DI	Prob
Sexe	Caractéristiques démographiques	St	0-1	-20,53	2,20	18 856	0	-8,17	1,89	18 812	0
Année scolaire		St	1-5	26,90	2,48	19 192	0	19,37	2,11	18 812	0
Nombre de livres à la maison		St	1-5	21,00	0,94	19 192	0	6,43	0,96	18 812	0
Niveau de scolarité de la mère		St	1-6	8,15	0,70	19 192	0	2,64	0,63	18 812	0
Taille de l'école		Sc	1-4	12,03	2,24	1672	0	6,84	2,19	1662	0,002
Administration de l'école		Sc	0-1	35,69	5,81	1672	0	20,99	4,79	1662	0
Taille de la collectivité		Sc	1-5	3,45	1,31	1672	0,009	-1,01	1,07	1662	0,346
Appréciation de l'école	Attitudes	St	Facteur (50,10)	1,41	0,13	19 192	0	0,22	0,12	18 812	0,082
Lire pour le plaisir		St	Facteur (50,10)	3,19	0,11	19 190	0	1,20	0,14	18 812	0
Être bonne ou bon en lecture		St	Facteur (50,10)	2,86	0,12	19 190	0	1,38	0,13	18 812	0
Attribution au fatalisme		St	Facteur (50,10)	-1,79	0,11	19 190	0	-0,30	0,10	18 812	0,004
Attribution de la réussite et de l'échec à l'habileté ou au travail à un motif endogène		St	Facteur (50,10)	0,74	0,14	19 190	0	0,21	0,12	18 812	0,072
Lecture et recherche faites hors de la classe	Habitudes de lecture	St	Facteur (50,10)	2,63	0,12	19 192	0	0,59	0,14	18 812	0
Lecture pour la recherche de sens		St	Facteur (50,10)	1,35	0,11	19 192	0	0,54	0,11	18 812	0
Décodage		St	Facteur (50,10)	-1,63	0,10	19 192	0	-1,16	0,11	18 812	0
Mode de lecture		St	Facteur (50,10)	1,81	0,12	19 192	0	0,54	0,11	18 812	0
Cours de soutien	Aide	St	0-1	-23,40	2,77	19 192	0	-22,80	2,64	18 812	0
Aide de la mère avec la lecture		St	1-3	18,82	1,79	19 192	0	4,10	1,65	18 812	0,013
Aide du père avec la lecture		St	1-3	8,94	1,56	19 192	0	-2,10	1,29	18 812	0,102
Activités scolaires/ culturelles		St	Facteur (50,10)	2,70	0,49	19 192	0	0,41	0,41	18 812	0,317

³² Variables liées au rendement en lecture

				Modèle bivarié				Modèle multivarié			
Variables	Catégorie	Niveau	Échelle	Coeff	ET	DI	Prob	Coeff	ET	DI	Prob
Aide des parents/tuteurs	Apprentissage dès le plus jeune âge	St	1-4	19,27	1,64	19 192	0	3,65	1,56	18 812	0,019
Aide du personnel enseignant		St	1-4	7,11	1,35	19 192	0	3,22	1,24	18 812	0,01
Encouragement des parents/tuteurs		St	Facteur (50,10)	1,92	0,12	19 192	0	0,14	0,13	18 812	0,275
Âge lors de l'apprentissage de la lecture		St	1-5					-2,13	1,36	18 812	0,116
Activités lors de l'apprentissage de la lecture		St	Facteur (50,10)	-1,12	0,13	19 192	0	-0,16	0,11	18 812	0,123
Nombre de minutes consacrées par semaine aux devoirs en arts du langage	Temps	Sc	1-5 quintile	-3,05	1,38	1672	0,027	-0,62	1,12	1662	0,576
Fréquence de distribution des devoirs		Sc	1-4	8,82	4,05	1672	0,029	-2,71	3,03	1662	0,371
Nombre de minutes par semaine de temps attendu pour les devoirs en arts du langage		Sc	1-4	10,31	3,21	1672	0,002	4,00	2,67	1662	0,134
Nombre de minutes consacrées par semaine aux devoirs		St	1-5	11,23	0,86	19 191	0	4,25	0,91	18 812	0
Nombre de minutes consacrées par semaine aux devoirs dans toutes les matières		St	1-4	5,42	1,28	19 191	0	-4,84	1,32	18 812	0
Jours d'absence		St	1-5	-3,67	1,04	19 191	0,001	-1,24	0,89	18 812	0,163
Usage des médias	Enseignement	St	Facteur (50,10)	-1,80	0,12	19 192	0	-1,20	0,11	18 812	0
Usage de documents littéraires/de la bibliothèque		St	Facteur (50,10)	1,10	0,12	19 192	0	0,31	0,10	18 812	0,002
Enseignement d'après le manuel		St	Facteur (50,10)	0,26	0,12	19 192	0,03	0,18	0,10	18 812	0,078
Usage de textes de création		Sc	Facteur (50,10)	0,70	0,23	1672	0,002	0,20	0,21	1662	0,345
Lecture en dehors de la classe		Sc	1-3	12,46	3,73	1672	0,001	5,90	3,18	1662	0,063
Rédiger des rapports		Sc	1-3	4,63	3,46	1672	0,181	-0,60	2,86	1662	0,833
Usage des items à réponse courte	Évaluation	St	Facteur (50,10)	-0,82	0,12	19 191	0	-0,49	0,10	18 812	0
Usage des items à réponse longue		St	Facteur (50,10)	1,22	0,11	19 191	0	0,53	0,10	18 812	0
Adaptation aux besoins spéciaux	Climat	Sc	Facteur (50,10)	-1,00	0,26	1672	0	-0,59	0,21	1662	0,006
Effectif de la classe		Sc	1-5	8,47	2,06	1672	0	4,72	1,76	1662	0,008