

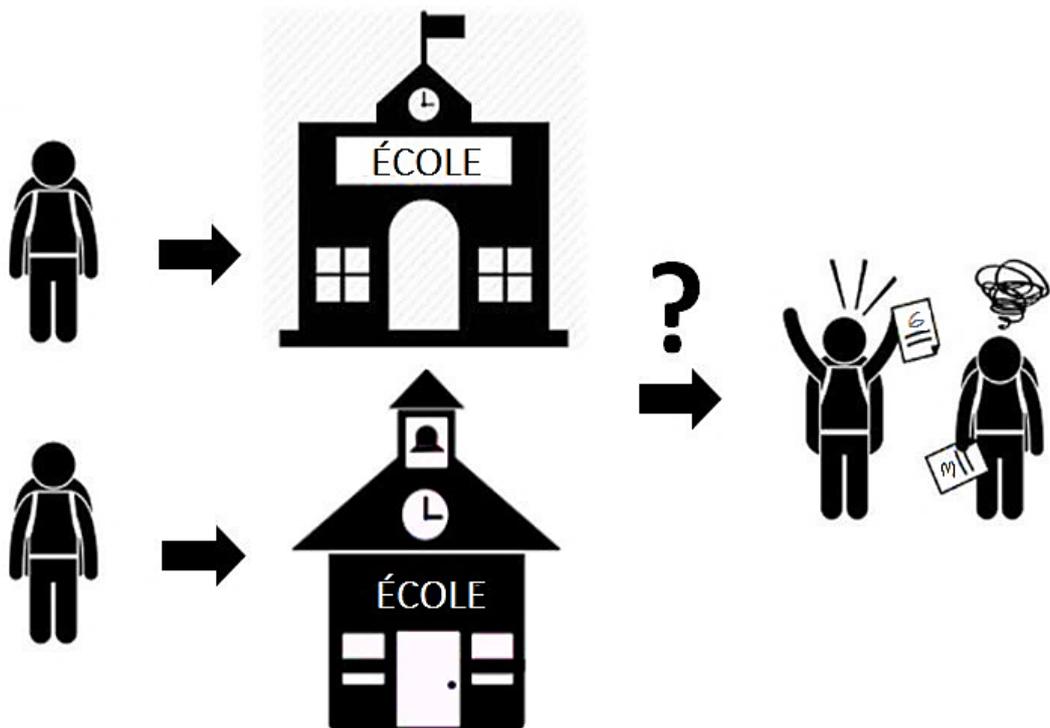


SUPSI

Effet établissements sur la réussite des élèves au Tessin et à Genève

Effetto-istituto sulla riuscita scolastica degli allievi in Ticino e a Ginevra

Franck Petrucci, Alice Ambrosetti, Sandra Fenaroli et Michele Egloff



Proposta di citazione:

Petrucci, F., Ambrosetti, A., Fenaroli, S., & Egloff, M. (2018). *Effet établissements sur la réussite des élèves au Tessin et à Genève*. Genève/Locarno: Service de la recherche en éducation & Centro innovazione e ricerca sui sistemi educativi.

Locarno, 2018

SRED - Service de la recherche en éducation

Quai du Rhône 12, 1205 Genève

sred@etat.ge.ch

CIRSE - Centro Innovazione e Ricerca sui Sistemi Educativi

Piazza San Francesco 19, 6600 Locarno

dfa.cirse@supsi.ch

ISBN 9788885585102

Autori: Franck Petrucci, Alice Ambrosetti, Sandra Fenaroli e Michele Egloff

Revisione: Narain Jagasia

Impaginazione: Elena Camerlo

Remerciements

Les auteurs remercient en premier lieu Daniela Di Mare, anciennement directrice du Service de la recherche en éducation du canton de Genève (SRED), et Dominique Gros, directeur adjoint, pour avoir accueilli positivement ce projet et en avoir assuré le suivi; ils remercient également Martin Benninghoff, actuel directeur du SRED, pour en avoir garanti la réalisation jusqu'à la publication.

Un remerciement tout particulier est adressé à Anne Soussi et Christian Nidegger du SRED et Miriam Salvisberg du *Centro innovazione e ricerca sui sistemi educativi* (CIRSE) pour leur relecture attentive et critique des brouillons du manuscrit et le partage de leurs connaissances approfondies de l'enquête et des données PISA. Merci à Miriam Salvisberg pour avoir pris l'initiative d'organiser une formation interne au CIRSE sur l'analyse multiniveau qui a été l'étincelle à l'origine de la présente recherche.

Les auteurs sont, par ailleurs, reconnaissants envers Odile Le Roy-Zen Ruffinen, François Rastoldo, et Marion Dutrevis (SRED) pour leurs précieux encouragements et pour les réflexions pertinentes avec lesquelles ils ont accompagné le processus d'analyse des données.

Un chaleureux remerciement est aussi adressé à Marina Mikulic et Adriano Varetta du Centre de compétence GAS-GAGI (*Gestione Amministrativa delle Scuole - Gestione Allievi e Gestione Istituti*) du Département de l'éducation, de la culture et du sport (DECS) du Canton du Tessin et à Annick Evrard, Bernard Engel, Youssef Hrizi et Rami Mouad (SRED) qui, avec leur disponibilité habituelle, ont fourni les données cantonales sur les élèves et les enseignants indispensables à la réalisation de ce travail.

Enfin, un grand merci à Narain Jagasia (SRED) pour sa relecture du texte et à Elena Camerlo (CIRSE) qui a réalisé avec beaucoup de soin la mise en page de ce rapport.

Ringraziamenti

Gli autori ringraziano innanzitutto Daniela Di Mare, già direttrice del Service de la recherche en éducation de l'Etat de Genève (SRED), e Dominique Gros, direttore aggiunto, per aver accolto positivamente il progetto e per averne sostenuto lo sviluppo; ringraziano Martin Benninghoff, direttore in carica dello SRED, per averne garantito la realizzazione fino alla pubblicazione.

Un sentito ringraziamento va ad Anne Soussi e Christian Nidegger dello SRED e a Miriam Salvisberg del Centro innovazione e ricerca sui sistemi educativi (CIRSE) per la rilettura attenta e critica delle bozze del manoscritto e per la condivisione delle loro solide conoscenze dell'indagine e dei dati PISA.

A Miriam Salvisberg va inoltre riconosciuta con gratitudine l'iniziativa di proporre una formazione interna al CIRSE sull'analisi multilivello dei dati, iniziativa che è stata la scintilla che dato il via al presente studio.

Gli autori sono altresì riconoscenti a François Rastoldo, Odile Le Roy-Zen Ruffinen e Marion Dutrevis (SRED) per i preziosi incoraggiamenti e per le riflessioni argute con le quali hanno accompagnato l'intero processo di analisi.

Un pensiero di gratitudine va inoltre a Marina Mikulic e Adriano Varetta del Centro di competenza GAS-GAGI (Gestione Amministrativa delle Scuole - Gestione Allievi e Gestione Istituti) del Dipartimento dell'educazione, della cultura e dello sport del Cantone Ticino (DECS) e a Annick Evrard, Bernard Engel, Youssef Hrizi et Rami Mouad (SRED) che, con la consueta disponibilità, hanno fornito i dati cantonali su allievi e docenti, indispensabili per lo svolgimento di questo lavoro.

Infine, un grazie particolare a Narain Jagasia (SRED) per l'attenta revisione del testo e a Elena Camerlo (CIRSE) che ha pazientemente curato l'impaginazione e la formattazione del presente rapporto.

Table des matières / Sommario

L'ESSENTIEL EN BREF	7
L'ESSENZIALE IN BREVE	13
1. INTRODUCTION	17
2. OBJECTIF DE LA RECHERCHE ET METHODOLOGIE	19
2.1 UN DESIGN DE RECHERCHE PARTICULIEREMENT ADAPTE A LA MESURE DE L'EFFET DE COMPOSITION AU TESSIN ET A GENEVE .	20
3. RAPPEL DE QUELQUES ELEMENTS CONTEXTUELS	23
3.1 ORGANISATION DE L'ENSEIGNEMENT SECONDAIRE I DANS LE CANTON DU TESSIN.....	23
3.2 ORGANISATION DE L'ENSEIGNEMENT SECONDAIRE I DANS LE CANTON DE GENEVE	23
3.3 ILLUSTRATION DE LA DIVERSITE DES ETABLISSEMENTS DE L'ENSEIGNEMENT SECONDAIRE I DANS CHAQUE CANTON	24
4. ECHANTILLONNAGE ET DONNEES	27
4.1 TAILLE DES ECHANTILLONS	27
4.2 LES TROIS DOMAINES DE COMPETENCE EVALUES DANS PISA 2012	27
4.3 QUELQUES RESULTATS A PISA POUR LES CANTONS DU TESSIN ET DE GENEVE.....	29
4.3.1 <i>Les mathématiques</i>	29
4.3.2 <i>La lecture et les sciences</i>	30
4.4 LES CARACTERISTIQUES DES ELEVES.....	31
4.4.1 <i>Les caractéristiques sociodémographiques usuelles (d'après l'enquête PISA 2012)</i>	31
4.4.2 <i>Le contrôle du niveau scolaire antérieur / initial (d'après fichiers cantonaux)</i>	31
4.5 LES CARACTERISTIQUES DES ETABLISSEMENTS.....	34
4.5.1 <i>Les caractéristiques sociodémographiques (d'après fichiers cantonaux)</i>	34
4.5.2 <i>Les caractéristiques scolaires (d'après fichiers cantonaux)</i>	36
4.5.3 <i>Les caractéristiques organisationnelles (d'après fichiers cantonaux)</i>	37
4.5.4 <i>Caractéristiques du corps enseignant</i>	39
4.5.5 <i>Climat et gestion de l'établissement (d'après enquête PISA 2012)</i>	41
5. PRESENTATION DU MODELE MULTINIVEAU	45
5.1 QUELQUES RAPPELS METHODOLOGIQUES SUR LE MODELE MULTINIVEAU	45
5.2 SPECIFICATION DES MODELES ET STRATEGIE DE MISE EN ŒUVRE.....	49
<i>Etape 1 : le modèle «vide»</i>	49
<i>Etape 2 : l'introduction des caractéristiques individuelles de l'élève (niveau 1)</i>	50
<i>Etape 3 : la prise en compte des caractéristiques de l'établissement (niveau 2)</i>	50
6. PRESENTATION DES RESULTATS	53
6.1 DES EFFETS ETABLISSEMENT RELATIVEMENT MODESTES, AU TESSIN COMME A GENEVE	53
6.1.1 <i>Des effets probablement un peu surestimés sur la base des données PISA</i>	55
6.2 ANALYSE DES SCORES OBTENUS EN MATHEMATIQUES	56
6.3 LA LECTURE.....	62
6.4 LES SCIENCES.....	68
7. ELEMENTS DE SYNTHESE ET DISCUSSION	75
8. RÉFÉRENCES	83

L'essentiel en bref

Introduction

Depuis la fin des années 50, de nombreux travaux de recherche ont montré que d'importantes inégalités de réussite scolaire sont liées à l'origine sociale et culturelle des élèves.

Par ailleurs, le rôle de facteurs contextuels dans ces inégalités fait également l'objet d'un relatif consensus au sein de la communauté scientifique. De façon plus précise, l'école dans laquelle un élève est scolarisé serait une unité d'analyse pertinente pour expliquer certaines variations d'acquisitions de compétences.

Le courant du *School mix*, qui a émergé à partir des années 80, s'est plus particulièrement attaché à explorer l'effet des caractéristiques propres aux écoles sous l'angle de leur composition et de la manière dont cette dernière influe sur les comportements et acquisitions des élèves.

Selon ce courant, le niveau scolaire des élèves (*academic mix*, ou composition académique) et leur profil sociologique (*social mix*, ou composition sociale) sont les deux dimensions qui affecteraient principalement les progressions académiques des élèves.

Fournir des éléments de réponse à la question de l'effet établissement est un enjeu majeur et on dispose d'une littérature scientifique abondante sur le sujet (Duru-Bellat, 2003; Caldas et Bankston, 1997; Crahay et Monseur, 2006). Toutefois, bon nombre de résultats sont très dépendants des caractéristiques des systèmes scolaires dans lesquels ils ont été obtenus et il n'existe pas encore de véritable consensus sur l'ampleur et la nature de cet effet.

Dans ce contexte, le *Centro innovazione e ricerca sui sistemi educativi* (CIRSE) du canton du Tessin et le Service de la recherche en éducation (SRED) du canton de Genève se sont penchés sur les relations qui existent, dans ces deux cantons, entre l'acquisition de compétences des élèves et le contexte de scolarisation.

Les enjeux d'une telle recherche sont multiples. Tout d'abord, il apparaît nécessaire de constituer un savoir spécifique sur l'effet établissement en Suisse et plus particulièrement au Tessin et à Genève, où la question n'a été que peu traitée à ce jour. Par ailleurs, identifier les facteurs qui permettent à certains établissements d'être plus efficaces que d'autres pourrait potentiellement contribuer à améliorer les apprentissages d'un grand nombre d'élèves et plus particulièrement de ceux issus des milieux les moins favorisés (Bressoux, 1994).

Pour les autorités scolaires, une telle recherche pourrait également s'avérer utile au pilotage des systèmes de formation en alimentant les réflexions autour de l'égalité des chances et de l'équité de traitement.

Enfin, il est également intéressant de s'interroger sur les liens qui existent entre l'effet établissement et l'organisation du système éducatif. Au Tessin et à Genève, l'enseignement secondaire I est organisé de manière différente, le canton italophone ne répartissant pas les élèves en filières, contrairement au canton lémanique majoritairement organisé en regroupements différenciés qui séparent les élèves selon leur niveau académique. Dès lors, l'effet établissement, sa taille et sa nature, diffèrent-ils en fonction de l'organisation ?

Objectif de la recherche

Ces travaux visent à identifier et à qualifier un éventuel effet de l'établissement fréquenté sur l'acquisition des compétences des élèves dans les trois domaines évalués par PISA (à savoir la lecture, les mathématiques et les sciences). En d'autres termes, la question de départ pourrait se résumer par «**dans quelle mesure deux élèves semblables en début d'une période de formation vont-ils évoluer différemment en fonction des caractéristiques des autres élèves de leur école?**» (Dumay et Dupriez, 2009, p. 463).

Dans la lignée des travaux du *School mix*, on postule ici que l'effet établissement est lié à la composition sociale (profil socio-économique des élèves) et/ou académique (niveau scolaire des élèves) de la population scolarisée. Une hypothèse est que la ségrégation d'élèves en difficulté scolaire ou de milieu socio-économique défavorisé crée des dynamiques de groupe qui, toutes choses

égales par ailleurs, se révèlent peu propices à l'apprentissage et génératrices de performances scolaires plus modestes.

Les données utilisées

Les données de base sont issues des échantillons tessinois et genevois de l'enquête PISA 2012. Elles ont été complétées par des informations extraites des bases de données scolaires cantonales.

Les données PISA se composent d'un échantillon de jeunes de 11^e année, c'est-à-dire la dernière année de la scolarité obligatoire (1'081 élèves scolarisés dans 35 établissements pour le Tessin et 920 élèves dans 19 établissements pour Genève). Le contexte de scolarisation repose quant à lui sur les données exhaustives de l'ensemble de la population scolaire de la Scuola Media et du Cycle d'orientation. Ainsi la composition sociale et académique des établissements est mesurée en tenant compte des caractéristiques sociodémographiques et scolaires de tous les élèves et pas seulement de ceux de l'échantillon PISA.

Les analyses effectuées dans cette recherche reposent principalement sur la mise en œuvre de modèles hiérarchiques linéaires (HLM), aussi appelés modèles multiniveaux, qui ont été développés pour traiter de l'influence du contexte sur les individus et qui s'appliquent à des données possédant une structure hiérarchisée avec des niveaux emboîtés les uns dans les autres : ici des élèves (niveau 1 – échantillon PISA) scolarisés dans des établissements (niveau 2 – données PISA et bases de données scolaires cantonales). Dans ces analyses, on considère que les performances d'un élève à PISA dépendent de caractéristiques qui lui sont propres mais aussi de caractéristiques inhérentes à son établissement.

Les variables d'analyse

On cherche à expliquer les scores obtenus dans les différents domaines évalués par PISA. Pour ce faire on dispose de toutes les caractéristiques sociodémographiques usuelles des élèves que sont le genre, l'âge, le statut migratoire, la langue et le statut socioéconomique et culturel. Ces données ont été recueillies lors de l'enquête PISA au moyen d'un questionnaire rempli par les élèves après la passation des tests dans les trois disciplines (mathématiques, lecture et sciences). Le niveau initial des élèves est quant à lui extrait des données cantonales. Il est mesuré à la fois par la moyenne annuelle dans la discipline cible à la fin de l'année scolaire précédente (à savoir mathématiques, sciences/biologie ou italien/français), et par la filière d'étude pour Genève (Regroupement A ou B) ou le profil pour le Tessin (niveaux A ou B en mathématiques et allemand). Il importe de noter que la prise en compte du niveau initial des élèves est incontournable pour déterminer l'existence d'un effet établissement et disposer d'une estimation fiable et robuste de sa taille. En effet, l'idée est de s'assurer que deux élèves sont semblables en début d'une période de formation, en particulier du point de vue de leur niveau scolaire, pour être en mesure de déterminer que c'est bien l'exposition à un contexte de scolarisation différent qui génère potentiellement, en fin de période de formation, des différences de résultats. Willms (1992) et Lauder et Hughes (1990) ont d'ailleurs émis l'hypothèse que l'effet de composition mis en évidence par d'autres chercheurs pourrait, en réalité, être un artefact qui découle d'un mauvais contrôle de variables caractérisant les élèves, et tout particulièrement d'une prise en compte insuffisante de leur niveau initial.

Les variables qui caractérisent les établissements ont été construites grâce aux données cantonales (caractéristiques de la population scolaire et du corps enseignant) et au questionnaire école de PISA auquel ont répondu les chefs d'établissement (informations relatives aux ressources, à la gestion et au climat d'établissement).

Concernant la population scolaire, la composition sociale et la composition académique de l'établissement, qui sont les deux variables fondamentales de cette étude, ont été définies de la manière suivante :

- la composition sociale a été caractérisée à partir d'un indice de déplacement (Duru-Bellat, Danner, Le Bastard-Landrier et Piquée, 2004) qui a permis de classer les établissements comme étant plutôt favorisés, mixtes ou défavorisés. Plus précisément, cet indice repose sur le calcul du pourcentage d'élèves qu'il faudrait déplacer pour que chaque établissement possède une structure sociale comparable à celle qui prévaut dans l'ensemble de la population (situation d'équi-répartition);

- la composition académique a été mesurée en utilisant différents indices, notamment la proportion d'élèves scolarisés dans la filière à exigences élevées (Genève) ou suivant les niveaux les plus exigeants en mathématiques et en allemand (Tessin), ainsi que la proportion d'élèves non promus dans chaque établissement.

Enfin, les autres variables retenues pour décrire l'établissement peuvent être divisées en différentes catégories : les variables relatives au corps enseignant (ancienneté, âge, genre, etc.), celles organisationnelles (taille de l'établissement et des classes), celles relatives à la population scolarisée (proportion de filles, proportion d'élèves migrants, etc.) et celles relatives au climat et à la gestion (indices de pénurie d'enseignants qualifiés, de participation des enseignants à la gestion de l'établissement, etc.).

Les résultats

La présentation des résultats reprend les différentes étapes de la démarche analytique multiniveau. Existe-t-il un effet établissement sur l'apprentissage des élèves, c'est-à-dire sur les scores obtenus à PISA ? Dans quelle mesure les caractéristiques individuelles des élèves (niveau 1) influencent-elles ces scores ? Enfin, quelle est la nature de l'effet établissement, autrement dit parvient-on à le caractériser en ajoutant dans le modèle les variables du deuxième niveau qui décrivent les établissements scolaires ?

La première étape permet d'établir que, dans les deux cantons, il y a effectivement un effet de l'établissement sur les compétences des élèves. Autrement dit, le « même » élève est susceptible d'obtenir un score différent dans les domaines PISA selon l'établissement fréquenté, ce qui signifie qu'il y a des établissements diversement efficaces. Toutefois la taille de cet effet est relativement modeste puisque seuls 5% à 13% de la variance des scores des élèves lui sont imputables.

Selon la littérature sur le sujet, ces résultats se situent dans la norme. Duru-Bellat et al. (2004) ont relevé que dans bon nombre de travaux, la part de la variance des résultats des élèves due à des différences entre écoles oscille entre 8% et 15%, la plupart des estimations se situant plutôt dans le bas de cette fourchette.

Différentes explications de la faiblesse de l'effet établissement observé dans les deux cantons concernés peuvent être avancées. En premier lieu, Bressoux (1994) relève que, dans les pays industrialisés, il y a en général plus d'hétérogénéité à l'intérieur des écoles qu'entre elles, ce qui signifie que les différences entre élèves expliquent l'essentiel des différences d'apprentissage. De plus, Crahay et Monseur (2008) avancent l'hypothèse que l'effet établissement est, en réalité, tributaire des caractéristiques des systèmes éducatifs. Au Tessin et à Genève, les possibilités de voir émerger un effet contextuel relativement important sont vraisemblablement assez réduites car la Scuola Media et le Cycle d'orientation fonctionnent de manière relativement homogène. Bien que les établissements disposent d'un certain degré d'autonomie, la plupart des principaux facteurs qui régissent leur fonctionnement (plans d'études, niveau de formation des enseignants, taille des classes, ressources, carte scolaire, etc.) sont en réalité fixés par voie réglementaire et s'appliquent de façon identique sur l'ensemble du territoire cantonal, ce qui atténue la variabilité entre établissements.

La différence dans les scores obtenus en mathématiques, lecture et sciences par les élèves est donc surtout le fruit de différences individuelles et non pas de différences liées aux établissements. Pour autant, au Tessin comme à Genève, on ne réussit pas de la même manière dans toutes les écoles et ce quand bien même on possède les mêmes caractéristiques individuelles. L'effet de ces variables telles que le genre, l'âge, l'origine sociale ou le statut migratoire n'est d'ailleurs pas détaillé ici car l'objet principal de cette recherche était avant tout de mesurer l'effet du contexte. Remarquons simplement que l'on retrouve les effets usuels des caractéristiques sociodémographiques des élèves sur leurs performances maintes fois décrits dans la littérature (Salvisberg et Zampieri, 2015; Nidegger (éd.), 2014; Consortium PISA.ch, 2011).

À ce sujet on relèvera encore que le poids de ces caractéristiques dans l'explication des différences de résultats est plus faible au Tessin qu'à Genève. Ceci signifie, par exemple, que l'effet du genre ou du statut migratoire a moins d'impact sur le score obtenu à PISA au Tessin qu'à Genève. Se pose alors la question d'une plus grande équité du système éducatif tessinois et de son origine qui, si l'on s'en remet aux hypothèses formulées par certains auteurs (Felouzis et Charmillot, 2017), serait liée à une organisation intégrée tendancielle plus équitable que celle articulée autour de filières prévalant à Genève.

La troisième étape de l'analyse concerne la nature de l'effet établissement et permet de se pencher sur l'explication des différences entre écoles. L'hypothèse principale que nous souhaitons tester était celle d'un effet de la composition sociale et/ou académique. Il s'avère que l'effet établissement mis en évidence au Tessin et à Genève n'est pas de cette nature. Quel que soit le domaine de compétences considéré, il n'a en effet jamais été possible de montrer que la composition sociale ou académique d'un établissement de la Scuola Media ou du Cycle d'orientation expliquait une partie des différences de performances des élèves. Autrement dit, si au Tessin comme à Genève, certains établissements sont plus efficaces que d'autres et, qu'à caractéristiques individuelles identiques (notamment niveau initial), on ne réussit pas de la même manière en fonction de l'école dans laquelle on est scolarisé, ces variations ne sont pas dues à la tonalité sociale ou académique des établissements. Ceci constitue d'ailleurs un élément positif dans l'évaluation du fonctionnement des systèmes éducatifs genevois et tessinois. En effet, aucun de ces deux systèmes ne semble renforcer la (re-)production d'inégalités sociales de réussite scolaire en mettant en œuvre des modes de regroupement des élèves au sein des établissements qui génèrent une ségrégation des publics telle qu'elle impacte les performances individuelles.

En l'absence d'effet de composition sociale et/ou académique, la nature de l'effet établissement reste encore à déterminer. Pourquoi certains établissements sont-ils plus efficaces que d'autres au Tessin et à Genève ? Quelques facteurs expliquent, en partie, les différences de résultats imputables à l'établissement mais ils varient à la fois d'un domaine de compétences à l'autre et d'un canton à l'autre, ce qui ne nous permet pas de dégager des conclusions univoques et généralisables.

Au Tessin, dans les trois disciplines du test PISA, on identifie un effet de la taille moyenne des classes. Ainsi, toutes choses égales par ailleurs, les élèves obtiennent de meilleurs scores dans les établissements où la taille des classes est, en moyenne, inférieure à ce qui existe dans l'ensemble des établissements du canton. Ce résultat pourrait certes s'interpréter comme un impact positif de la réduction de la taille des classes sur les acquisitions des élèves (Piketty et Valdenaire, 2006; Meuret, 2001) mais on ne peut toutefois exclure qu'il s'agisse aussi d'un effet d'une toute autre nature, dans la mesure où tous les établissements concernés appartiennent à une même zone dans laquelle il se peut que les conditions d'enseignement soient particulières, ces dernières étant d'ailleurs plus le fruit d'une situation géographique atypique que le résultat d'un véritable choix de politique éducative.

Toujours pour le Tessin, dans le domaine des sciences, il apparaît que dans les écoles où le corps enseignant est plus qualifié (proportion plus importante d'enseignants pleinement qualifiés disposant de l'habilitation à enseigner), les élèves obtiennent de meilleurs résultats. Ce résultat est de même nature que celui mis en évidence pour Genève en mathématiques où, à caractéristiques individuelles identiques, les élèves obtiennent de moins bonnes performances lorsqu'ils sont scolarisés dans un établissement où les enseignants de mathématiques ont une ancienneté plus faible que celle qui existe en moyenne dans l'ensemble du canton.

Ces deux résultats, qui semblent montrer que le niveau de qualification et l'expérience des enseignants sont des caractéristiques assez fortement corrélées aux compétences des élèves, sont assez conformes à ce que l'on trouve dans la littérature.

Pour le domaine de la lecture, on peut aussi mentionner qu'au Tessin les élèves tendent, toutes choses égales par ailleurs, à obtenir des résultats plus faibles lorsqu'ils sont scolarisés dans des établissements où le taux de non-promotion est supérieur à la norme cantonale. À Genève, bien que l'on parvienne aussi à démontrer que les établissements sont diversement efficaces dans le domaine de la lecture, on ne parvient en revanche pas à identifier des caractéristiques propres aux écoles qui puissent expliquer les différences de résultats.

En ce qui concerne les sciences, pour le canton de Genève il a été possible d'identifier deux caractéristiques des établissements ayant un effet sur les performances des élèves. Premièrement, les élèves réussissent mieux, toutes choses égales par ailleurs, lorsque les enseignants participent davantage à la gestion de l'établissement. Deuxièmement, à caractéristiques individuelles identiques, les élèves scolarisés dans des établissements de taille supérieure à la norme (une centaine d'élèves de plus que la moyenne cantonale) obtiennent de moins bonnes performances. Bien que la question du lien entre la taille des établissements et la réussite des élèves doive encore être considérée comme ouverte, on pourrait penser qu'il s'agit là d'une illustration de ce que relève Afsa (2014), sur la base des revues de la littérature produites par Slate et Jones (2005) et Leithwood et Jantzi (2009), à savoir que «la taille a un effet négatif sur la performance scolaire de l'établissement». Cependant, dans le cadre de notre étude, on ne peut pas exclure l'hypothèse que cet effet observé à Genève soit en réalité d'une toute autre nature. Nous avons en fait constaté que les directeurs des établissements

de plus grande taille déclarent avoir davantage de problèmes de ressources que leurs homologues d'établissements de taille plus modeste (et notamment des difficultés de recrutement de personnel qualifié et un nombre insuffisant d'ordinateurs pour les élèves).

Conclusions et développements possibles

La question de la nature de l'effet établissement, que la présente étude contribue à baliser en partie, reste encore largement ouverte au Tessin et à Genève. Si les résultats incitent à rejeter l'hypothèse d'un effet de composition sociale et académique sur les performances des élèves, ils ne permettent pas en revanche de dégager d'autres pistes qui auraient une portée relativement générale. Par ailleurs, certains des résultats laissent penser que d'autres unités d'analyse, telles que la classe, sont au moins aussi pertinentes que l'établissement pour comprendre les variations d'acquisitions de compétences. Il serait donc utile, à l'avenir, de conduire des recherches qui considèrent simultanément tous ces niveaux (élèves, classes, établissements) afin de mieux comprendre comment ils s'articulent entre eux. Il s'agirait, par exemple, d'analyser comment ce qui se passe au niveau de l'établissement en termes d'organisation, de gestion ou de leadership est susceptible d'affecter les processus qui prennent place au sein des classes. En matière d'exploration de la problématique des effets de contexte, la présente recherche constitue, au Tessin et à Genève, un premier pas qui apporte quelques éléments de réponse et permet le développement de questionnements futurs.

L'essenziale in breve

Introduzione

Dalla fine degli anni 1950, numerose ricerche hanno dimostrato il nesso tra la riuscita scolastica e l'origine sociale e culturale degli allievi.

Inoltre, la comunità scientifica è relativamente concorde nel riconoscere che sul rendimento possono influire anche fattori contestuali. In altre parole, la scuola frequentata da un allievo sarebbe un'unità d'analisi pertinente per spiegare alcune differenze nell'acquisizione delle competenze.

La corrente del cosiddetto *School mix*, diffusasi a partire dagli anni 1980, si è concentrata in particolare sulla composizione della popolazione scolastica degli istituti e sull'effetto che tale composizione esercita sui comportamenti degli allievi e sulla loro acquisizione delle competenze.

Secondo questa corrente, il livello scolastico degli allievi (*academic mix*, o composizione accademica) e il loro profilo sociologico (*social mix*, o composizione sociale) sono le due dimensioni che condizionerebbero principalmente i progressi accademici degli allievi.

Fornire elementi per rispondere alla questione dell'effetto-istituto è particolarmente importante. Lo attesta la vasta letteratura scientifica sull'argomento (Duru-Bellat, 2003 ; Caldas & Bankston, 1997 ; Crahay & Monseur, 2006). Numerosi risultati dipendono tuttavia dalle caratteristiche dei sistemi scolastici nei quali sono stati ottenuti e non esiste ancora un vero e proprio consenso sulla portata e sulla natura di questo effetto.

A tale proposito, il Centro innovazione e ricerca sui sistemi educativi (CIRSE) del Cantone Ticino e il Service de la recherche en éducation (SRED) del Cantone di Ginevra hanno esaminato le relazioni esistenti nei rispettivi Cantoni tra l'acquisizione delle competenze degli allievi e il contesto di scolarizzazione.

Le implicazioni di una simile ricerca sono molteplici. Innanzitutto è necessario costituire un sapere specifico sull'effetto-istituto in Svizzera e in particolare in Ticino e a Ginevra, dove questa tematica è stata poco esplorata finora. In più, il fatto di individuare i fattori che consentono ad alcuni istituti di essere più efficaci rispetto ad altri potrebbe contribuire a migliorare l'apprendimento di numerosi allievi, in particolare di quelli provenienti da contesti più svantaggiati (Bressoux, 1994).

La presente ricerca potrebbe risultare utile anche alle autorità scolastiche, poiché potrebbe fungere da spunto nell'impostazione dei sistemi di formazione e alimentare le riflessioni in materia di pari opportunità e pari trattamento.

Infine, è interessante anche esaminare il nesso tra l'effetto-istituto e l'organizzazione del sistema educativo. Il Cantone Ticino e il Cantone di Ginevra hanno impostato l'insegnamento del grado secondario I in modo diverso: in Ticino gli allievi non sono suddivisi in filiere, come succede invece a Ginevra, dove il sistema è prevalentemente strutturato per gruppi diversificati nei quali gli allievi sono suddivisi in base al livello accademico. Ci si chiede pertanto se l'effetto-istituto, le sue dimensioni e la sua natura variano a seconda dell'organizzazione.

Obiettivo della ricerca

Il presente lavoro di ricerca mira a individuare e a spiegare un eventuale effetto dell'istituto frequentato sull'acquisizione delle competenze da parte degli allievi nei tre ambiti considerati in PISA (lettura, matematica e scienze). In altri termini, la domanda di fondo potrebbe essere riassunta così: **«In che misura due allievi simili all'inizio di un periodo di formazione evolveranno diversamente a seconda delle caratteristiche degli altri allievi della stessa scuola?»** (Dumay & Dupriez, 2009, p. 463).

Sulla scia dei lavori dello *School mix*, si suppone che l'effetto-istituto sia legato alla composizione sociale (profilo socio-economico degli allievi) e/o accademico (livello scolastico degli allievi) della popolazione scolastica. Un'ipotesi è che la segregazione di allievi con difficoltà scolastiche o provenienti da un contesto socioeconomico svantaggiato crei dinamiche di gruppo che, a parità di tutte le altre condizioni, risultano poco propizie all'apprendimento e condizionano il rendimento scolastico.

I dati utilizzati

I dati di base provengono dai campioni ticinesi e ginevrini dell'indagine PISA 2012 e sono stati completati con informazioni estratte dalle banche dati scolastiche cantonali.

I dati PISA sono composti da un campione di allievi dell'11° anno, l'ultimo della scuola dell'obbligo (1081 allievi di 35 istituti in Ticino e 920 allievi di 19 istituti a Ginevra). La descrizione del contesto di scolarizzazione si basa sui dati esaustivi dell'intera popolazione scolastica della Scuola Media e del Cycle d'orientation. Ciò significa che la composizione sociale e accademica degli istituti è misurata tenendo conto delle caratteristiche sociodemografiche e scolastiche di tutti gli allievi e non solo di quelli del campione PISA.

Le analisi effettuate nell'ambito della presente ricerca poggiano prevalentemente sulla metodologia dei modelli lineari gerarchici (HLM), detti anche modelli multilivello, sviluppati per stabilire l'influenza del contesto sugli individui. Questi modelli si applicano a dati che possiedono una struttura gerarchica: nello specifico, degli allievi (livello 1 – campione PISA) sono raggruppati secondo gli istituti che frequentano (livello 2 – dati PISA e banche dati scolastiche cantonali). In queste analisi si parte dal principio che i risultati ottenuti da un allievo in PISA dipendano da caratteristiche proprie alla sua persona ma anche da caratteristiche legate all'istituto che frequenta.

Le variabili d'analisi

Si tenta di spiegare i risultati ottenuti nei tre ambiti indagati in PISA (matematica, lettura e scienze) con l'ausilio dei dati raccolti mediante un questionario compilato dagli allievi al termine dei test e relativi a tutte le caratteristiche sociodemografiche abituali degli allievi, ovvero genere, età, statuto migratorio, lingua e statuto socioeconomico e culturale. Va sottolineato che per determinare l'esistenza di un effetto-istituto e ottenere una stima attendibile e robusta della sua portata è fondamentale tenere conto del livello iniziale degli allievi. Infatti, per poter determinare se l'esposizione a un contesto di scolarizzazione diverso genera potenzialmente risultati diversi al termine della formazione occorre assicurarsi che gli allievi osservati abbiano un livello scolastico simile all'inizio della formazione. Nel presente studio, il livello iniziale di competenza è stimato con l'ausilio dei dati cantonali: è misurato sia con il valore della nota ricevuta alla fine dell'anno scolastico precedente nella disciplina considerata (matematica, scienze/biologia o italiano/francese) sia in base alla filiera di studio (raggruppamento A o B) per Ginevra o ai livelli attitudinali o di base in matematica e tedesco per il Ticino. Willms (1992), Lauder e Hughes (1990) hanno peraltro ipotizzato che l'effetto della composizione evidenziato da altri ricercatori potrebbe in realtà essere un artefatto derivante da un controllo lacunoso delle variabili che caratterizzano gli allievi e in particolare da una considerazione insufficiente del loro livello iniziale.

Le variabili che caratterizzano gli istituti sono state costruite grazie ai dati cantonali (caratteristiche della popolazione scolastica e del corpo insegnante) e al questionario PISA per le scuole somministrato ai direttori degli istituti scolastici (informazioni sulle risorse, sulla gestione e sul clima nell'istituto).

Riguardo alla popolazione scolastica, le due variabili fondamentali della presente ricerca, ovvero la composizione sociale e la composizione accademica dell'istituto sono state definite come segue:

- la composizione sociale è stata determinata sulla scorta di un «indice di spostamento» (Duru-Bellat, Danner, Le Bastard-Landrier & Piquée, 2004) che ha permesso di ripartire gli istituti in base a tre categorie: piuttosto favoriti, misti o piuttosto svantaggiati. Più precisamente, questo indice si basa sul calcolo della percentuale di allievi che bisognerebbe spostare affinché ogni istituto disponga di una struttura sociale paragonabile a quella che prevale nell'insieme della popolazione (situazione di equi-ripartizione).
- la composizione accademica è stata determinata con l'ausilio di vari indici, segnatamente la proporzione di allievi che seguono la filiera a esigenze elevate (Ginevra) o i corsi attitudinali in matematica e tedesco (Ticino) nonché la proporzione di allievi bocciati in ogni istituto.

Infine, le altre variabili considerate per descrivere l'istituto possono essere ripartite in varie categorie: le variabili relative al corpo insegnante (anzianità di servizio, età, genere, ecc.), le variabili organizzative (dimensione dell'istituto e delle classi), le variabili relative alla popolazione scolastica (proporzione di ragazze, di allievi migranti, ecc.) e le variabili relative al clima e alla gestione (indice di penuria di insegnanti qualificati, di partecipazione degli insegnanti alla gestione dell'istituto, ecc.).

I risultati

La presentazione dei risultati riprende le varie tappe della procedura analitica multilivello. Esiste un effetto-istituto sull'apprendimento degli allievi, ovvero sui risultati ottenuti in PISA? In che misura le caratteristiche individuali degli allievi (livello 1) influenzano tali risultati? E infine, qual è la natura dell'effetto-istituto o, detto altrimenti, è possibile caratterizzarlo integrando nel modello le variabili di secondo livello che descrivono gli istituti scolastici?

La prima tappa consente di stabilire che, nei due Cantoni, l'istituto ha un effetto sulle competenze degli allievi. In altre parole, uno «stesso» allievo potrebbe ottenere un risultato diverso negli ambiti PISA a seconda dell'istituto frequentato, il che significa che l'efficacia varia da un istituto all'altro. Tuttavia, avendo riscontrato che a questo effetto si può imputare solo tra il 5 % e il 13 % della varianza dei risultati, si può affermare che la sua portata è relativamente modesta.

Secondo la letteratura specializzata, questi risultati sono nella norma. Duru-Bellat et al. (2004) hanno osservato che in numerosi lavori la varianza dei risultati degli allievi dovuta a differenze tra scuole oscilla tra l'8 % e il 15 %, con una prevalenza di stime nella parte inferiore di questa forbice.

La debolezza dell'effetto-istituto osservata nei due Cantoni considerati può essere ricondotta a vari fattori. In primo luogo, Bressoux (1994) osserva che nei Paesi industrializzati vi è in generale una maggiore eterogeneità in seno a una stessa scuola che tra una scuola e l'altra, il che significa che le differenze di apprendimento vanno ricondotte essenzialmente alle differenze tra allievi. Dal canto loro, Crahay e Monseur (2008) ipotizzano che l'effetto-istituto dipenda in realtà dalle caratteristiche dei sistemi educativi. In Ticino e a Ginevra le possibilità di veder emergere un effetto contestuale relativamente importante appaiono piuttosto ridotte in quanto sia la Scuola Media sia il Cycle d'orientation funzionano in modo relativamente omogeneo. Benché gli istituti dispongano di un certo grado di autonomia, la maggior parte dei principali fattori che ne determinano il funzionamento (programmi di studio, livello di formazione degli insegnanti, numero di allievi per classe, risorse, carta della scuola, ecc.) sono stabiliti dal regolamento e si applicano in modo identico su tutto il territorio cantonale, il che attenua le variazioni tra istituti.

I diversi risultati ottenuti dagli allievi nei test di matematica, lettura e scienze in PISA vanno quindi ricondotti prevalentemente a differenze individuali e non a differenze legate agli istituti. Eppure, sia in Ticino che a Ginevra, gli allievi non ottengono esattamente gli stessi risultati in tutte le scuole, anche quando dispongono delle stesse caratteristiche individuali. L'effetto di queste variabili (genere, età, origine sociale e statuto migratorio) non viene peraltro esposto nei dettagli nel presente rapporto, in quanto l'obiettivo principale della ricerca è di misurare l'effetto esercitato dal contesto. Ci limitiamo a osservare che gli effetti abituali delle caratteristiche sociodemografiche degli allievi sul loro rendimento sono descritti con frequenza nella letteratura scientifica sull'argomento (Salvisberg & Zampieri, 2015; Nidegger (ed.), 2014; Consorzio PISA.ch, 2011).

A tale proposito, si osserva che il peso di queste caratteristiche nello spiegare le differenze dei risultati è più basso in Ticino che a Ginevra. Ciò significa, ad esempio, che in Ticino l'effetto del genere o dello statuto migratorio ha un impatto inferiore sul risultato ottenuto in PISA rispetto a Ginevra. Si potrebbe quindi pensare che il sistema scolastico ticinese sia più equo, come ipotizzato da alcuni autori (Felouzis & Charmillot, 2017), visto che poggerebbe su un'organizzazione maggiormente integrata rispetto all'organizzazione per filiere più diffusa a Ginevra.

La terza tappa dell'analisi riguarda la natura dell'effetto-istituto e consente di soffermarsi sulla spiegazione delle differenze tra scuole. L'ipotesi principale che ci premeva esaminare era quella di un effetto della composizione sociale e/o accademica. Risulta però che l'effetto-istituto rilevato in Ticino e a Ginevra non è di questa natura. A prescindere dall'ambito di competenza esaminato, non è mai stato infatti possibile dimostrare se una parte delle differenze riscontrate nel rendimento degli allievi era attribuibile alla composizione sociale o accademica di una sede di Scuola Media o di un istituto del Cycle d'orientation. In altre parole, se tanto in Ticino quanto a Ginevra alcuni istituti sono più efficaci di altri e se a parità di caratteristiche individuali (in particolare inizialmente) la riuscita varia a seconda della scuola frequentata, queste variazioni non sono dovute all'impronta sociale o accademica degli istituti. Questo costituisce peraltro un elemento positivo nella valutazione del funzionamento dei sistemi educativi ginevrino e ticinese. In effetti, né l'uno né l'altro sistema sembra rafforzare la (ri)produzione di disparità sociali nella riuscita scolastica attraverso forme di raggruppamento degli allievi in seno agli istituti, che generano una segregazione tale da ripercuotersi sul rendimento individuale.

In assenza di effetti di composizione sociale e/o accademica, resta ancora da stabilire la natura dell'effetto-istituto. Perché alcuni istituti sono più efficaci di altri sia in Ticino che a Ginevra? Determinati fattori spiegano almeno in parte le differenze tra i risultati imputabili all'istituto, ma variano sia da un ambito di competenza all'altro sia da un Cantone all'altro, il che non ci consente di formulare conclusioni univoche e generalizzabili.

In Ticino, si osserva nelle tre discipline PISA un effetto della grandezza media delle classi. Infatti, a parità di tutte le altre caratteristiche, gli allievi ottengono risultati migliori nelle sedi in cui la grandezza delle classi è mediamente inferiore a quella dell'insieme degli istituti del Cantone. Questo risultato potrebbe sicuramente essere interpretato come un impatto positivo della riduzione della grandezza delle classi sulle acquisizioni degli allievi (Piketty & Valdenaire, 2006 ; Meuret, 2001), ma non si può neppure escludere che si tratti di un effetto di tutt'altra natura in quanto tutti gli istituti considerati appartengono a una stessa zona nella quale è possibile che le condizioni d'insegnamento siano particolari e derivino più da una situazione geografica atipica che da una vera e propria scelta in termini di politica dell'istruzione.

Sempre in Ticino, nell'ambito delle scienze, emerge che gli allievi ottengono risultati migliori nelle scuole in cui il corpo insegnante è più qualificato (proporzione superiore di insegnanti qualificati titolari dell'abilitazione all'insegnamento). Questo risultato è della stessa natura di quello evidenziato in matematica a Ginevra, dove, a caratteristiche individuali identiche, ottengono punteggi inferiori gli allievi di istituti con insegnanti di matematica con una minore anzianità di servizio rispetto a quella della media cantonale.

Questi due risultati, che sembrano dimostrare una correlazione piuttosto marcata tra il livello di qualifiche e di esperienza degli insegnanti e le competenze degli allievi, sono in linea con quanto riportato dalla letteratura scientifica.

Riguardo all'ambito della lettura si può altresì menzionare che in Ticino gli allievi tendono, a parità delle altre caratteristiche, a ottenere risultati inferiori se frequentano istituti con un tasso di bocciature superiore alla media cantonale. A Ginevra, pur potendo dimostrare la diversa efficacia degli istituti nell'ambito della lettura, non si riesce a individuare caratteristiche proprie alle scuole che possano spiegare le differenze dei risultati.

Per quanto riguarda le scienze, per il Cantone di Ginevra è stato possibile individuare due caratteristiche degli istituti con un effetto sul rendimento degli allievi. Primo: gli allievi riescono meglio, a parità di tutte le altre caratteristiche, quando gli insegnanti partecipano maggiormente alla gestione dell'istituto. Secondo: a parità di caratteristiche individuali, gli allievi che frequentano istituti di dimensioni superiori alla media (un centinaio di allievi in più della media cantonale) ottengono risultati inferiori. Benché il nesso tra la dimensione dell'istituto e la riuscita degli allievi debba ancora essere considerato aperto, si potrebbe pensare che si tratti dell'illustrazione di quanto rilevato da Afsa (2014) in base alle analisi della letteratura effettuate da Slate e Jones (2005) e Leithwood et Jantzi (2009), cioè che «la dimensione ha un effetto negativo sui risultati scolastici dell'istituto». Tuttavia, nell'ambito del presente studio, non si può escludere che questo effetto osservato a Ginevra sia in realtà di tutt'altra natura. Abbiamo constatato infatti che i direttori degli istituti più grandi dichiarano di avere maggiori problemi di risorse rispetto ai loro omologhi di sedi più piccole (e segnatamente difficoltà di reclutamento di personale qualificato oltre a un numero insufficiente di computer per gli allievi).

Conclusioni e possibili sviluppi

La questione della natura dell'effetto-istituto, che la presente ricerca contribuisce in parte a delineare, resta ancora ampiamente aperta sia in Ticino che a Ginevra. Se i risultati invitano a respingere l'ipotesi di un effetto della composizione sociale e accademica sulle prestazioni degli allievi, non permettono nemmeno di individuare altre piste che avrebbero una portata relativamente generale. Peraltro alcuni risultati fanno pensare che altre unità di analisi, come la classe, siano almeno altrettanto pertinenti dell'istituto per comprendere le variazioni nell'acquisizione delle competenze. Sarebbe quindi utile, in futuro, svolgere ricerche che considerino parallelamente tutti questi livelli (allievi, classi, istituti) per meglio capire come si articolano tra loro. Si tratterebbe per esempio di analizzare in che modo ciò che accade in ogni istituto in termini di organizzazione, di gestione o di leadership sia in grado di incidere sui processi che si creano nelle classi. In materia di esplorazione della problematica degli effetti contestuali, la presente ricerca costituisce, sia in Ticino che a Ginevra, un primo passo che apporta alcuni elementi di risposta e consente di sviluppare interrogativi futuri.

1. Introduction

Depuis la fin des années 50, de nombreux travaux de recherche ont montré que d'importantes inégalités de réussite scolaire sont liées à l'origine sociale des élèves. Par ailleurs, l'importance de facteurs contextuels dans l'explication de ces inégalités fait l'objet d'un relatif consensus au sein de la communauté scientifique.

De façon plus précise, l'école dans laquelle un élève est scolarisé serait une des unités d'analyse pertinentes des variations d'acquisitions de compétences. Ainsi, d'après Bressoux (1994), la recherche sur les effets écoles qui a émergé dès les années 60 dans les pays anglo-saxons, a permis de démontrer que «tout ne se joue pas dans le milieu familial et que l'École joue un rôle autonome sur les acquis des élèves» (p. 127).

Comme le relevait Dumay (2004), l'établissement scolaire serait, en lui-même, une source d'inégalité des acquis. Ce constat se retrouve également dans les résultats d'enquêtes à large échelle comme PISA. En effet, sur la base d'une analyse des données 2006, Monseur et Crahay (2008) parvenaient eux aussi à la conclusion qu'au-delà des caractéristiques individuelles, sur lesquelles les autorités scolaires ont souvent peu de prise, il existe bien des variables contextuelles qui affectent l'importance des inégalités sociales de résultats. En d'autres termes, l'école elle-même participerait à la production d'inégalités sociales de réussite scolaire en définissant notamment le mode de regroupement des élèves au sein des établissements, en générant une certaine ségrégation des publics.

Ce résultat semble, par ailleurs, assez pérenne puisque plus récemment, l'OCDE a indiqué dans l'analyse de l'édition PISA 2012 (OCDE, 2014), que les résultats scolaires varient fortement entre les élèves de 15 ans et que cette variation est due à la fois à leurs caractéristiques individuelles et, pour une part non négligeable, aux caractéristiques de leur établissement : «il ressort des chiffres de tous les pays et économies qui ont participé à l'enquête PISA en 2012 que 23% de la variation de la performance s'observe entre les systèmes d'éducation, 31% entre les établissements, et 46%, entre les élèves» (p. 29).

Si l'existence d'un effet établissement sur les résultats des élèves semble donc désormais acquise, la taille et la nature de cet effet restent, en revanche, des sujets d'exploration qui donnent lieu à de multiples pistes de recherche. Le courant du *school mix*, qui a émergé à partir des années 80, s'est plus particulièrement attaché à explorer l'effet des caractéristiques propres aux écoles sous l'angle de leur composition et sur la manière dont cette dernière influe sur les comportements et acquisitions des élèves. L'idée générale de ces travaux est, selon Duru-Bellat (2003), conforme à un postulat classique de la sociologie qui indique que l'environnement influe sur les conduites individuelles, au-delà des caractéristiques personnelles des acteurs. Les apprentissages ont un caractère contextualisé et, parmi les paramètres du contexte scolaire, les caractéristiques du public qui fréquente l'école jouent un rôle déterminant. Ce sont plus particulièrement le niveau scolaire des élèves (*academic mix*) et leur profil sociologique (*social mix*) qui affecteraient les progressions académiques en modifiant la quantité et la qualité de l'instruction délivrée et les attitudes et comportements des élèves.

Selon Duru-Bellat (2003), qui s'appuie sur des travaux de Caldas et Bankston (1997), «une part de l'avantage dont jouissent les enfants de milieu favorisé ne passe pas par un 'héritage culturel' mais par l'accès à des contextes plus formateurs. La réciproque est vraie pour les élèves de milieu défavorisé dont une part du 'handicap' scolaire s'explique par le fait que leurs condisciples sont le plus souvent majoritairement aussi de milieu défavorisé» (p. 200).

La politique de regroupement des élèves au sein des établissements, qui façonne le contexte de scolarisation, constitue donc un enjeu majeur pour les autorités scolaires dans la mesure où elle a des répercussions sur l'équité et l'efficacité des systèmes éducatifs. Ainsi Monseur et Crahay (2008) rappellent que «plus le fonctionnement d'un système d'enseignement engendre des phénomènes d'agrégation sociale ou académique, plus il amplifie les différences entre élèves faibles et élèves forts et plus il amplifie ainsi les inégalités sociales devant l'école» (p. 60).

D'une manière plus générale, la question de l'importance du contexte sur le devenir scolaire est sous-jacente à bon nombre de préoccupations de politique éducative et justifie que la recherche s'y attarde (Duru-Bellat, Danner, Le Bastard-Landrier et Piquée, 2004). Ainsi, Bressoux (1994) rappelle que «si l'on peut déceler des écoles plus efficaces que d'autres ainsi que les facteurs qui y sont associés, alors l'école n'est pas que le simple révélateur des inégalités sociales de réussite scolaire, mais elle a un poids spécifique, et il est dès lors possible d'améliorer les acquisitions d'un grand nombre d'élèves,

en particulier ceux issus des classes sociales défavorisées» (p. 91). Davezies (2005) synthétise lui aussi assez bien les enjeux associés à l'étude des effets de contexte en rappelant que cette problématique est au cœur des réflexions récentes en matière de pilotage du système éducatif dans la mesure où la gestion des politiques de discrimination positive et de mixité sociale en dépend fortement.

À l'automne 2015, le *Centro innovazione e ricerca sui sistemi educativi* (CIRSE) du canton du Tessin et le Service de la recherche en éducation (SRED) du canton de Genève ont amorcé une collaboration méthodologique autour de la mesure des effets de contexte et ont manifesté un intérêt commun pour les modèles multiniveaux.

Ces méthodes d'analyse statistique, développées au milieu des années 80, sont de plus en plus utilisées dans le domaine des sciences de l'éducation car elles permettent de traiter la question fondamentale des relations entre les individus et leur environnement : dans quelle mesure et selon quels processus l'environnement influe-t-il sur les comportements, attitudes, opinions des individus (Bressoux, Coustère et Leroy-Audouin, 1997) ? Ce type de modélisation est d'ailleurs présenté par Paterson et Goldstein (1991), cités par Bressoux (1994), comme une «avancée à la fois statistique et conceptuelle» (p. 127).

De façon plus précise, afin de développer une connaissance partagée de ces méthodes d'analyse et de leur mise en œuvre, le CIRSE et le SRED ont décidé de se pencher sur la question des relations qui existent, dans leurs cantons respectifs, entre acquisitions des élèves et contexte de scolarisation, notamment au niveau de l'établissement. Pour ce faire ils ont décidé de se livrer à une analyse secondaire des résultats de l'enquête PISA 2012 à laquelle a participé la Suisse et pour laquelle nous disposons d'échantillons cantonaux de 11^e année (fin de scolarité obligatoire).

Les enjeux d'une telle recherche sont multiples. Tout d'abord, au-delà de la maîtrise des aspects méthodologiques, il apparaît nécessaire de constituer un savoir spécifique sur l'effet école en Suisse et plus particulièrement au Tessin et à Genève car la question a encore été peu traitée à ce jour ¹.

Comme nous l'avons mentionné précédemment, il existe bien une littérature scientifique abondante sur l'effet de composition mais bon nombre de résultats, comme bien souvent dans le domaine des sciences de l'éducation, sont très dépendants des caractéristiques des systèmes scolaires dans lesquels ils ont été obtenus. Existe-t-il aussi, au Tessin et à Genève, un effet de l'établissement fréquenté sur les compétences des élèves en mathématiques, lecture et sciences ?

Par ailleurs, la comparaison des résultats entre le Tessin et Genève pourrait aussi apporter quelques éléments sur une facette particulière du contexte qui est celle du mode de regroupement à l'échelle d'un système scolaire. En effet, dans ces deux cantons, l'organisation de l'enseignement secondaire I est assez différente puisque le Tessin a fait le choix d'une organisation «intégrée» (élèves dans des classes hétérogènes avec des regroupements par niveau dans certaines matières) alors que Genève a opté pour un système majoritairement organisé autour de regroupements différenciés qui séparent les élèves selon leur niveau académique. Dès lors, l'effet établissement est-il de même taille et de même nature dans ces deux environnements ?

Enfin, de tels travaux sont, d'une manière plus générale, susceptibles d'intéresser les autorités scolaires des deux cantons concernés dans la mesure où ils peuvent potentiellement contribuer à l'identification de facteurs permettant une réflexion et un travail autour des notions d'égalité des chances et d'équité au sein des systèmes de formation.

¹ Origoni (2007) avec les données de PISA 2003 avait déjà fait une analyse de la variance inter-établissements et de celle intra-établissement afin d'explorer l'équité et la qualité dans le système éducatif tessinois, en comparaison avec d'autres cantons.

2. Objectif de la recherche et méthodologie

Le CIRSE et le SRED se proposent d'examiner conjointement dans quelle mesure les résultats obtenus par les élèves tessinois et genevois aux tests PISA 2012 dépendent de leurs caractéristiques individuelles mais aussi de celles de l'établissement dans lequel ils sont scolarisés, notamment de sa composition sociale et académique.

La problématique qui nous intéresse ici est donc, au sens large, celle de l'effet de composition défini par Dumay et Dupriez (2009) comme l'impact des caractéristiques agrégées des élèves après que leur effet sur le plan individuel ait été pris en compte. En d'autres termes, la question de départ pourrait se résumer par «dans quelle mesure deux élèves semblables en début d'une période de formation vont évoluer différemment en fonction des caractéristiques des autres élèves de leur école» (Dumay et Dupriez, 2009, p. 463).

Il convient de relever ici que, comme le mentionne Dumay (2004), l'effet de composition ne doit pas être envisagé comme un phénomène général mais plutôt comme un phénomène spécifique associé à la variable qui a été mesurée. On conçoit, par exemple, assez bien qu'un effet de composition ethnique ne puisse pas s'interpréter de la même manière qu'un effet mesuré à partir des dispositions cognitives des élèves.

Dans la présente étude, la question initiale porte principalement sur les dimensions sociale et académique de l'effet de composition. Par ailleurs, certains travaux (Thrupp, 1999) ont montré que l'effet de composition sur la performance scolaire aurait un caractère indirect, médié par des processus internes à l'établissement. Ainsi, comme le rappellent Duru-Bellat, Danner, Le Bastard-Landrier et Piquée (2004), cet effet n'aurait pas de rôle en lui-même mais il engendrerait des processus pédagogiques ou psychologiques spécifiques qui seraient les variables véritablement causales. En l'occurrence, les processus par lesquels cet effet opère seraient de trois types :

- influence sur l'organisation, le management et le climat général de l'établissement;
- influence sur la quantité et la qualité de l'instruction délivrée en classe;
- influence sur les attitudes, aspirations et comportements des élèves.

En d'autres termes, la ségrégation d'élèves en difficulté scolaire ou de milieu socio-économique défavorisé créerait des dynamiques groupales peu propices à l'apprentissage qui, toutes choses égales par ailleurs², se révéleraient génératrices de performances scolaires plus modestes.

En réalité, les travaux qui ont porté sur les effets de composition sont extrêmement nombreux mais leurs résultats sont très contrastés et comme le relèvent Dumay, Dupriez et Maroy (2009), il n'existe pas à l'heure actuelle de consensus complet dans la littérature scientifique autour de la notion d'effet de composition. Ces auteurs attribuent en partie cette absence de consensus à la diversité des méthodologies mises en œuvre et au manque de pertinence de certaines d'entre elles : «à vrai dire, nous manquons encore aujourd'hui d'un nombre suffisant d'études rigoureuses centrées sur l'effet de composition» (p. 2).

Ces dernières années, plusieurs chercheurs se sont penchés sur les problèmes méthodologiques et statistiques que pose l'estimation d'un tel effet (Thrupp, Lauder et Robinson, 2002; Gorard, 2006; Harker et Tymms, 2004).

Thrupp et al. (2002), cités par Dumay et al. (2009), ont établi un certain nombre de recommandations méthodologiques qui doivent idéalement guider l'étude de l'effet de composition sur les performances des élèves :

- L'échantillon doit inclure des écoles situées aux deux extrêmes du spectre socioéconomique et socioculturel. Cette disposition vise à éviter l'écueil auquel ont été confrontés de nombreux travaux de recherche qui ont été réalisés sur des échantillons relativement homogènes d'écoles plutôt urbaines et défavorisées (Duru-Bellat, 2003). Dans ce cas, la réflexion sur l'effet de composition est rendue plus difficile dans la mesure où le public qui compose ces établissements est peu diversifié.
- Le modèle d'estimation doit incorporer un ensemble complet de prédicteurs individuels et au minimum une mesure de l'état des connaissances antérieures des élèves.

² Cette expression, synonyme de *ceteris paribus*, traduit simplement le fait que toutes les autres variables explicatives introduites dans la modélisation sont maintenues constantes.

- De multiples mesures de la composition doivent être construites (à la fois des indicateurs de moyenne et des indicateurs de proportion).
- Le modèle d'analyse doit également intégrer des variables qui, potentiellement, co-varient avec la composition (par exemple la qualité de l'enseignement ou le climat de travail). Il s'agit ici de distinguer une influence directe de la composition liée à des relations entre élèves et une influence indirecte qui relève davantage de ce que propose l'école et qui est affectée par la composition de celle-ci.
- Enfin, dans la mesure du possible, le design de recherche devrait être longitudinal.

Par ailleurs, le choix de la méthodologie à mettre en œuvre est également déterminant pour mesurer un effet de la composition des établissements scolaires. Ainsi, Duru-Bellat (2003) relève que bon nombre des travaux qui ont cherché à mesurer cet effet ne reposent pas sur des méthodes statistiques adaptées : «les méthodologies courantes sont bien trop frustrées pour appréhender la notion même d'effet de contexte (effet établissement et effet classe) qui ne peut être pensée et a fortiori évaluée sans une instrumentation de type analyse multivariée» (p. 187).

Parmi la multitude de méthodes d'analyse multivariées disponibles, les modèles multiniveaux apparaissent comme les plus appropriés pour traiter la problématique de l'effet du contexte sur les performances des élèves. En effet, ces modèles opèrent une décomposition de la variance totale des performances en une variance individuelle et une variance contextuelle (variance due à l'école) de façon simultanée. Ceci conduit Bressoux, Coustère et Leroy-Audouin (1997) à considérer que les modèles multiniveaux proposent une solution adéquate au problème longtemps posé dans les sciences sociales, celui de rendre compte du fait que l'individu s'insère dans un milieu qui ne peut manquer d'influer sur lui» (p. 89).

En synthèse, selon Dumay et al. (2009) un design de recherche pertinent pour estimer l'effet de composition nécessite au minima deux mesures d'acquis scolaires (dont une mesure du niveau antérieur / initial des élèves), un échantillonnage approprié et des modèles d'analyse multiniveaux.

2.1 Un design de recherche particulièrement adapté à la mesure de l'effet de composition au Tessin et à Genève

La philosophie générale de la recherche envisagée par le CIRSE et le SRED repose sur l'idée de compléter les données des échantillons tessinois et genevois issus de l'enquête PISA 2012 par des informations extraites des bases de données scolaires cantonales, notamment en termes de résultats scolaires antérieurs et de composition des établissements. Ceci devrait permettre de s'affranchir de bon nombre de limites couramment associées aux travaux sur la mesure des effets de contexte et de réaliser une recherche dont le design et la méthodologie correspondent à ce qui est minimalement considéré comme pertinent pour mesurer l'effet de la composition sociale et académique des établissements sur les performances des élèves.

De façon plus précise, voici dans quelle mesure la présente recherche se conforme aux recommandations émises par Thrupp, Lauder et Robinson (2002) évoquées précédemment :

- **Mesure de l'état des connaissances antérieures des élèves** : sur la base des enquêtes PISA 2000 et 2006, Monseur et Crahay (2006, 2008) ont conclu à la présence d'un effet de composition socio-économique des écoles sur les performances en lecture et en sciences dans la plupart des pays. Toutefois, bien qu'extrêmement intéressants, leurs travaux ne s'affranchissent pas d'une limite majeure mentionnée précédemment, à savoir l'absence de contrôle du niveau antérieur des élèves. Il en va de même pour les travaux de Moreau (2007) portant sur l'analyse des résultats des élèves vaudois à l'enquête PISA 2003. Certains auteurs (Willms, 1992; Lauder et Hughes, 1990) ont émis l'hypothèse que les effets de composition mis en évidence dans de telles recherches pourraient, en réalité, être dus à un mauvais contrôle de variables caractérisant les individus, et tout particulièrement à une prise en compte insuffisante du niveau initial des élèves. En fait, comme le relèvent Duru-Bellat, Danner, Le Bastard-Landrier et Piquée (2004), «les données PISA, dans la mesure où elles ne fournissent pas une mesure du niveau initial (mais seulement une performance ponctuelle) ne permettent pas de parler d'effet du *social mix* de

l'établissement en tant que tel mais uniquement d'une corrélation avec ladite performance» (p. 16).

En effet, l'absence de contrôle du niveau antérieur des compétences empêche d'affirmer avec certitude qu'il existe bien un effet de la composition de l'établissement sur les performances des élèves. À caractéristiques individuelles identiques, est-ce l'exposition à un contexte différent qui influe sur les performances d'un élève (en l'occurrence un contexte dont la tonalité sociale ou académique ne serait pas la même) ou les différences de résultats observées sont-elles simplement liées au fait que certains élèves sont initialement plus performants ?

Disposer d'une seconde mesure d'acquisitions permet en fait de raisonner sur les progressions des élèves et donc en termes de valeur ajoutée par l'école.

Les données disponibles dans les cantons du Tessin et de Genève vont permettre de se positionner dans ce cadre pour des élèves qui ont participé à PISA puisqu'il sera possible de contrôler leur niveau initial en lecture, mathématiques et sciences en se basant sur leurs résultats scolaires antérieurs.

- **Echantillonnage d'écoles couvrant la totalité du spectre socioéconomique et socioculturel** : concernant l'échantillonnage des écoles, tous les établissements du Tessin (soit 35 établissements) et la quasi-totalité des établissements du canton de Genève³ (soit 19 établissements) ont participé à l'enquête PISA 2012.

Dès lors, les données dont nous disposons couvrent bien l'ensemble du spectre socioéconomique et socioculturel des écoles pour chacun des cantons concernés, ce qui correspond aux recommandations mentionnées précédemment.

Par ailleurs, dans les travaux antérieurs réalisés sur les données PISA, la composition sociale des établissements scolaires a été définie sur la base des caractéristiques individuelles des seuls élèves ayant participé à l'enquête, notamment de leur statut socio-économique et culturel (ESCS). Or cette variable n'est pas un critère d'échantillonnage dans l'enquête PISA où seul le critère de l'âge des élèves est utilisé⁴.

L'utilisation d'informations extraites des bases de données scolaires cantonales devrait permettre de parvenir à une caractérisation plus judicieuse du profil socio-économique des établissements en utilisant les caractéristiques de toute la population d'élèves qui y est scolarisée. Elle permettra aussi de proposer, comme le recommandent Thrupp, Lauder et Robinson (2002), plusieurs mesures de la composition qui seront introduites dans la modélisation, quand les travaux antérieurs reposant sur les seules données PISA n'en avaient souvent qu'une seule (moyenne du statut socio-économique des élèves de l'établissement).

- **Variables qui co-varient avec la composition** : d'après les chercheurs qui ont formulé des recommandations méthodologiques relatives à l'étude de l'effet de composition sur les performances des élèves, il est souhaitable que le modèle d'analyse intègre également des variables qui co-varient potentiellement avec la composition de l'établissement scolaire.

Dans l'enquête PISA, les chefs d'établissement ont été interrogés sur l'environnement d'apprentissage de leur établissement. Les informations collectées ont permis à l'OCDE (2014) de définir un certain nombre d'indices synthétiques relatifs à la gestion, aux ressources ainsi qu'à différentes dimensions du climat d'établissement.

Ces données seront utilisées dans l'analyse en complément de celles extraites des fichiers cantonaux qui permettront de caractériser la composition sociale et académique des établissements.

- **Modèle multiniveau** : en ce qui concerne la méthode d'analyse envisagée, le CIRSE et le SRED auront bien recours au modèle multiniveau préconisé dans le traitement de la problématique de l'effet de composition.

Les données dont nous disposons présentent une structure hiérarchisée, constituée de niveaux emboîtés les uns dans les autres (des élèves scolarisés dans des établissements) qui s'avère tout à fait propice à la mise en œuvre de ce type d'analyse.

Pour résumer, les données dont le CIRSE et le SRED disposent permettent de réaliser, de façon adéquate, une recherche visant à mesurer l'effet de la composition sociale et académique des établissements scolaires sur les performances des élèves en mathématiques, lecture et sciences. Par ailleurs, ces travaux portent sur une problématique relativement peu explorée à ce jour avec des

³ Seul un Collège est absent de l'échantillon PISA. Il a été exclu de l'enquête en raison d'un problème survenu lors de la passation des tests (questionnaires reçus en allemand plutôt qu'en français).

⁴ En Suisse, le critère du degré scolaire est aussi utilisé pour la constitution d'échantillons régionaux et cantonaux (11^e année, soit dernière année de scolarité obligatoire). Consortium Pisa.ch (2011).

données PISA, en particulier dans le contexte Suisse, et sur laquelle les connaissances sont encore à développer.

3. Rappel de quelques éléments contextuels

Avant d'entrer de façon plus approfondie dans la méthodologie et les données qui permettent de se livrer à l'étude de l'effet de composition des établissements scolaires sur les compétences des élèves au Tessin et à Genève, il apparaît nécessaire de procéder à un bref rappel sur l'organisation de l'enseignement secondaire I dans chacun de ces cantons puisque cette dernière diffère sensiblement.

Par ailleurs, nous présenterons également une synthèse des principaux résultats obtenus dans chacun des trois domaines testés afin de rappeler certaines spécificités cantonales mises en évidence dans des travaux antérieurs (Nidegger (éd.), 2014; Salvisberg et Zampieri, 2015).

Enfin nous illustrerons brièvement, de façon descriptive, les différences de composition qui existent au Tessin et à Genève entre établissements scolaires. Il s'agira de montrer qu'il y a bien une certaine hétérogénéité des établissements qui justifie que l'on s'intéresse à un possible effet du contexte sur les compétences des élèves.

3.1 Organisation de l'enseignement secondaire I dans le Canton du Tessin

Au Tessin, l'enseignement secondaire I est divisé en deux cycles : l'un, dénommé cycle d'observation («*ciclo d'osservazione*») couvre la 8^e et la 9^e année⁵ (*prima e seconda media*), et l'autre, défini comme cycle d'orientation («*ciclo d'orientamento*»), couvre la 10^e et la 11^e année («*terza e quarta media*»).

Dès le deuxième cycle, soit à partir de la 10^e année, les élèves sont répartis dans des niveaux pour les cours de mathématiques et d'allemand. En général les élèves qui ont obtenu au moins la note de 4.5 à la fin de la 9^e en mathématiques et en allemand vont suivre le niveau A (*corso attitudinale*). Les élèves qui ont quelques difficultés de plus suivent le niveau B (*corso base*). La répartition fait aussi parfois l'objet de discussions au sein du conseil de classe et les parents peuvent explicitement demander que leur fils/fille fréquente un niveau A. Enfin, les élèves qui rencontrent encore plus de difficultés ont la possibilité de suivre des cours appelés «*differenziazione curriculare*» qui leur sont spécifiquement dédiés.

Pour toutes les autres disciplines, les élèves ne sont pas triés selon leur niveau scolaire et fréquentent les cours avec toute la classe (à l'exception des cours optionnels⁶). Il n'y a donc pas de différenciation curriculaire dans l'enseignement secondaire I.

On relèvera également qu'à l'issue de ce degré d'enseignement, l'accès aux différentes écoles du secondaire II dépend principalement de la moyenne des notes dans toutes les matières, mais également des niveaux en mathématiques et allemand, ainsi que de la note d'italien. Par exemple, pour accéder au lycée ou à une quelconque autre école professionnelle, il est nécessaire d'avoir les deux niveaux A, une moyenne générale d'au minimum 4.65, une seule note insuffisante et la note d'italien d'au minimum 4.5⁷.

3.2 Organisation de l'enseignement secondaire I dans le Canton de Genève

Lors de la réalisation de l'enquête PISA 2012 (année scolaire 2011-2012), l'enseignement secondaire I genevois est majoritairement organisé en regroupements différenciés qui séparent les élèves selon

⁵ On utilise ici comme référence les années de scolarité HarmoS.

⁶ Par exemple le cours de latin, les cours d'orientation, de capacités techniques (éducation alimentaire, techniques d'habillement, etc.) ou de capacités expressives (éducation aux arts visuels, éducation musical, etc.).

⁷ Si l'élève ne remplit pas les critères pour accéder au lycée il a la possibilité de passer un examen d'admission. En outre, si seulement un critère n'est pas respecté il existe d'autres dispositions qui permettent d'accéder au lycée sans examen. Ces critères sont détaillés dans l'Art. 65[88] du «*Regolamento della scuola media*» (du 18 septembre 1996).

leur niveau académique (17 établissements). Le regroupement A se caractérise par des exigences scolaires étendues et des effectifs ordinaires alors que les regroupements B et C correspondent à des exigences scolaires moindres et à des effectifs réduits. On notera également que le regroupement C n'existe pas dans tous les établissements⁸.

Par ailleurs, en parallèle de ce système, trois établissements fonctionnent avec des classes hétérogènes dans lesquelles les élèves n'ont fait l'objet d'aucune sélection sur la base de leurs résultats scolaires antérieurs. Ils sont, en revanche, affectés dans des niveaux différents en mathématiques et en allemand (B, R, C en 10^e et A, B, C en 11^e). Enfin, au sein de chacun de ces systèmes des options sont proposées (à partir de la 10^e) : langues anciennes (latin), sciences et arts dans le Regroupement A et les classes hétérogènes alors que seules les options sciences et arts sont disponibles dans le Regroupement B.

Pour les besoins de la présente recherche, des profils comparables à ceux des regroupements A et B ont été définis pour les classes hétérogènes à partir des niveaux en allemand et en mathématiques. Ainsi on retiendra le découpage suivant :

- Regroupement A ou assimilé (exigences élevées) : élèves du regroupement A et des classes hétérogènes avec les configurations de niveaux en allemand et mathématiques suivantes :
 - BB, BR ou RB pour la 10^e CO
 - AA, AB ou BA pour la 11^e CO
- Regroupement B ou assimilé (exigences moyennes ou élémentaires) : élèves du regroupement B et des classes hétérogènes de 10^e et 11^e avec des configurations de niveaux en allemand et mathématiques différentes de celles retenues pour la catégorie «Regroupement A et assimilé».

À ce stade de la description, il convient de relever qu'en 2011-2012 l'organisation décrite précédemment ne s'applique qu'aux élèves de 10^e et 11^e années. En effet, pour les élèves de 9^e année (soit la 1^{re} année de scolarité dans le secondaire I), il s'agit d'une période de mutation qui correspond à la mise en œuvre progressive de la nouvelle organisation de l'enseignement secondaire I. Cette dernière correspond à l'instauration de trois regroupements (R1, R2, R3) dans un système unifié commun à l'ensemble des établissements du canton. Le regroupement R3 correspond à la filière aux exigences élevées alors que les regroupements R2 et R1 seront assimilés à la filière aux exigences moyennes et élémentaires.

3.3 Illustration de la diversité des établissements de l'enseignement secondaire I dans chaque canton

Comme nous l'avons mentionné précédemment, la problématique centrale de cette recherche est celle de l'effet de composition des établissements scolaires.

De façon plus précise, il s'agit de voir dans quelle mesure les résultats obtenus par les élèves tessinois et genevois aux tests PISA 2012 dépendent de leurs caractéristiques individuelles mais aussi de celles de l'établissement dans lequel ils sont scolarisés, notamment de sa composition sociale et académique.

Cette question de la composition des établissements et plus particulièrement de sa dimension sociale et académique n'est légitime que dans des contextes où il existe bel et bien des différences entre les établissements. Est-ce le cas dans les cantons du Tessin et de Genève ?

De façon relativement descriptive on peut essayer de représenter l'hétérogénéité des publics accueillis au sein des établissements de l'enseignement secondaire I en mobilisant les deux indicateurs de la composition sociale et académique que sont la proportion d'élèves issus d'un milieu favorisé (définitions différentes dans chaque canton) et la proportion d'élèves considérés comme scolairement plus performants, soit parce qu'ils suivent des cours dans les niveaux les plus exigeants

⁸ Par ailleurs ce regroupement ne concerne théoriquement que la 9^e année (1^{re} année de scolarité du secondaire I). Toutefois, en raison de la mise en place progressive de la nouvelle organisation du CO dès la rentrée scolaire 2011-2012, on ne dénombre plus aucun élève dans le regroupement C au moment de la passation de PISA 2012.

(deux niveaux A en mathématiques et allemand au Tessin), soit parce qu'ils sont scolarisés dans la filière à exigences élevées (regroupement A ou assimilé pour le canton de Genève).

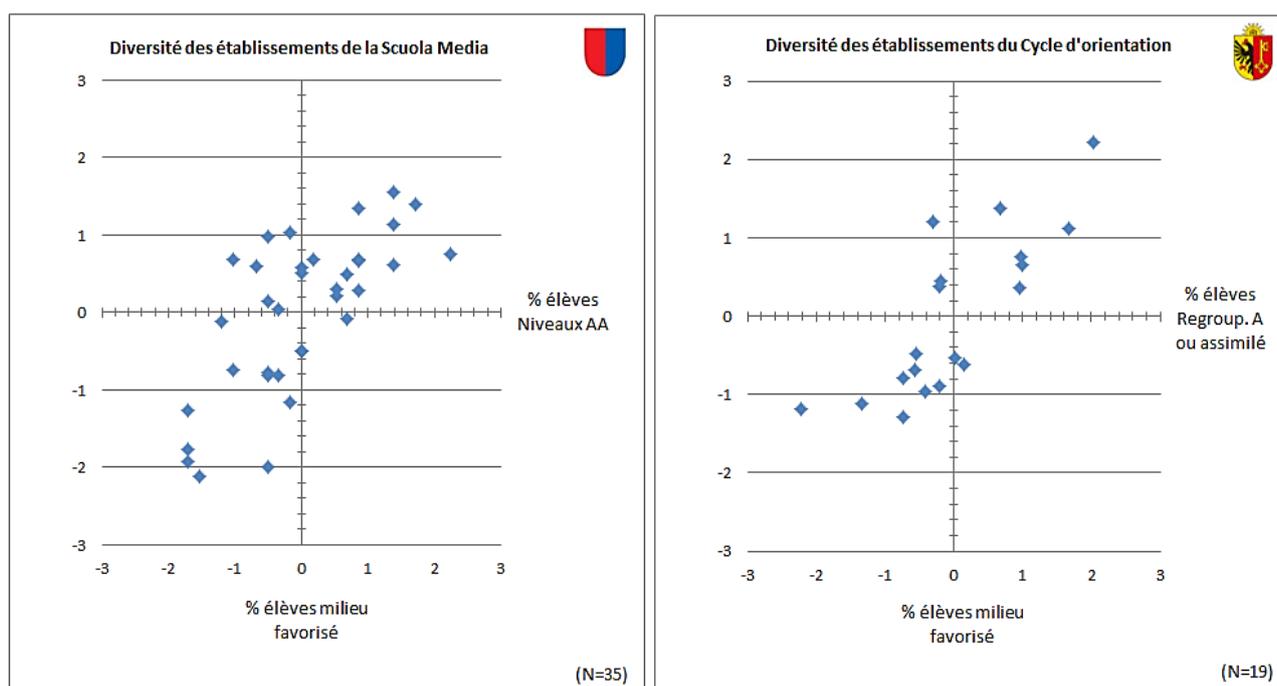
Afin de rendre les choses plus comparables, nous avons centré et réduit ces proportions pour que chaque établissement soit situé par rapport aux valeurs moyennes qui prévalent dans son canton (Figure 1). Ainsi, plus l'établissement est éloigné de l'intersection des deux axes et plus sa situation diverge d'une certaine «norme» cantonale. La dispersion des points sur chacun des graphiques de la Figure 1 traduit assez bien le fait qu'au Tessin comme à Genève, il existe une certaine diversité dans la composition des établissements.

Pour illustrer encore davantage le propos, on notera par exemple que la proportion d'élèves issus d'un milieu favorisé varie de moins de 30% à près de deux tiers au sein des établissements de la Scuola Media (moyenne cantonale = 50.4%) ou que la proportion d'élèves scolarisés dans la filière à exigences élevées varie de 52% à 82% dans les établissements du Cycle d'orientation (moyenne cantonale = 67.6%).

Ces différences relativement importantes justifient dès lors que l'on s'intéresse à l'impact de l'hétérogénéité des contextes de scolarisation sur les résultats des élèves.

Figure 1 – Diversité des établissements de la Scuola Media et du Cycle d'orientation par rapport aux moyennes cantonales

Proportions d'élèves de milieu favorisé et d'élèves de niveaux AA (TI) ou scolarisés dans la filière à exigences élevées (GE) – Valeurs centrées et réduites



Source : SRED et CIRSE

Note de lecture : les points représentent les établissements. Par exemple, ceux qui se trouvent à la droite de l'axe vertical sont des établissements qui ont un pourcentage d'élèves de milieu favorisé supérieur à la moyenne cantonale. Avec la même logique, les établissements qui se trouvent sous l'axe horizontal ont un pourcentage d'élèves de niveaux AA ou de Regroupement A inférieur à la moyenne cantonale.

4. Echantillonnage et données

4.1 Taille des échantillons

Les échantillons cantonaux d'élèves de 11^e année qui ont passé les tests PISA en 2012 au Tessin et à Genève ont les caractéristiques suivantes⁹ :

- Tessin : 1'081 élèves scolarisés dans 35 établissements soit, en moyenne, 31 élèves par établissement (ce nombre varie de 25 à 35 élèves selon l'établissement considéré);
- Genève : 920 élèves scolarisés dans 19 établissements soit, en moyenne, 48 élèves par établissement (ce nombre varie de 31 à 69 élèves selon l'établissement considéré).

Bien que la collecte de ces données représente un effort substantiel pour les cantons concernés, on peut tout de même légitimement se demander si leur volume est suffisant pour mettre en œuvre une analyse de type multiniveau indispensable au traitement de la problématique de l'effet de composition.

En réalité, dans ces modèles le nombre d'observations s'appréhende simultanément à deux niveaux : le nombre de sujets au sein de chaque groupe (en l'occurrence des élèves dans des établissements) et le nombre total de groupes. Les travaux de Kreft et de Leeuw (1998) ont montré que pour obtenir des résultats de qualité avec les modèles multiniveaux, il est recommandé de disposer d'un échantillon qui comporte au minimum 30 groupes dont chacun est composé de 30 sujets (soit au total environ un millier de sujets).

Toutefois, ces mêmes travaux ont aussi montré qu'il existe un arbitrage entre ces deux paramètres et que lorsqu'on augmente le nombre de groupes, on peut diminuer le nombre de sujets au sein de chacun d'entre eux et réciproquement. Ainsi les auteurs indiquent par exemple qu'un échantillon de 150 groupes avec cinq sujets par groupe permet de parvenir à des résultats de même qualité que ceux obtenus à partir d'un échantillon qui comporterait 30 groupes dont chacun serait composé de 30 sujets.

Par ailleurs, les modèles multiniveaux permettent également de travailler sur des données déséquilibrées, c'est-à-dire n'ayant pas le même nombre d'observations dans chaque groupe (dans notre cas, pas le même nombre d'élèves par établissement).

Sur la base des éléments qui précèdent, on peut raisonnablement considérer que les données dont disposent les cantons du Tessin et de Genève sont globalement suffisantes pour mettre en œuvre une analyse multiniveau.

Toutefois, ces dernières correspondent en quelque sorte aux volumes minimalement requis pour ce type d'analyse. Dès lors ceci aura des répercussions sur la spécification des modèles et plus particulièrement sur le nombre de variables qui pourront être introduites dans la modélisation pour caractériser l'établissement (voir 5. Présentation du modèle multiniveau).

4.2 Les trois domaines de compétence évalués dans PISA 2012¹⁰

Dans l'enquête PISA, les compétences et connaissances des élèves âgés de 15 ans sont testées dans les trois domaines importants que sont les mathématiques, la lecture (compréhension de l'écrit) et les sciences.

Ces trois domaines sont considérés comme essentiels pour l'avenir personnel et professionnel des élèves dans une société hautement développée. Il ne s'agit donc pas de vérifier que ces derniers maîtrisent les contenus des différents curriculums auxquels ils sont soumis mais plutôt de savoir dans quelle mesure leurs compétences leur permettent de maîtriser les situations de la vie quotidienne et de répondre aux défis de leur vie future. De façon plus précise, la définition des trois compétences évaluées est la suivante (Nidegger (ed.), 2014) :

⁹ Il s'agit des échantillons qui ont été collectés en plus de l'échantillon d'élèves de 15 ans utilisé pour caractériser la Suisse dans les comparaisons internationales et pour permettre une comparaison inter-cantonale.

¹⁰ D'après Nidegger (éd.) (2014) et Suchaut (2013).

- **Culture mathématique (mathématiques)** : il s'agit de l'aptitude d'un individu à formuler, employer et interpréter les mathématiques dans un éventail de contextes, c'est-à-dire à raisonner en termes mathématiques et à utiliser des concepts, procédures, faits et outils mathématiques pour décrire, expliquer et prévoir des phénomènes, ainsi qu'à comprendre le rôle que les mathématiques jouent dans le monde et à se comporter en citoyen constructif, engagé, réfléchi.
- **Compréhension de l'écrit (lecture)** : c'est non seulement comprendre et utiliser des textes écrits, mais aussi réfléchir à leur propos et s'y engager. Cette capacité devrait permettre à chacun de réaliser ses objectifs, de développer ses connaissances et son potentiel, et de prendre une part active dans la société.
- **Culture scientifique (sciences)** : comprend les connaissances scientifiques de l'individu et sa capacité à utiliser ces connaissances pour identifier les questions auxquelles la science peut apporter une réponse, pour acquérir de nouvelles connaissances, pour expliquer des phénomènes scientifiques et pour tirer des conclusions fondées sur des faits scientifiques; la compréhension des éléments caractéristiques de la science en tant que forme de recherche et de connaissance humaines; la conscience du rôle de la science et de la technologie dans la constitution de notre environnement matériel, intellectuel et culturel; la volonté de s'engager en qualité de citoyen réfléchi à propos de problèmes à caractère scientifique et touchant à des notions relatives à la science.

Ce sont les scores obtenus par les élèves tessinois et genevois dans ces trois domaines, ou plus précisément leur variance, que nous allons chercher à expliquer dans la présente recherche. Comme le rappelle Suchaut (2013),

«L'enquête PISA fixe la moyenne des scores des élèves à 500 avec un écart-type de 100, cette échelle arbitraire présentant un avantage majeur qui est celui de la distinguer, d'une part d'autres échelles utilisées dans d'autres tests très connus et, d'autre part, des barèmes de notation traditionnels des élèves en vigueur dans les systèmes scolaires. En outre, la distribution des scores des élèves suit la loi normale dont la connaissance des propriétés statistiques est essentielle pour comprendre la manière dont ces scores se répartissent autour de la moyenne. Sans développer ici les propriétés de la loi normale, on signalera seulement que la courbe dérivée de cette loi a une allure bien connue en forme de cloche qui met en évidence la concentration des valeurs autour de la moyenne ce qui a pour conséquence que plus les valeurs s'éloignent de cette moyenne, plus leur fréquence d'apparition devient faible. Ainsi, environ les deux tiers des valeurs sont comprises entre plus ou moins un écart-type autour de la moyenne, 95% des valeurs sont comprises entre la moyenne et plus ou moins deux écart-types et presque la totalité des valeurs entre la moyenne et plus ou moins trois écart-types. Concrètement, dans l'échelle de PISA, les deux tiers des élèves ont donc un score compris entre 400 et 600 points (puisqu' l'écart-type est de 100), 95% entre 300 et 700 points et 99% entre 200 et 800 points» (p. 4).

D'ailleurs, dans l'enquête PISA chaque élève se voit attribuer dans un domaine de compétences donné non pas une mais cinq valeurs de score différentes (les cinq *plausible values*). Ces dernières permettent, en quelque sorte, d'estimer l'incertitude liée au fait que les compétences ne sont en réalité pas directement observées. En effet, comme le rappelle Nidegger (éd.) (2014), chaque élève répond bien par écrit à une épreuve de mathématiques, compréhension de l'écrit et sciences mais les cahiers de tests sont différents et personne ne répond à la totalité des questions testées¹¹.

Des analyses complexes, basées sur la théorie des réponses aux items (IRT), permettent de positionner sur une même échelle, à l'aide de procédures itératives, à la fois le niveau de difficulté des items et les compétences des élèves. Enfin, on notera que les analyses multiniveaux qui figurent dans la présente recherche ont été réalisées avec la version 7 du logiciel HLM qui permet non seulement de travailler avec les cinq *plausible values* définies pour chacun des domaines de compétences mais aussi de tenir compte de la pondération des données réalisée par le consortium international responsable du programme PISA et qui assure la représentativité de ces dernières.

¹¹ Il s'agit de pouvoir évaluer un grand nombre de tâches et d'items tout en limitant la durée (2 heures) et le coût de l'épreuve et en évitant la démotivation et la fatigue des personnes interrogées.

4.3 Quelques résultats à PISA pour les cantons du Tessin et de Genève

Avant de s'intéresser plus directement à la problématique de l'effet établissement sur les performances des élèves dans les différents domaines évalués, il nous a semblé nécessaire de rappeler ici, à des fins de cadrage, quelques éléments généraux sur les résultats obtenus à PISA 2012 dans les deux cantons concernés. Il s'agira notamment de situer les résultats du Tessin et de Genève les uns par rapport aux autres et de voir dans quelle mesure ils divergent de ceux observés pour l'ensemble de la Suisse.

Ces éléments ont, de plus, fait l'objet d'une littérature relativement abondante et le lecteur désireux d'en savoir davantage pourra se reporter à Salvisberg et Zampieri (2015) ou à Nidegger (éd.) (2014).

4.3.1 Les mathématiques

En mathématiques, domaine principal de l'enquête PISA 2012, le Tessin obtient un score moyen de 515 points, statistiquement inférieur à la moyenne suisse (531 points) (Figure 2). Le canton de Genève obtient lui aussi des résultats plus faibles que la moyenne nationale avec un score de 502 points.

En termes de dispersion, on peut relever qu'au Tessin les résultats sont plus homogènes qu'à Genève et qu'en Suisse de manière générale, l'écart entre les élèves les meilleurs et les plus faibles y étant plus modeste.

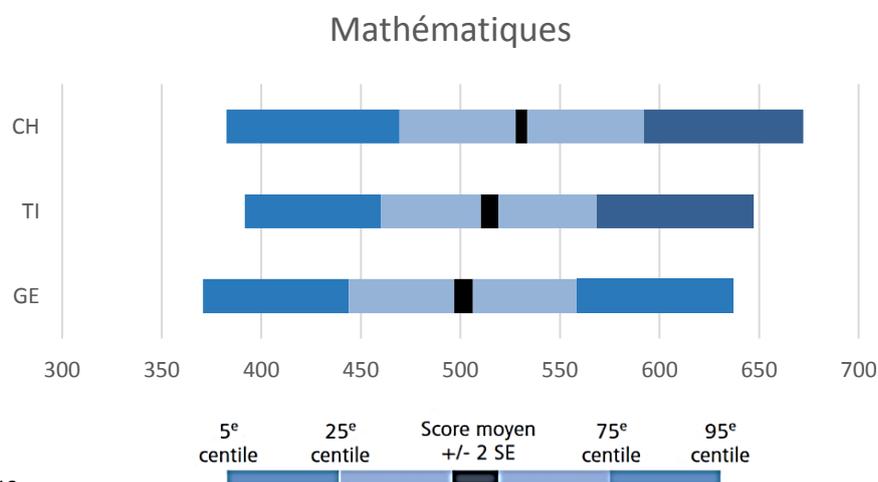
En ce qui concerne les facteurs qui influencent les performances des élèves, au Tessin le genre ne possède pas d'effet statistiquement significatif alors qu'à Genève le fait d'être un garçon constitue, toutes choses égales par ailleurs, un avantage (Salvisberg et Zampieri, 2015).

Dans les deux cantons concernés, le niveau socio-économique a un impact sur les résultats. De façon plus précise, les élèves tessinois issus d'un milieu défavorisé obtiennent des résultats statistiquement inférieurs et, à l'inverse, ceux issus des milieux les plus favorisés obtiennent de meilleurs résultats. À Genève, seuls les élèves issus d'un milieu favorisé obtiennent des scores plus élevés.

Par ailleurs, la langue parlée à la maison n'a pas d'impact sur les performances des jeunes Tessinois alors qu'à Genève les élèves allophones réussissent moins bien.

Enfin, dans les deux cantons, le statut migratoire influence négativement les résultats. En effet, on peut montrer que les natifs obtiennent des résultats en mathématiques statistiquement supérieurs à ceux des élèves issus de la migration.

Figure 2 – Résultats moyens des élèves en mathématiques à PISA 2012 dans les cantons du Tessin et de Genève



Source : PISA 2012

À l'issue de cette brève présentation, il convient de relever que l'impact des principales caractéristiques sociodémographiques des élèves sur leurs performances en mathématiques pourra se révéler différent dans les résultats d'analyses multiniveaux qui suivront, dans la mesure où d'autres informations que celles contenues dans PISA seront prises en compte, notamment le niveau initial des élèves.

4.3.2 La lecture et les sciences

Dans les domaines de la lecture et des sciences, les résultats obtenus par les élèves tessinois et genevois sont, là aussi, inférieurs à la moyenne nationale. En effet, en lecture le score moyen du Tessin s'élève à 485 points et à 501 points pour Genève, alors que le score pour l'ensemble de la Suisse est de 507 points (Figure 3). En sciences, la moyenne nationale est de 513 points alors que Genève et le Tessin obtiennent respectivement 489 points et 490 points, soit des résultats très similaires (Figure 4).

Concernant la dispersion des résultats, on constate que pour la lecture les deux cantons ont des résultats relativement proches qui correspondent à ce que l'on observe au niveau national.

En revanche, pour les sciences, Genève et le Tessin ne se distinguent pas l'un de l'autre mais présentent des résultats plus homogènes que ce qui s'observe au niveau national.

Figure 3 – Résultats moyens des élèves en lecture à PISA 2012 dans les cantons du Tessin et de Genève

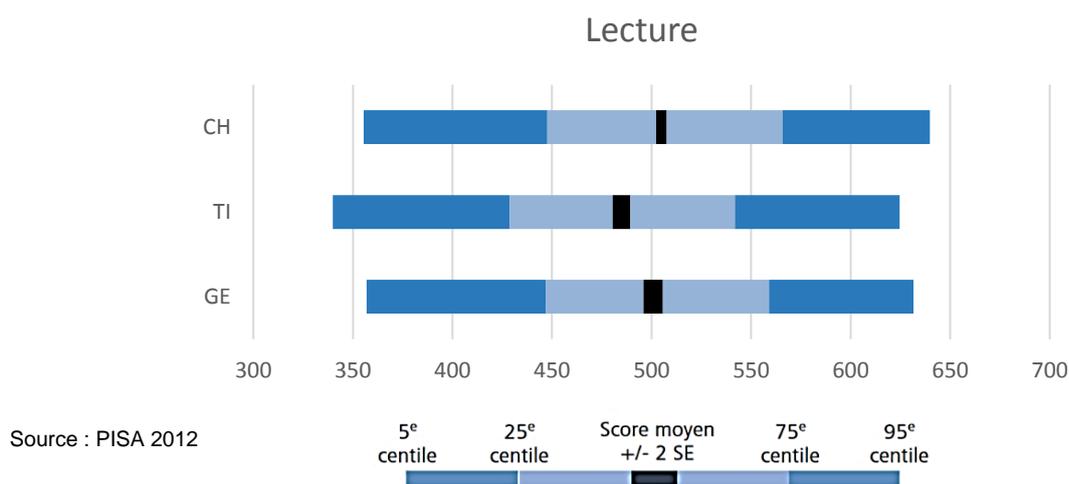
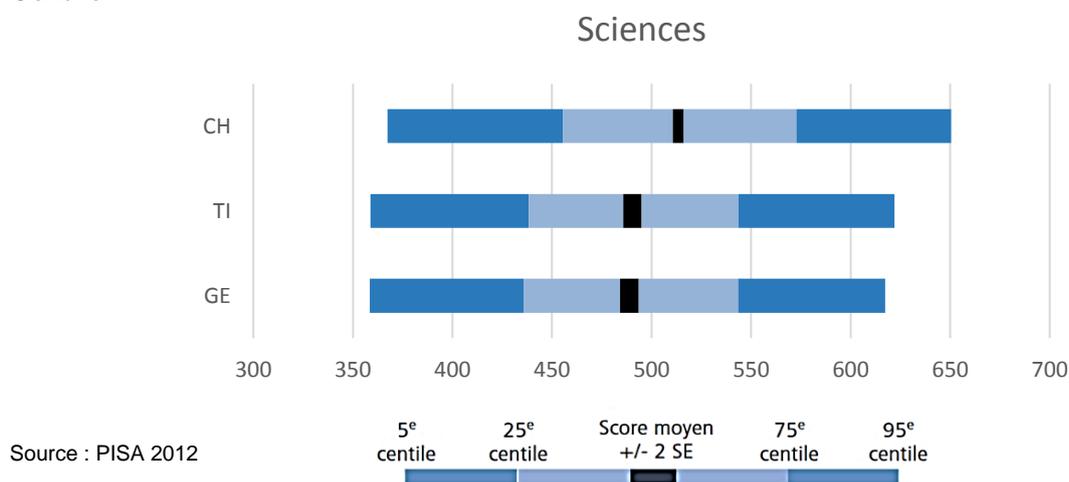


Figure 4 – Résultats moyens des élèves en sciences à PISA 2012 dans les cantons du Tessin et de Genève



4.4 Les caractéristiques des élèves

4.4.1 Les caractéristiques sociodémographiques usuelles (d'après l'enquête PISA 2012)

En plus de répondre aux tests en mathématiques, lecture et sciences, les élèves qui participent à l'enquête PISA 2012 sont interrogés sur leur milieu familial, leurs stratégies d'apprentissage, leurs attitudes à l'égard des mathématiques, leur engagement et leur motivation (OCDE, 2014).

Ces informations ont été utilisées pour définir les principales caractéristiques sociodémographiques des élèves tessinois et genevois qui seront introduites dans la modélisation. Il s'agit plus particulièrement des variables :

- **Genre** : garçon et fille (catégorie de référence). Dans l'échantillon PISA, environ la moitié des élèves sont des filles dans les cantons de Genève et du Tessin (respectivement 50.3% et 49.7%).
- **Age** : correspond à la différence entre le mois et l'année de l'évaluation et le mois et l'année de naissance des élèves (est exprimé en années et en mois).
- **Statut migratoire**: l'indice d'ascendance allochtone (IMMIG) comporte les catégories suivantes : (1) les élèves autochtones (élèves nés dans le pays de l'évaluation ou dont au moins un parent est né dans ce pays [les élèves nés à l'étranger d'au moins un parent né dans le pays de l'évaluation font également partie de cette catégorie]); (2) les élèves de la deuxième génération (élèves nés dans le pays de l'évaluation de parents nés à l'étranger); et (3) les élèves de la première génération (élèves nés à l'étranger de parents nés à l'étranger). Des données sont déclarées manquantes si les élèves n'ont pas répondu à la question les concernant, à celles concernant leurs parents ou s'ils n'ont répondu à aucune des trois questions.
À partir de cet indice détaillé on constitue une variable qui ne comporte que deux modalités, à savoir les natifs (référence) et les immigrés (1^{er} ou 2^e génération). Dans l'échantillon PISA, seuls 52% des élèves genevois sont considérés comme natifs alors que cette proportion est de 70.1% au Tessin.
- **Langue** : les élèves ont indiqué la langue qu'ils parlent le plus souvent en famille. Les données sont collectées via un code de langue spécifique à chaque pays qui est ensuite recodé en deux modalités : la langue parlée en famille est identique à la langue de l'évaluation (français ou italien pour les cantons concernés; référence) et la langue parlée en famille est différente de la langue de l'évaluation (allophone). Dans l'échantillon PISA, 74.2% des élèves genevois parlent le plus souvent la langue du test alors qu'au Tessin cette proportion dépasse 78%.
- **Statut socioéconomique et culturel (ESCS)** : l'indice PISA de statut économique, social et culturel (ESCS) est dérivé de trois indices : le statut professionnel le plus élevé des parents (HISEI), le niveau de formation le plus élevé des parents convertis en années d'études d'après la CITE (PARED) et le patrimoine familial qui inclut notamment des éléments sur la bibliothèque familiale (nombre de livres). Plus précisément, l'indice ESCS est obtenu à partir d'une analyse en composantes principales de ces variables normalisées (la moyenne de l'OCDE de toutes ces variables est égale à 0, et leur écart-type à 1).

4.4.2 Le contrôle du niveau scolaire antérieur / initial (d'après fichiers cantonaux)

Le système genevois d'enseignement et de formation ainsi que le système éducatif tessinois se sont depuis longtemps dotés d'épreuves de référence visant à évaluer les connaissances et compétences acquises par les élèves en fonction des objectifs d'apprentissage définis dans les programmes d'enseignement.

Ces épreuves constituent, à l'échelle d'un canton, un matériau précieux et privilégié pour les travaux portant sur l'acquisition des compétences scolaires dans la mesure où, contrairement aux évaluations internes mises en œuvre par les enseignants, elles sont standardisées du point de vue de leur contenu, des conditions de passation, des modalités de correction et des barèmes appliqués. Dès lors, elles permettent de comparer les résultats obtenus par les élèves car elles sont soumises à tous les jeunes scolarisés au sein d'un même degré.

Comme nous l'avons mentionné précédemment, la mesure d'un effet de composition de l'établissement scolaire fréquenté sur les compétences des élèves nécessite un contrôle du niveau antérieur de ces mêmes compétences.

En effet, l'idée est de s'assurer que deux élèves sont semblables en début d'une période de formation, en particulier du point de vue de leur niveau scolaire, pour être en mesure de déterminer que c'est bien l'exposition à un contexte différent qui génère potentiellement, en fin de période de formation, des différences de résultats.

Les élèves de 11^e année des cantons concernés ont été soumis aux tests PISA au printemps 2012. Idéalement il aurait donc fallu disposer d'une mesure initiale de leurs compétences dans les trois domaines en début d'année scolaire 2011-2012 (soit en août ou septembre 2011). Malheureusement l'OCDE ne fournit pas une telle mesure puisque l'enquête PISA correspond à une évaluation de performance «ponctuelle».

Dès lors, le CIRSE et le SRED ont envisagé d'utiliser les résultats scolaires de mathématiques, langue d'enseignement (français ou italien) et sciences disponibles en fin d'année scolaire 2010-2011 (soit en fin de 10^e année) pour pallier ce manque d'information. Initialement, il avait été prévu de contrôler le niveau initial des élèves qui ont passé PISA à partir des résultats de 10^e aux épreuves de référence qui présentent, en partie, les caractéristiques requises pour remplir cette fonction (notamment en termes de comparabilité des élèves). Si à Genève ces données existent, elles ne sont toutefois disponibles que pour les mathématiques et le français. Au Tessin, en revanche, le CIRSE ne dispose pas de résultats semblables pour 2010-2011 dans la mesure où, dans ce canton, les épreuves standardisées, bien que très développées, ne sont pas réalisées chaque année dans tous les degrés pour chacun des domaines considérés. Etant donné que l'on souhaite construire, à des fins de comparaison, des modèles similaires reposant sur des informations de même nature dans les deux cantons, il est apparu que la seule information commune que l'on puisse mobiliser pour contrôler le niveau initial des élèves correspond aux notes attribuées par les enseignants au cours de l'année scolaire.

De façon plus précise, il s'agit des moyennes annuelles en mathématiques, langue d'enseignement et sciences en fin de 10^e année qui, à Genève comme au Tessin, sont exprimées sur une échelle allant de 0 à 6. Toutefois, garantir la comparabilité du niveau scolaire des élèves sur la seule base de ces moyennes apparaît relativement délicat dans la mesure où, comme l'ont montré certains travaux de recherche, évaluer les acquis des élèves et en rendre compte par l'attribution de notes est un phénomène complexe qui peut être sujet à un certain nombre de biais liés au contexte ou à la classe, à l'enseignant et à ses valeurs ou normes, ou encore aux caractéristiques individuelles des élèves (Leclercq, Nicaise et Demeuse, 2004). Ainsi, d'après Merle (2007), la notation n'est pas directement liée aux performances des élèves et résulte davantage d'un processus de fabrication que de l'application automatique d'un barème. Par ailleurs, ces moyennes annuelles sont le fruit de divers travaux d'évaluation plus ou moins nombreux, plus ou moins longs et reposant sur des parties de programme de longueur variable.

À Genève, des travaux antérieurs ont également montré que les enseignants du Cycle d'orientation n'utilisent pas l'échelle des notes de la même manière d'une filière de formation à l'autre (Petrucci, Soussi, Rastoldo, Guilley et Nidegger, 2015).

Pour l'ensemble de ces raisons, il est apparu indispensable d'ajouter dans la modélisation d'autres informations servant à contrôler le niveau initial. Ainsi des profils d'élèves en fin de 10^e ont été constitués sur la base de la filière de formation pour le canton de Genève et sur la base des niveaux d'enseignement en mathématiques et en allemand pour le Tessin (voir 3. Rappel de quelques éléments contextuels).

De façon détaillée, les variables retenues pour le profil de 10^e sont les suivantes :

- Pour le Tessin :
 - Deux niveaux A : l'élève est scolarisé dans les deux niveaux les plus exigeants pour l'allemand et les mathématiques. Cette modalité (référence) correspond aux élèves qui sont scolairement les plus performants.
 - Niveaux mixtes : l'élève est inscrit dans un cours de niveau de base en allemand et un niveau plus exigeant en mathématiques ou inversement.
 - Deux niveaux B : l'élève est inscrit dans des cours de niveaux de base en allemand et en mathématiques. Cette modalité correspond aux élèves dont le niveau scolaire est le moins élevé.

- Pour Genève :
 - Regroupement A ou assimilé : l'élève est scolarisé dans la filière aux exigences élevées (référence).
 - Regroupement B ou assimilé : l'élève est scolarisé dans la filière aux exigences moyennes ou élémentaires.

Ainsi, dans l'analyse multiniveau, le contrôle du niveau initial d'un élève se fera à la fois par la moyenne annuelle obtenue dans une discipline et par le profil en fin de 10^e. En d'autres termes, on considérera qu'au sein d'un même canton deux élèves en début de 11^e sont semblables du point de vue de leur niveau scolaire dans un domaine donné lorsqu'ils ont, en fin de 10^e, la même moyenne annuelle et le même profil.

Enfin, concernant le domaine des sciences, rappelons que dans le test PISA administré aux élèves, environ la moitié des items correspondent à de la physique et l'autre moitié à de la biologie.

Dans le canton du Tessin, l'enseignement des sciences naturelles englobe ces deux disciplines et la note dispensée par les enseignants porte donc sur un champ relativement proche de ce qui est évalué dans PISA.

En revanche, il n'en va pas de même à Genève où la physique n'est pas enseignée à l'ensemble des élèves de 10^e puisque seuls les élèves de la section LS (Littéraire-Scientifique) avec un profil S (Sciences) peuvent se voir dispenser cet enseignement. Dès lors, seules les notes de biologie, disponibles pour l'ensemble des élèves, ont pu être utilisées pour contrôler le niveau initial en sciences dans ce canton, ce qui constitue une limite des résultats qui seront présentés ultérieurement pour ce domaine, notamment en termes de comparabilité stricte avec les résultats du Tessin.

Par ailleurs, avant d'utiliser les éléments qui précèdent dans une analyse multiniveau plus complexe, il nous est apparu souhaitable d'avoir une idée du sens et de l'intensité de la relation qui existe entre les moyennes annuelles de fin de 10^e et les scores obtenus dans les différents domaines à PISA.

Pour ce faire, nous avons calculé les coefficients de corrélation¹² présentés dans la Figure 5. On relève tout d'abord, sans grande surprise, qu'à Genève comme au Tessin tous ces coefficients sont positifs, ce qui traduit simplement le fait que, quel que soit le domaine considéré, il existe un lien positif entre la moyenne annuelle de fin de 10^e et le score obtenu à PISA. Autrement dit, en règle générale, plus la moyenne de fin de 10^e est élevée et mieux on réussit à PISA et inversement. Ce constat se vérifie, en outre, quel que soit le profil de 10^e considéré.

D'après les critères communément admis en sciences sociales, la force du lien peut être globalement considérée comme allant de modérée à forte selon le canton et le domaine¹³. On relèvera que le lien est tendanciellement moins fort chez les élèves ayant un profil de 10^e correspondant à ceux qui sont scolairement les moins performants.

Enfin, il apparaît qu'à Genève la corrélation entre les notes de biologie et les scores de sciences obtenus dans PISA est du même ordre de grandeur que celle observée au Tessin (0.56 contre 0.54) alors même que les notes de sciences dans ce dernier canton correspondent à un enseignement de biologie et de physique plus conforme, en théorie, à ce qui est évalué dans PISA.

Dès lors pour le domaine des sciences, même si la comparaison des résultats entre Genève et le Tessin est, au sens strict, rendue plus difficile par ces notes de 10^e d'une nature un peu différente, on peut vraisemblablement considérer dans les analyses qui suivront que l'utilisation des seules notes de biologie à des fins de contrôle du niveau initial en sciences ne constitue pas, pour le canton de Genève, un choix moins judicieux sur le plan statistique que celui opéré pour le Tessin avec les notes de sciences.

¹² Pour chacun des domaines évalués, ces coefficients correspondent à la moyenne des coefficients de corrélation de Pearson calculés entre chacune des cinq *plausible values* disponibles dans PISA et la note annuelle moyenne de l'élève.

¹³ Lorsque la valeur absolue arrondie du coefficient se situe entre 0.4 et 0.5, on parle de lien modéré et on parlera de lien fort lorsque la valeur est comprise entre 0.6 et 0.7.

Figure 5 – Corrélations entre moyennes annuelles de 10^e et scores PISA selon le canton, le domaine et le profil de 10^e année

Ensemble des élèves		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e CO	Français	0.53		
	Mathématiques		0.62	
	Biologie			0.56

Ensemble des élèves		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e année	Italien	0.52		
	Mathématiques		0.44	
	Sciences			0.54

Regroupement A et assimilé (10e CO)		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e CO	Français	0.52		
	Mathématiques		0.65	
	Biologie			0.53

2 Niveaux A (10e année)		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e année	Italien	0.38		
	Mathématiques		0.44	
	Sciences			0.41

Regroupement B et assimilé (10e CO)		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e CO	Français	0.13		
	Mathématiques		0.34	
	Biologie			0.12

2 Niveaux B (10e année)		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e année	Italien	0.23		
	Mathématiques		0.36	
	Sciences			0.27

Niveaux A et B "Mixtes" (10e année)		Scores Pisa		
		Lecture	Mathématiques	Sciences
Moyennes annuelles 10e année	Italien	0.18		
	Mathématiques		0.16	
	Sciences			0.26

Source : SRED et CIRSE

4.5 Les caractéristiques des établissements

4.5.1 Les caractéristiques sociodémographiques (d'après fichiers cantonaux)

À partir des informations individuelles enregistrées dans la base de données scolaires du DIP, pour le canton de Genève, on peut connaître les principales caractéristiques sociodémographiques de tous les élèves scolarisés dans les 19 établissements du Cycle d'orientation qui ont participé à l'enquête, soit 12'356 élèves (situation au 31.12.2011). Ces informations sont extraites de la «fiche élève» que renseignent les parents en début d'année scolaire.

On notera toutefois que dans PISA, une faible proportion d'élèves est exclue de la passation des tests en raison d'un manque de maîtrise suffisante de la langue d'enseignement (classes d'accueil) ou de très grandes difficultés scolaires (classes atelier) (Nidegger (éd.), 2014). Si leur exclusion de l'échantillon semble assez légitime du point de vue de la mesure des compétences, ces élèves doivent en revanche être conservés dans la caractérisation des établissements puisqu'ils font partie intégrante de la population scolarisée et qu'ils constituent un des éléments du contexte dont on cherche à mesurer l'effet sur les performances individuelles.

Par ailleurs, les effectifs concernés sont faibles et les structures dans lesquelles ils sont scolarisés sont présentes dans la totalité ou la quasi-totalité des établissements¹⁴. Ainsi leur prise en compte n'induit pas de biais dans la caractérisation des contextes.

Pour le canton du Tessin, des informations analogues ont été extraites de la base de données scolaires du DECS. Elles ont là aussi permis de caractériser toute la population scolarisée dans les 35 établissements de la Scuola Media (soit 12'353 élèves à la rentrée 2011).

De façon plus précise, les variables retenues pour caractériser la composition sociodémographique des établissements des deux cantons concernés sont les suivantes :

¹⁴ Il y a des classes d'accueil dans tous les établissements du CO et des classes atelier dans 16 établissements sur 19.

- **Proportion de filles**
- **Proportion d'élèves allophones** : pour Genève, il s'agit des élèves dont la première langue parlée n'est pas le français. Il convient d'avoir à l'esprit que d'après cette définition, le fait d'être allophone traduit, dans le contexte genevois, des réalités parfois très différentes. En effet, si dans certains cas, il s'agit d'une véritable absence de maîtrise de la langue française, dans d'autres cela correspond plutôt à une appartenance culturelle. Ainsi, certains élèves considérés comme allophones peuvent avoir effectué toute leur scolarité à Genève et être parfaitement bilingues. Dans le canton du Tessin, les élèves allophones sont ceux dont la langue maternelle diffère de l'italien et des dialectes tessinois/italiens.
- **Proportion d'élèves migrants** : le statut migratoire est une variable composite construite pour le seul canton de Genève. Elle croise la première langue parlée avec la date d'arrivée dans le canton. Plus précisément, il s'agit en quelque sorte d'affiner l'information fournie par la variable précédente en considérant que les élèves allophones nés à Genève ou arrivés dans le canton avant l'âge de la scolarité obligatoire (soit 4 ans) sont assimilés à des francophones. Dans les autres cas de figure, les élèves dont la première langue parlée diffère du français restent considérés comme allophones. Pour le Tessin, l'absence d'information sur la date d'arrivée dans le canton n'a pas permis de construire une variable similaire.
- **Proportion d'élèves étrangers** : il s'agit de la proportion d'élèves n'ayant pas la nationalité suisse.

Ces variables, exprimées en pourcentages, ont dans un premier temps été utilisées dans l'analyse sous leur forme « brute ».

Dans un second temps, elles ont été utilisées en se plaçant dans une perspective catégorielle qui a consisté à répartir les établissements en trois groupes sur la base des déciles : les 10% d'établissements ayant les proportions les plus faibles sont considérés comme ayant des valeurs inférieures à la norme et, à l'inverse, les 10% d'établissements ayant les proportions les plus élevées sont considérés comme ayant des valeurs supérieures à la norme. Le reste des établissements (soit 80%) sont considérés comme étant dans la moyenne.

- **La composition sociale de l'établissement** :
 - À Genève, la catégorie socioprofessionnelle (csp) de l'élève est définie sur la base de la profession des parents. Ainsi on dispose d'une variable qui comporte initialement 15 modalités que l'on regroupe en trois modalités :
 - csp favorisée : cadres supérieurs et dirigeants,
 - csp moyenne : employés, cadres intermédiaires et petits indépendants,
 - csp défavorisée : ouvriers, divers et sans indication.
 - Pour chaque établissement scolaire on connaît donc la répartition de la population scolarisée selon les trois modalités de l'origine sociale décrites ci-dessus.
 - À partir de l'information qui précède, plusieurs pistes ont été explorées pour caractériser la composition sociale des établissements du Cycle d'orientation¹⁵. Au final nous avons retenu l'indicateur de déplacement utilisé par Duru-Bellat, Danner, Le Bastard-Landrier et Piquée (2004). Cet indicateur se définit de la manière suivante pour un établissement i :

$$Depl_i = | \%Défavo_i - \%Défavo_{Total} | + | \%CSP Moy_i - \%CSP Moy_{Total} | + | \%Favo_i - \%Favo_{Total} |$$

Il s'agit plus précisément d'évaluer le pourcentage d'élèves qu'il faudrait déplacer pour réaliser une situation d'équi-répartition (homogénéité de la répartition d'un groupe dans les différents établissements étudiés). Autrement dit, on obtient le pourcentage d'élèves qu'il faudrait déplacer pour que chaque établissement possède une structure sociale comparable à celle qui prévaut dans l'ensemble de la population. Toutefois, cet indicateur ne dit rien des groupes qui sont sur- ou sous-représentés. En effet, on peut avoir des établissements avec des valeurs de l'indice très proches mais des tonalités sociales très différentes.

¹⁵ Essais de constitution de typologies synthétiques à partir d'analyses de classification ascendante hiérarchique (« clusters »), calcul de différents indices dont le *coefficient social de l'établissement* élaboré par le DIP pour l'attribution des ressources. Ce coefficient est égal à (Coeff = % d'ouvriers - % d'employés - (2x%cadres sup et dirigeants)). Il peut donc varier de 1 (il n'y a que des ouvriers dans l'établissement considéré) à -2 (il n'y a que des élèves issus des milieux favorisés).

Dans un premier temps, à partir des valeurs obtenues pour chaque établissement, on calcule la valeur moyenne de l'indice. On considèrera qu'un établissement a une composition qui diffère de la moyenne lorsque la valeur de l'indice est supérieure à la valeur moyenne plus un écart-type. Ainsi, dans le canton de Genève, quatre établissements sur 19 peuvent être considérés comme ayant une composition sociale qui diffère de la norme. L'observation de la proportion d'élèves issus d'un milieu favorisé dans chacun d'entre eux permet de déterminer que deux établissements ont une tonalité sociale plutôt défavorisée et, qu'à l'inverse, deux autres peuvent être considérés comme des établissements plutôt favorisés du point de vue de leur composition sociale. Les autres établissements seront considérés comme des établissements «mixtes».

- Dans le canton du Tessin, on a utilisé une démarche relativement similaire pour caractériser la composition sociale des établissements de la Scuola Media. Concrètement on a, dans un premier temps, mobilisé l'information relative à la profession des parents (codée avec ISCO-08) pour déterminer la proportion d'élèves issus des différents milieux sociaux dans les établissements. Par analogie avec Genève, une catégorisation en trois modalités a été retenue :
 - La première modalité correspond aux «cols blancs hautement qualifiés»
 - La seconde aux «cols blancs peu qualifiés»
 - La troisième à tous les «cols bleus».
- Une première classification de la composition sociale des établissements a été définie à partir de cette catégorisation en trois modalités. Les quatre établissements possédant les proportions les plus élevées d'élèves de type «cols blancs hautement qualifiés» (soit environ 10% des écoles) ont été considérés comme favorisés et, à l'inverse, les quatre établissements dans lesquels ces proportions sont les plus faibles ont été considérés comme défavorisés (là aussi 10% des écoles).
- Dans un second temps, on a calculé l'indice de déplacement mentionné précédemment et on a adopté la même démarche que celle explicitée pour Genève. Cette dernière a permis de confirmer que la typologie des établissements était la même que celle initialement construite sur la seule base de la proportion d'élèves issus des milieux les plus favorisés. En raison de la valeur de l'indice, un 5^{ème} établissement a toutefois été ajouté au groupe des écoles considérées comme étant plutôt défavorisées de point de vue de leur composition sociale. Cette dernière catégorisation est celle qui a été retenue dans l'analyse.
- **Origine géographique** : les cantons de Genève et du Tessin ont des réalités géographiques très différentes. En effet, Genève est un canton que l'on peut qualifier de «canton-ville», dans lequel la distinction rural/urbain a peu de sens. En revanche, la situation au Tessin est beaucoup plus contrastée et justifie que l'on distingue les établissements implantés dans les zones urbaines de ceux situés dans des zones plus rurales. Pour ce faire, les communes d'implantation des établissements de la Scuola Media ont été caractérisées à partir de la nomenclature de l'Office fédéral de statistique sur les niveaux géographiques de Suisse. Un découpage en trois catégories a été retenu :
 - 1 = centre-ville
 - 2 = commune périurbaine
 - 3 = commune rurale.

4.5.2 Les caractéristiques scolaires (d'après fichiers cantonaux)

À partir des informations disponibles dans les bases de données scolaires du DIP et du DECS, on peut définir plusieurs indicateurs relatifs à la composition scolaire des établissements. De façon plus précise, ont été retenues :

- **La proportion d'élèves scolarisés dans la filière aux exigences élevées ou la proportion d'élèves ayant les niveaux les plus exigeants en mathématiques et allemand** : pour le canton de Genève, il s'agit de la proportion d'élèves qui, au sein d'un établissement, sont scolarisés dans le regroupement A (pour les 10^e et 11^e années) ou dans le regroupement R3 (pour la 9^e année) ainsi que des élèves des classes hétérogènes avec un «profil A» (voir 3.1 et 3.2). Cette proportion est de 67.6% pour l'ensemble des établissements du canton et varie de 52.5% à 81.6% selon l'établissement considéré.

Dans le canton du Tessin, à partir de la répartition des élèves dans les différents niveaux de mathématiques et allemand (élèves avec deux niveaux A, un niveau A et un B et deux niveaux B), on calcule la proportion d'élèves ayant un profil «AA» au sein de chaque école. Cette proportion varie de 43% à 66% selon l'établissement considéré.

- **La proportion d'élèves qui ont eu des difficultés dans leur parcours scolaire antérieur :** à partir des données longitudinales contenues dans la base de données scolaires genevoise, on peut construire quatre marqueurs de difficultés antérieures dans le parcours scolaire des élèves, à savoir le fait d'avoir été confronté au redoublement, à un passage par les classes d'accueil (structures pour non francophones), par les classes atelier (élèves avec de très grandes difficultés scolaires) ou par l'enseignement spécialisé. La proportion d'élèves ayant connu au moins une des difficultés mentionnées précédemment s'élève à 22.6% pour l'ensemble des établissements du canton et varie de 13.4% à 31.7% selon l'établissement considéré.

Pour le canton de Tessin, de telles données n'ont pu être mobilisées.

- **Les proportions d'élèves en avance, à l'heure ou en retard dans leur scolarité :** pour le canton de Genève, ces proportions se définissent par comparaison entre l'âge théorique de la formation suivie et l'âge scolaire de l'élève. De façon plus précise :
 - La proportion d'élèves en avance dans leur scolarité s'élève à 3.1% pour l'ensemble des établissements du canton et varie de 1% à 5.8% selon l'établissement considéré.
 - La proportion d'élèves à l'heure dans leur scolarité s'élève à 75.2% pour l'ensemble des établissements du canton et varie de 67.3% à 80.2% selon l'établissement considéré.
 - La proportion d'élèves en retard dans leur scolarité s'élève à 21.7% pour l'ensemble des établissements du canton et varie de 14% à 31.4% selon l'établissement considéré.
 - Dans le canton du Tessin de telles données n'ont pu être mobilisées dans la présente recherche.
- **La proportion d'élèves non promus :** il s'agit de la proportion d'élèves qui, quelle que soit l'année de scolarité, ont des résultats scolaires insuffisants en fin d'année et ne remplissent pas les conditions de promotion définies par les autorités scolaires. Dans le canton de Genève ces élèves sont, en règle générale, orientés dans un autre type de regroupement l'année suivante. Dans certains cas de non-promotion ou d'un résultat insuffisant dans une discipline, les directions d'établissement peuvent prendre des mesures particulières telles que la dérogation, l'essai ou le redoublement. Au Cycle d'orientation, la proportion d'élèves non promus s'élève à 17.7% pour l'ensemble des établissements du canton et varie de 11.3% à 25.5% selon l'établissement considéré.
 Au Tessin, la proportion de non-promus est d'environ 2% pour l'ensemble de la Scuola Media et varie de 0% à 6% selon l'établissement considéré. On notera que cet indicateur peut être le reflet du niveau scolaire de la population scolarisée dans un établissement mais qu'il peut aussi traduire le caractère plus ou moins sélectif de la politique mise en œuvre dans un établissement.

Comme pour les variables relatives aux caractéristiques sociodémographiques, les variables qui renseignent sur la composition scolaire d'un établissement sont exprimées en pourcentages. Dans la mise en œuvre des analyses, elles ont fait également l'objet des deux stratégies de traitement de l'information qui ont été appliquées aux variables des caractéristiques sociodémographiques (introduction de la variable «brute» et catégorisation en trois groupes de la variable sur la base des déciles, voir 4.5.1 pour plus de détails).

4.5.3 Les caractéristiques organisationnelles (d'après fichiers cantonaux)

- **Taille des établissements :** en moyenne, dans le canton de Genève, un établissement du Cycle d'orientation accueille 650 élèves au 31.12.2011. On considérera que lorsque la taille d'un établissement est comprise dans l'intervalle défini par la valeur moyenne plus ou moins un écart-type, soit entre 583 et 717 élèves, l'établissement est de taille moyenne. Au-delà de 717 élèves, la taille sera considérée comme plutôt supérieure à la moyenne, et en-dessous de 583 élèves comme plutôt inférieure à la moyenne. À partir de ces critères, on dénombre deux établissements de taille supérieure à la moyenne (respectivement 743 et 751 élèves) et trois établissements de taille inférieure (entre 513 et 545 élèves).

Au Tessin, un établissement de la Scuola Media comporte en moyenne 353 élèves, ce chiffre variant de 103 à 615 élèves selon l'établissement considéré. En appliquant la même règle que pour le canton de Genève, on identifie sept établissements que l'on peut considérer comme ayant une taille supérieure à la moyenne (au-delà des 495 élèves) et quatre établissements de taille plutôt inférieure (moins de 225 élèves).

- **Taille moyenne des classes** : dans la présente étude, le CIRSE et le SRED ont souhaité identifier des établissements qui auraient éventuellement des politiques de constitution des classes un peu différentes de ce qui se fait, en moyenne, dans le canton (tendance à constituer des classes un peu plus petites ou un peu plus grandes).

À Genève, le nombre moyen d'élèves par classe dépend principalement de la filière de formation (autrement dit, il s'agit d'un critère fixé par voie réglementaire). Dès lors, le calcul d'un indicateur relatif à la taille des classes d'un établissement nécessite la prise en compte du poids de chaque filière. En effet, un simple calcul de moyenne de la taille des classes ne reflèterait que des différences de composition par filière d'un établissement à l'autre puisque, sans surprise, les établissements dont la taille moyenne des classes est la plus faible sont ceux dans lesquels on dénombre proportionnellement le plus d'élèves dans la filière à exigences moyennes et élémentaires (Regroupement B et R1, R2). Par ailleurs, pour réaliser un tel indicateur il convient de ne considérer que les classes ordinaires, les classes atelier, d'accueil et sports et arts, dont les effectifs sont par définition plus faibles, n'étant pas systématiquement présentes dans l'ensemble des établissements. Le principe de l'indicateur repose sur le calcul de l'écart entre une taille moyenne des classes observée et une taille moyenne théorique qui tient compte du poids des différentes filières présentes dans chaque établissement. Là où l'écart est relativement important, cela peut permettre d'identifier des établissements qui mettent en œuvre des politiques de constitution des classes un peu particulières. De façon plus précise, on procède de la manière suivante :

- **Taille moyenne des classes ordinaires observée** : elle correspond au nombre total d'élèves scolarisés dans les classes ordinaires divisé par le nombre total de classes ordinaires. Par exemple, pour le Cycle d'orientation dont les effectifs sont détaillés dans la **Figure 6**, la taille moyenne observée = $595/33 = 18.03$ élèves par classe.
- **Taille moyenne théorique des classes** : on connaît, pour chaque établissement, le nombre de classes ordinaires et leur type (donc la répartition des classes ordinaires en fonction de leur type). Par ailleurs, on connaît également la taille moyenne de chaque type de classe observée sur l'ensemble du canton. Par exemple, pour l'établissement n°1 de la Figure 6, la taille moyenne théorique est égale à :

$$\text{Taille moy. théorique} : (0.394 \times 22.7) + (0.242 \times 14.4) + (0.061 \times 9.8) + (0.121 \times 13.7) + (0.182 \times 23.2) = 18.90 \text{ élèves par classe}$$

Figure 6 – Exemple de taille moyenne des classes à Genève

		Cycle d'orientation n°1			Ensemble du canton
		Nbre de classes	% classes	Nbre élèves	
Type de classes ordinaires	Regroupement A	13	39.4%	269	Taille moyenne des classes (nbre d'élèves /classe)
	Regroupement B	8	24.2%	111	22.7
	R1	2	6.1%	18	14.4
	R2	4	12.1%	54	9.8
	R3	6	18.2%	143	13.7
Total		33	100.0%	595	23.2

Source : SRED et CIRSE

- À partir des deux valeurs qui précèdent, on calcule l'écart entre la taille moyenne observée et la taille moyenne théorique. On fixe un seuil égal à 0.4 en valeur absolue pour considérer que l'écart est «significatif». À ce stade, il convient de relever que ce choix est relativement arbitraire et discutable. Par ailleurs, on notera que, d'une

manière générale, il existe peu de différences entre valeurs théorique et observée. En effet, à Genève il y a peu de variance dans la taille des classes entre les établissements car cette dernière est fixée par voie réglementaire et les établissements ont assez peu de marge de manœuvre en la matière.

- L'application de la démarche explicitée ci-dessus permet d'identifier deux établissements qui ont des classes dont la taille est, en moyenne, sensiblement plus faible que ce qui se fait dans le canton et, à l'inverse, quatre établissements dont les classes ont, en moyenne, une taille plus importante.
- Au Tessin, en l'absence de filières, la taille moyenne des classes de chaque établissement se calcule plus simplement en rapportant le nombre total d'élèves au nombre total de classes. Ce chiffre varie de 16,4 élèves à 22,1 élèves selon l'établissement considéré, la moyenne cantonale s'élevant à 20,4 élèves. Seuls deux établissements de la Scuola Media ont été considérés comme constituant des classes dont la taille moyenne est plutôt inférieure à la norme (moins de 17 élèves).

4.5.4 Caractéristiques du corps enseignant

Les données contenues dans les fichiers de ressources humaines du DIP et les fichiers GAGI¹⁶ du DECS permettent de disposer d'informations sur le corps enseignant rattaché à un établissement du Cycle d'orientation ou de la Scuola Media¹⁷. De façon plus précise, on peut déterminer :

- **La proportion de femmes dans le corps enseignant** : à Genève, cette proportion s'élève à 57% pour l'ensemble des établissements et varie de 49% à 70% selon l'établissement considéré. Au Tessin cette proportion s'élève, en moyenne, à 58% et varie de 37% à 69% selon l'école considérée.
- **La proportion d'enseignants qui travaillent à temps partiel** : un temps partiel correspond ici aux taux d'activité inférieurs à 95%. On distinguera en outre le temps partiel au sein d'un établissement et le temps partiel calculé sur un taux d'activité global qui cumule l'ensemble des activités assurées par un enseignant. Ainsi, au cours de l'année scolaire 2011-2012, près de 72% des enseignants genevois travaillent à temps partiel dans un établissement (de 62% à 81% selon l'établissement considéré) et 70% ont un taux d'activité global qui correspond à un temps partiel (de 60.4% à 79.4%). Dans le canton du Tessin, le pourcentage d'enseignants travaillant à temps partiel dans un établissement fait lui aussi référence aux enseignants qui, dans une école donnée, travaillent à un taux inférieur à 95%. Ces chiffres varient de 56% à 91% selon l'établissement considéré. Pour le calcul de la part d'enseignants dont le taux d'activité global correspond à un temps partiel, toutes les heures effectuées sont comptabilisées. Ainsi, deux tiers des enseignants de la Scuola Media ont un taux d'activité global qui correspond à un temps partiel (de 50% à 88%). On notera que dans chacun des cantons on dispose des mêmes informations par discipline.
- **La proportion d'enseignants qui travaillent dans plusieurs établissements** : à Genève ceci concerne globalement un peu moins de 5% des enseignants. Toutefois, les différences sont assez importantes d'un établissement à l'autre puisque dans certains cas on ne dénombre aucun enseignant travaillant sur plusieurs sites alors que dans d'autres, près de 10% du personnel exerce son activité dans plusieurs établissements. Au Tessin, l'identification des enseignants qui exercent leur activité sur différents sites est relativement complexe. D'après les informations contenues dans la base de données GAGI on estime que cette proportion varie de 4% à 35% selon l'établissement considéré. On notera par ailleurs que dans les deux cantons considérés, le travail des enseignants sur plusieurs sites est également à mettre en lien avec les disciplines enseignées.
- **Expérience des enseignants** : afin de pouvoir faire un lien plus direct entre les caractéristiques des enseignants et les résultats obtenus à PISA dans les différents

¹⁶ Le fichier GAGI (*Gestione Allievi e Gestione Istituti*) est la base de données dans laquelle le département de l'éducation tessinois (DECS) recueille toutes les informations concernant les élèves et les enseignants.

¹⁷ Le corps enseignant est construit à partir du ou des établissement(s) dans lesquels l'activité professionnelle est exercée.

domaines évalués, il a été nécessaire de détailler certaines informations par discipline, notamment l'expérience des enseignants.

À Genève, cette dernière, exprimée en nombre d'années, est calculée par différence entre la date d'entrée en fonction et la date d'observation. Compte tenu des effectifs relativement faibles d'enseignants par discipline au sein de chaque établissement, nous avons préféré retenir la médiane comme indicateur de tendance centrale plutôt que la moyenne.

À partir de cette information, nous avons souhaité identifier les établissements dont les enseignants peuvent être considérés comme ayant, globalement, une expérience qui diffère un peu de la norme. Pour ce faire nous nous sommes, une fois encore, placés dans une perspective catégorielle basée sur les déciles. Ainsi, nous avons considéré que les 10% d'établissements ayant les valeurs médianes les plus faibles disposaient d'un corps enseignant plutôt moins expérimenté et, à l'inverse, que les 10% d'établissements ayant les valeurs médianes les plus élevées pouvaient être considérés comme ayant des enseignants plus expérimentés. Plus précisément, pour chaque discipline on dispose des éléments suivant:

- **Mathématiques:** le nombre total d'enseignants s'élève à 331 sujets et varie de 11 à 23 selon l'établissement considéré. La valeur médiane de l'expérience est de 9 ans pour l'ensemble des enseignants. Dans les deux établissements où l'expérience est considérée comme un peu plus faible qu'ailleurs, la valeur médiane est de 6 ans. Cette valeur s'élève à 13.5 ans dans les deux établissements où l'expérience peut être considérée comme un peu plus élevée.
- **Français:** le nombre total d'enseignants s'élève à 410 sujets et varie de 15 à 29 selon l'établissement considéré. Comme pour les mathématiques, la valeur médiane de l'expérience est de 9 ans pour l'ensemble des enseignants. Dans les deux établissements où l'expérience est considérée comme un peu plus faible qu'ailleurs, les valeurs médianes sont respectivement de 1.0 et 3.5 ans. Dans les deux établissements où les enseignants sont considérés comme plus expérimentés, les valeurs médianes sont respectivement de 17.5 et 21.5 ans.
- **Biologie:** le nombre total d'enseignants s'élève à 148 sujets et varie de 4 à 11 selon l'établissement considéré. La valeur médiane de l'expérience est de 8.5 ans pour l'ensemble des enseignants. Dans les deux établissements où les enseignants sont considérés comme un peu moins expérimentés, les valeurs médianes sont respectivement de 3.0 et 4.0 ans alors qu'elles s'élèvent à 17.0 et 24.0 ans dans les deux établissements où le corps enseignant peut être considéré comme un peu plus expérimenté.
- **Au Tessin, on a pu utiliser l'âge et l'ancienneté des enseignants.** Ces deux variables ont été construites à partir de la date de référence du 31.08.2011 pour chaque discipline. De façon plus précise on dispose des éléments suivants :
 - **Mathématiques :** le nombre total d'enseignants s'élève à 229 sujets et varie de 2 à 11 selon l'établissement considéré. En moyenne leur ancienneté s'élève à 15,4 ans (de 5 à 30 ans selon l'établissement).
 - **Italien :** le nombre total d'enseignants s'élève à 269 sujets et varie de 3 à 15 selon l'établissement considéré. En moyenne leur ancienneté s'élève à 12,4 ans (de 5 à 24 ans selon l'établissement).
 - **Sciences :** le nombre total d'enseignants s'élève à 153 sujets et varie de 1 à 8 selon l'établissement considéré. En moyenne leur ancienneté s'élève à 14,4 ans (de 2,5 à 28 ans selon l'établissement).

Le traitement des informations sur les caractéristiques organisationnelles (à l'exception de la variable de l'expérience des enseignants dans le canton de Genève) dans la modélisation a été effectué en suivant deux stratégies similaires à celles évoquées précédemment :

- **Stratégie 1 :** on introduit les variables sous leur forme «brute».
- **Stratégie 2 :** pour chaque variable, on se place dans une perspective catégorielle en répartissant les établissements en trois groupes sur la base d'un découpage qui repose sur le critère suivant :
 - La valeur de l'indice est comprise dans l'intervalle défini par la valeur moyenne plus ou moins un écart-type : on considère que la valeur de l'indice est «dans la moyenne» de ce qui s'observe sur l'ensemble des établissements du canton.

- La valeur de l'indice est supérieure à la valeur moyenne plus un écart-type : on considère que la valeur de l'indice est «plutôt supérieure» à ce qui s'observe sur l'ensemble des établissements du canton.
- La valeur de l'indice est inférieure à la valeur moyenne moins un écart-type : on considère que la valeur de l'indice est «plutôt inférieure» à ce qui s'observe sur l'ensemble des établissements du canton.

4.5.5 Climat et gestion de l'établissement (d'après enquête PISA 2012)

Les variables utilisées ici ont été construites par l'OCDE (2014) sur la base des informations collectées par le biais des questionnaires adressés aux chefs d'établissement dans l'enquête PISA 2012 (voir Figure 7). Une description détaillée de ces indices est disponible dans la documentation fournie par l'OCDE (2014) et le lecteur intéressé pourra s'y reporter. Seul un tableau synthétique sera présenté ici. On y détaillera principalement la façon dont ces indices s'interprètent et les items retenus dans leur élaboration. Dans la phase de spécification des modèles, ces variables seront introduites en utilisant les deux stratégies décrites pour les variables des caractéristiques organisationnelles (introduction variable «brute» et catégorisation en trois groupes selon le critère de la valeur moyenne plus ou moins un écart type).

Figure 7: Variables collectées par le questionnaire adressé aux chefs d'établissement PISA 2012

Libellé	Items utilisés dans l'indice	Interprétation
% d'enseignants pleinement qualifiés		Varie de 0 à 100%
% d'enseignants diplômés de niveau 5A de la CITE		Varie de 0 à 100%
% d'enseignants de mathématiques diplômés de niveau 5A de la CITE		Varie de 0 à 100%
Indice de pénurie d'enseignants qualifiés	L'enseignement que votre établissement est à même de dispenser est-il affecté par les problèmes suivants ? (Pas du tout → Beaucoup). Indice basé sur 4 items qui concernent la pénurie de: i) professeurs de sciences qualifiés; ii) professeurs de mathématiques qualifiés; iii) professeurs de français (Ge) ou d'italien (Ti) qualifiés; iv) professeurs qualifiés dans d'autres matières	Plus la valeur de l'indice est élevée et plus le chef d'établissement fait état d'une pénurie d'enseignants qualifiés
Indice activités extrascolaires en mathématiques proposées par l'établissement	Parmi les activités suivantes, lesquelles votre établissement propose-t-il aux élèves de 9e année (11e HarmoS) durant l'année scolaire en cours ? (Oui/Non). Indice basé sur les activités suivantes: i) Club de mathématiques; ii) Compétitions de mathématiques; iii) Club axé sur l'informatique/les technologies de l'information et de la communication; iv) Cours de mathématiques supplémentaires	Plus la valeur de l'indice est élevée et plus l'établissement propose d'activités extrascolaires en maths
Indice de gestion de l'établissement : communication des objectifs de l'établissement et développement des programmes	Fréquence à laquelle vous avez adopté les attitudes suivantes dans l'établissement lors de l'année scolaire passée (Ce n'est pas arrivé → Plus d'une fois par semaine). Il s'agit plus particulièrement des items: i) Je me sers des résultats des élèves pour élaborer les objectifs pédagogiques de l'établissement; ii) Je fais en sorte que les activités de formation continue des enseignants soient en accord avec les objectifs d'enseignement de l'établissement; iii) Je veille à ce que les enseignants travaillent en accord avec les objectifs pédagogiques de l'établissement; iv) Je discute des objectifs pédagogiques de l'établissement avec les enseignants lors de réunions collégiales	Les valeurs plus élevées de l'indice traduisent une implication plus importante du chef d'établissement
Indice de gestion de l'établissement : direction pédagogique	Fréquence à laquelle vous avez adopté les attitudes suivantes dans l'établissement lors de l'année scolaire passée (Ce n'est pas arrivé → Plus d'une fois par semaine). Il s'agit plus particulièrement des items: i) J'encourage les méthodes d'enseignement fondées sur les résultats de recherches récentes en éducation; ii) Je complimente les enseignants dont les élèves participent activement aux activités d'apprentissage; iii) J'attire l'attention des enseignants sur l'importance du développement de l'esprit critique et du sens social chez les élèves.	Les valeurs plus élevées de l'indice traduisent une implication plus importante du chef d'établissement
Indice de gestion de l'établissement : promotion de l'amélioration de la pédagogie et du développement professionnel	Fréquence à laquelle vous avez adopté les attitudes suivantes dans l'établissement lors de l'année scolaire passée (Ce n'est pas arrivé → Plus d'une fois par semaine). Il s'agit plus particulièrement des items: i) Quand un professeur rencontre un problème dans sa classe, je prends l'initiative d'en discuter avec lui; ii) Je suis attentif aux comportements qui risquent de perturber le travail en classe; iii) Quand un professeur signale un problème rencontré en classe, nous cherchons une solution ensemble.	Les valeurs plus élevées de l'indice traduisent une implication plus importante du chef d'établissement

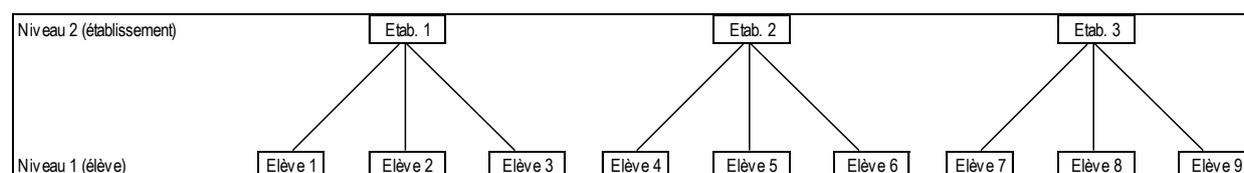
Libellé	Items utilisés dans indice	Interprétation
Indice gestion étab. : participation des enseignants	Fréquence à laquelle vous avez adopté les attitudes suivantes dans l'établissement lors de l'année scolaire passée (Ce n'est pas arrivé → Plus d'une fois par semaine). Il s'agit plus particulièrement des items: i) Je donne au personnel enseignant la possibilité d'intervenir dans des décisions concernant l'établissement; ii) J'incite les enseignants à instaurer une culture de l'établissement axée sur l'amélioration continue; iii) Je demande aux enseignants de participer à l'évaluation des pratiques de gestion.	Les valeurs plus élevées de l'indice traduisent une implication plus importante du chef d'établissement
Qualité des ressources éducatives de l'établissement	L'enseignement que votre établissement est à même de dispenser est-il affecté par les problèmes suivants ? (Pas du tout → Beaucoup). Indice basé sur 6 items qui concernent les problèmes de pénurie ou d'inadéquation en matière de: i) équipement des laboratoires de sciences; ii) matériel pédagogique; iii) ordinateurs pour le travail en classe; iv) connexion à internet; v) logiciels pour le travail en classe; iv) ressources de la bibliothèque.	Les valeurs plus élevées de l'indice traduisent une meilleure qualité des ressources éducatives
Qualité des infrastructures de l'établissement	L'enseignement que votre établissement est à même de dispenser est-il affecté par les problèmes suivants ? (Pas du tout → Beaucoup). Indice basé sur 3 items qui concernent les problèmes de pénurie ou d'inadéquation en matière de: i) locaux et terrains scolaires; ii) installations de chauffage/air conditionné/éclairage; iii) locaux destinés à l'enseignement (par ex., les classes)	Les valeurs plus élevées de l'indice traduisent une meilleure qualité des infrastructures
Indice comportement positif des enseignants	Dans votre établissement, dans quelle mesure l'apprentissage des élèves est-il gêné par les facteurs suivants ? (Pas du tout → Beaucoup). L'indice porte plus précisément sur les items suivant: i) Les élèves ne sont pas encouragés à donner la pleine mesure de leurs capacités ii) Les relations médiocres entre élèves et enseignants; iii) Les enseignants doivent donner cours à des élèves présentant différents niveaux d'aptitudes dans une même classe; iv) Les enseignants doivent donner cours à des élèves d'origines ethniques différentes (c'est-à-dire de langue et culture différentes) dans une même classe; v) Les enseignants ont un niveau d'attentes trop bas à l'égard des élèves; vi) Les enseignants ne satisfont pas aux besoins individuels des élèves; vii) L'absentéisme des enseignants; viii) La résistance du personnel au changement; ix) La sévérité excessive des enseignants à l'égard des élèves; x) Les enseignants arrivent en retard au cours; xi) Les enseignants ne sont pas assez bien préparés pour leurs cours.	Les valeurs plus élevées dénotent un comportement positif des enseignants
Indice comportement positif des élèves	Dans votre établissement, dans quelle mesure l'apprentissage des élèves est-il gêné par les facteurs suivants ? (Pas du tout → Beaucoup). L'indice porte plus précisément sur les items suivant: i) absentéisme des élèves; ii) Les élèves sèchent les cours; iii) Les élèves arrivent en retard en classe; iv) Les élèves ne se présentent pas aux activités scolaires obligatoires (par ex., une journée sportive) ou aux excursions; v) Le manque de respect des élèves envers les enseignants; vi) Les élèves perturbent les cours; vii) La consommation d'alcool ou de substances illicites par les élèves; viii) Les élèves menacent ou brutalisent d'autres élèves.	Les valeurs plus élevées dénotent un comportement positif des élèves
Indice du moral des enseignants	Pensez au personnel enseignant de votre établissement. Dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ? (Tout à fait → Pas du tout). L'indice mesure le degré d'assentiment des chefs d'établissement aux items suivants: i) Le moral des enseignants est très bon dans cet établissement; ii) Les enseignants travaillent avec enthousiasme; iii) Les enseignants sont fiers de cet établissement; iv) Les enseignants attachent beaucoup d'importance aux performances scolaires.	Les valeurs plus élevées dénotent un moral positif des enseignants

5. Présentation du modèle multiniveau

5.1 Quelques rappels méthodologiques sur le modèle multiniveau

Les analyses effectuées dans cette recherche reposent principalement sur les modèles hiérarchiques linéaires (HLM), aussi appelés modèles multiniveaux, qui ont été développés pour traiter de l'influence du contexte sur les individus. Ils s'appliquent à des données possédant une structure hiérarchisée qui comporte plusieurs niveaux. Dans le cas présent, les élèves, qui constituent le premier niveau, sont scolarisés dans des établissements (le contexte), qui représentent le deuxième niveau (Figure 8)¹⁸.

Figure 8 – Exemple de structure hiérarchique des données



Source : SRED et CIRSE

Considérer que les performances d'un élève à PISA dépendent de caractéristiques qui lui sont propres mais aussi de caractéristiques inhérentes à son établissement revient donc à intégrer la structure hiérarchisée des données dans la démarche analytique.

Les modèles de régression classiques par les moindres carrés ordinaires (MCO) sont inadaptés à ce type d'exercice car ils ne permettent de raisonner qu'à un seul niveau. Par ailleurs, ces modèles ne permettent pas non plus de postuler que les relations modélisées varient d'un établissement à l'autre. Les modèles multiniveaux, dérivés des modèles classiques, apportent un certain nombre de réponses à ces limites.

La présente section vise à fournir un aperçu synthétique de la formulation théorique de ces modèles. Le lecteur intéressé par une présentation plus complète pourra se reporter à Bressoux (2010) ou à Snijders et Bosker (1999).

Etant donné que le modèle multiniveau est une extension des modèles de régression classiques par les MCO, on pourra repartir de la formulation de ce modèle pour exposer les particularités de la méthodologie mise en œuvre.

On rappelle donc que traditionnellement, le modèle de régression classique est exprimé dans sa formulation la plus simple sous la forme :

$$y_i = \beta_0 + \beta_1 x_i + e_i \quad \text{avec} \quad e_i \approx N(0, \sigma^2) \quad (1)$$

Où :

- y_i est le score obtenu à PISA par un élève i , quel que soit l'établissement dans lequel il est scolarisé
- x_i est une caractéristique sociodémographique ou scolaire d'un élève i , quel que soit l'établissement dans lequel il est scolarisé. Il s'agit, par exemple, du niveau initial de l'élève dans le domaine évalué par PISA.
- e_i correspond à l'erreur (ou perturbation). Il s'agit de l'écart entre la réalité, autrement dit le score réellement obtenu par l'élève i à PISA, et l'approximation de ce même score fournie par le modèle. Cette variable aléatoire comprend l'ensemble des déterminants du phénomène étudié qui ne sont pas explicitement pris en compte dans le modèle. La relation

¹⁸ On est conscient qu'il y a également le niveau classe qui façonne davantage le contexte d'apprentissage. Cependant, nous n'avons pas les données pour pouvoir faire une analyse aussi avec ce niveau.

entre X et Y n'est pas totalement déterministe. En effet, on conçoit assez facilement qu'il ne suffit pas de connaître le niveau initial d'un élève en fin de 10^e pour prévoir avec certitude le score qu'il obtiendra à PISA une année plus tard. Chaque élève est un cas particulier dont le résultat au test s'écarte plus ou moins du résultat prédit par le modèle théorique.

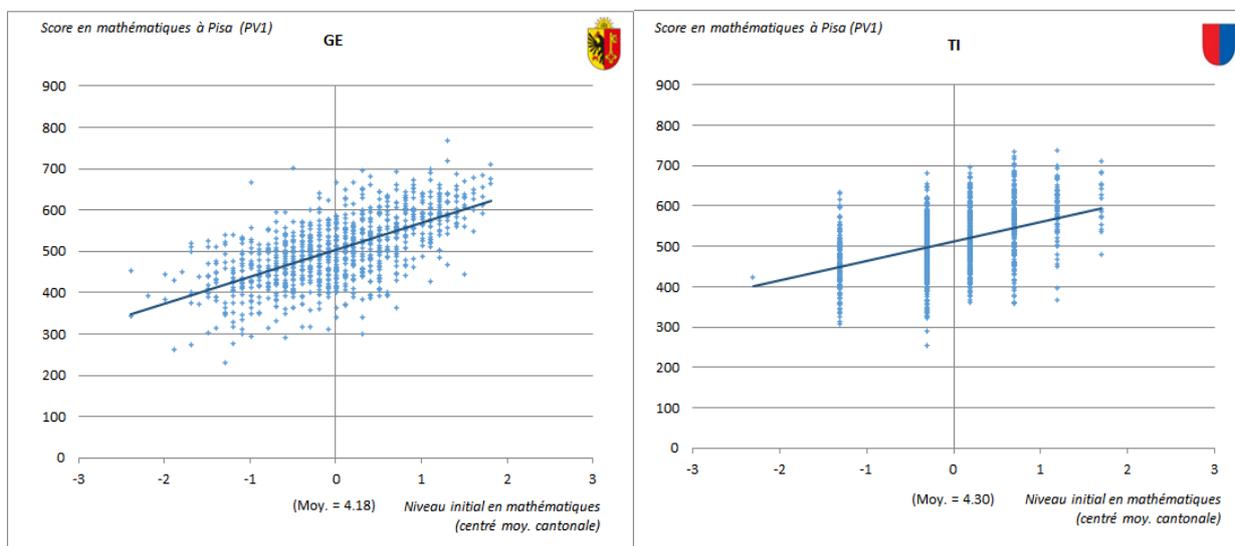
Ce qui nous intéresse dans la démarche présentée ci-dessus, c'est de mesurer correctement l'influence des variables figurant dans le modèle sur le phénomène étudié, en l'occurrence l'effet du niveau initial (X) sur le score obtenu à PISA (Y). Autrement dit, on va chercher à estimer les paramètres β_0 et β_1 .

Ces paramètres correspondent respectivement à :

- **la constante (β_0):** c'est la valeur à l'ordonnée pour une abscisse nulle (autrement dit la valeur de Y lorsque X = 0).
Il convient toutefois de relever que, de façon générale, la valeur 0 de X n'est pas toujours clairement interprétable. En effet, la valeur X=0 n'est parfois pas observée, voire elle peut même constituer une aberration. Pour ces raisons, la constante β_0 nous intéresse moins que la pente β_1 . Le coefficient β_0 sert surtout à la réalisation d'estimations ponctuelles. Dans notre étude, on pourra toutefois accorder une interprétation particulière à la valeur X=0 dans la mesure où on travaillera avec un niveau initial centré sur la moyenne de la population étudiée. Dès lors, la constante s'interprète ici comme le score obtenu à PISA par un élève de niveau initial moyen.
- **la pente (β_1):** traduit l'effet de la variation d'une unité de X sur Y. Dans notre cas, lorsque le niveau initial varie d'une unité, le score à PISA varie de β_1 unités.

L'estimation des deux paramètres présentés ci-dessus se fera par la méthode des moindres carrés ordinaires (MCO) qui permettra de trouver les valeurs de β_0 et β_1 qui minimisent la somme des carrés des erreurs (i.e. la somme du carré des écarts entre la réalité et son approximation par le modèle). D'un point de vue graphique, il va s'agir de trouver la droite qui passe «le plus près possible» de tous les points du nuage qui représente la relation qui existe, dans chaque canton, entre le niveau initial d'un élève de 11^e, mesuré par ses notes en fin de 10^e¹⁹, et le score qu'il a obtenu à PISA (Figure 9).

Figure 9 – Droite de régression par les MCO



Source : SRED et CIRSE

¹⁹ Les moyennes annuelles sont arrondies au demi-point pour le Tessin et au dixième pour Genève.

Dans la formulation la plus simple du modèle multiniveau, on va postuler que la constante β_0 est aléatoire, autrement dit qu'elle peut varier d'un établissement à l'autre. Dans ce cadre, un élève qui possède un niveau «initial» de mathématiques donné n'obtient donc pas nécessairement le même résultat à PISA en fonction de l'établissement qu'il fréquente.

Le modèle est alors défini de la manière suivante:

$$\text{Au niveau 1 (élève):} \quad y_{i_j} = \beta_{0_j} + \beta_1 x_{i_j} + e_{i_j} \quad \text{avec} \quad e_{i_j} \approx N(0, \sigma^2) \quad (2)$$

$$\text{Au niveau 2 (établissement):} \quad \beta_{0_j} = \beta_0 + \gamma_0 z_j + u_{0_j} \quad \text{avec} \quad u_{0_j} \approx N(0, \sigma_0^2) \quad (3)$$

où

- y_{i_j} est le score obtenu à PISA par un élève i scolarisé dans l'établissement j ;
- x_{i_j} est une caractéristique sociodémographique ou scolaire d'un élève scolarisé dans l'établissement j . Comme dans l'exemple précédent, il s'agira ici du niveau initial de l'élève dans le domaine évalué par PISA;
- z_j est une caractéristique de l'établissement j , par exemple la proportion d'élèves issus d'un milieu défavorisé.

D'un point de vue statistique, la volonté de permettre à des établissements d'être diversement «efficaces», autrement dit d'avoir des constantes β_{0_j} différentes, se matérialise par l'utilisation d'une variable aléatoire u_{0_j} dans la spécification de la variable β_{0_j} au deuxième niveau (équation (3)). Le coefficient β_0 représente alors une constante «moyenne» pour l'ensemble des établissements.

Le modèle combiné, qui intègre les deux niveaux (équations (2) et (3)), peut alors s'exprimer ainsi :

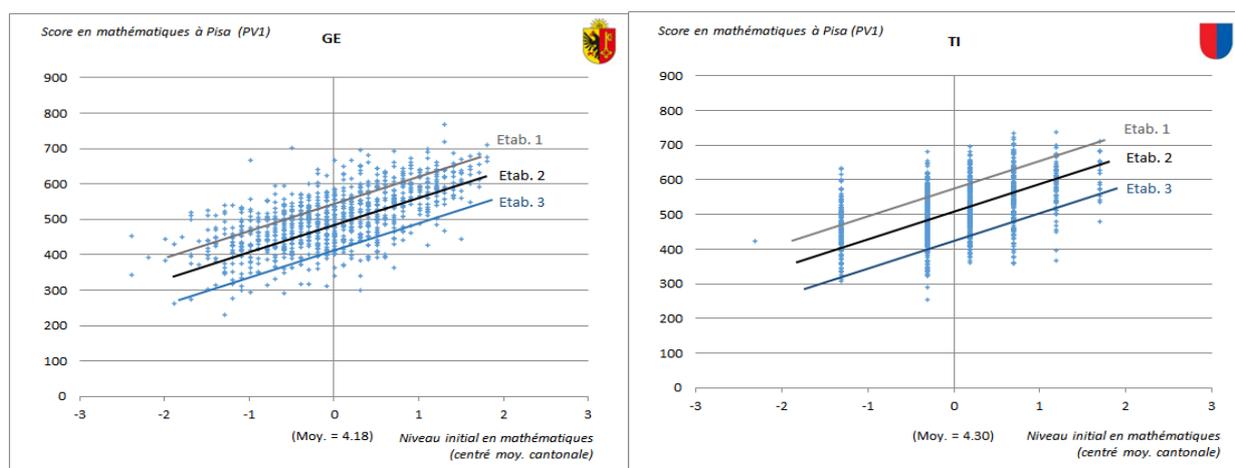
$$y_{i_j} = \beta_0 + \beta_1 x_{i_j} + \gamma_0 z_j + u_{0_j} + e_{i_j} \quad (4)$$

Dans le modèle ci-dessus, ce sont toujours les paramètres $\beta_0, \beta_1, \gamma_0, \sigma^2$ et σ_0^2 que l'on cherche à estimer. La structure complexe des erreurs rend toutefois inapplicable la méthode d'estimation classique de régression par les MCO et rend nécessaire l'utilisation de méthodes d'estimation par le maximum de vraisemblance.

Par ailleurs, d'après l'équation (4) ci-dessus on constate que le modèle multiniveau peut être décomposé en une partie fixe (les coefficients de régression β_0, β_1 et γ_0) et une partie aléatoire composée de deux variances: la variance individuelle σ^2 (niveau élève) et la variance de la constante σ_0^2 (niveau établissement). C'est cette partie aléatoire qui est fondamentale et qui distingue le modèle multiniveau du modèle de régression par les MCO. Dans le modèle multiniveau, la partie aléatoire apparaît en effet à chacun des niveaux de la structure hiérarchisée des données, c'est-à-dire au niveau des élèves et au niveau des établissements. Dans le modèle MCO, l'effet aléatoire n'intervient qu'au niveau des élèves, c'est-à-dire que l'on suppose que σ_0^2 est nulle.

D'un point de vue graphique, si le modèle de régression classique par les MCO se traduit par une droite de régression unique pour l'ensemble des élèves et des établissements, le modèle multiniveau fournit quant à lui une droite de régression pour chaque établissement j de la population d'élèves considérée. Dans le cas présenté ci-dessus, seule la constante est supposée aléatoire, ce qui implique que toutes les droites obtenues sont parallèles entre elles. Elles ne diffèrent que par leur origine (Figure 10).

Figure 10 – Droites de régression dans le modèle multiniveau – Constante aléatoire



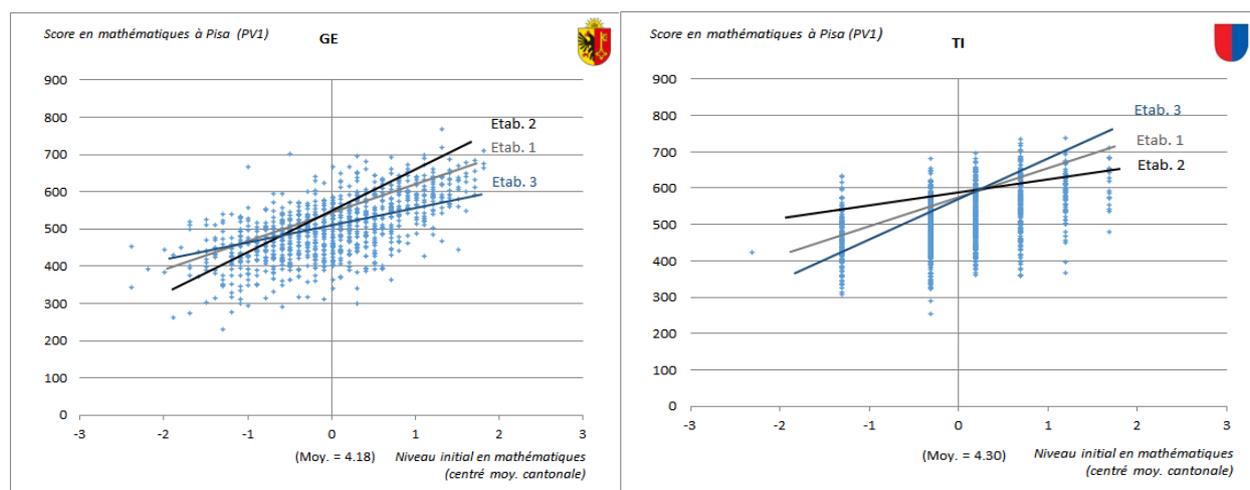
Source : SRED et CIRSE

Une formulation plus complexe du modèle permet aussi de supposer que les pentes sont aléatoires, autrement dit que les relations entre le score obtenu à PISA et le niveau initial ne sont pas les mêmes dans tous les établissements. Dans l'équation (2) on n'aura alors non plus β_1 mais β_{1j} avec une équation supplémentaire, au niveau 2, du type

$$\beta_{1j} = \beta_1 + \gamma_1 z_j + u_{1j} \text{ et } u_{1j} \approx N(0, \sigma_1^2) \quad (5)$$

Cette formulation du modèle permet de modéliser une certaine efficacité différentielle qui correspond au fait que certains établissements feraient mieux réussir certaines catégories d'élèves que d'autres. En plus, elle permet également d'estimer d'éventuels effets d'interaction entre les caractéristiques individuelles et celles de l'établissement. Ces effets traduisent l'idée que l'effet d'une caractéristique de l'environnement peut varier en fonction d'une caractéristique individuelle. Sur le plan graphique, on obtient là encore une droite de régression pour chaque établissement j de la population d'élèves considérée, mais toutes ces droites ne sont désormais plus parallèles entre elles (Figure 11).

Figure 11 – Droites de régression dans le modèle multiniveau – Constante et pente aléatoires



Source : SRED et CIRSE

En termes de restitution de résultats, il convient de relever qu'à l'issue d'une analyse multiniveau on communiquera traditionnellement :

- les effets fixes moyens, autrement dit l'estimation des paramètres β_0, β_1 (et γ_0). Ils correspondent, en quelque sorte, à la constante et à la pente d'une droite de régression moyenne autour de laquelle se distribuent les droites de régression correspondant aux établissements;
- les paramètres de variance σ^2 et σ_0^2 pour les effets aléatoires.

5.2 Spécification des modèles et stratégie de mise en œuvre

Comme l'ont mentionné Snijders et Bosker (1999), la phase de spécification d'un modèle est une tâche extrêmement complexe dans laquelle il faut manipuler simultanément des considérations en lien avec le sujet de la recherche effectuée et des considérations statistiques, la finalité étant de proposer un modèle qui décrit correctement les données sans complications inutiles.

La démarche analytique induite par la mise en œuvre du modèle multiniveau peut se décomposer en trois étapes distinctes qui vont être détaillées ci-dessous.

Etape 1 : le modèle «vide»

La première étape d'une analyse multiniveau est la réalisation d'un modèle dit «vide» qui n'intègre aucune variable explicative et qui correspond simplement à une décomposition de la variance totale du phénomène étudié en une part de variance inter-groupes et une part de variance intra-groupes.

Il s'agit en fait, dans le cas présent, d'étudier comment la variance totale des résultats obtenus à PISA se répartit entre les différents niveaux qui structurent les données, à savoir les élèves (niveau 1) et les établissements (niveau 2).

Cette étape a principalement pour but de mettre en évidence l'existence d'une variance systématique entre les établissements. Si une telle variance n'existe pas, il est alors inutile d'aller de l'avant avec une analyse multiniveau plus complexe. En effet, on conçoit aisément qu'une démarche qui vise à repérer et, le cas échéant, à expliquer un effet de l'établissement sur les performances de l'élève n'a aucune pertinence dès lors que l'on ne parvient pas à montrer qu'il existe bien des différences de résultats entre établissements.

Le modèle «vide» est donc essentiel puisqu'il délivre une information cruciale sur la répartition de la variance entre les différents niveaux de l'analyse. C'est lui qui va, en quelque sorte, nous renseigner sur l'existence et la taille de l'effet établissement.

Par ailleurs, il convient de relever que ce modèle joue également un rôle central dans le calcul du pouvoir explicatif des différents modèles mis en œuvre puisqu'il servira de référence tout au long des analyses (l'évolution de la part de variance résiduelle au fil des spécifications lui sera toujours rapportée).

Etape 2 : l'introduction des caractéristiques individuelles de l'élève (niveau 1)

Une fois que le modèle «vide» aura été mis en œuvre et que l'on se sera assuré de l'existence d'un effet établissement sur les performances des élèves, on pourra alors intégrer dans le modèle les caractéristiques individuelles de l'élève qui ont été décrites précédemment (voir 4.4 Les caractéristiques des élèves).

Il s'agira plus précisément des caractéristiques sociodémographiques usuelles que sont le genre, l'âge, le statut migratoire, la langue et le statut socio-économique et culturel ainsi que du niveau initial de l'élève mesuré par ses notes et son profil en fin de 10^e.

Afin de faciliter la comparaison des résultats entre les différents domaines (mathématiques, lecture et sciences) et entre les cantons concernés, le CIRSE et le SRED ont fait le choix d'intégrer de façon systématique toutes ces variables de premier niveau dans l'analyse. Elles constituent, par ailleurs, un ensemble relativement complet de prédicteurs individuels du phénomène étudié dont la prise en compte est, nous semble-t-il, conforme aux recommandations formulées par Thrupp et al. (2002) sur la spécification des modèles à mettre en œuvre pour mesurer un éventuel effet de composition de l'établissement sur les acquis des élèves (voir 2. Objectif de la recherche et méthodologie). En outre, il nous paraît également important de rappeler ici que l'éventuel effet des caractéristiques individuelles qui pourrait être mis en évidence à ce stade de l'analyse ne doit pas être interprété comme un lien de causalité (il en sera d'ailleurs de même pour l'effet des caractéristiques de l'établissement qui seront ultérieurement intégrées dans la modélisation). En effet, le modèle multiniveau permet simplement de formaliser la relation qui existe, «toutes choses égales par ailleurs», entre une caractéristique de l'élève et le score que ce dernier a obtenu à l'enquête PISA. Enfin, c'est dans cette étape de l'analyse que l'on doit déterminer quelle est la formulation théorique du modèle retenu. Dans la présente étude, on se situe dans le cas le plus simple où seule la constante est aléatoire. Les effets associés aux caractéristiques individuelles sont considérés ici comme des effets fixes car il n'a pas été possible de montrer que les pentes étaient aléatoires.

Etape 3 : la prise en compte des caractéristiques de l'établissement (niveau 2)

À l'issue des deux premières étapes de l'analyse, on devrait être en mesure de savoir s'il existe un effet de l'établissement sur les scores obtenus à PISA (le modèle «vide» nous indique quelle est la part des différences de résultats imputable à l'établissement) et, le cas échéant, de conclure que l'on est bien en présence d'établissements diversement efficaces. Autrement dit, deux élèves qui ont les mêmes caractéristiques individuelles, et en particulier le même niveau initial dans un domaine donné, ne réussissent pas de la même façon à PISA en fonction de l'établissement qu'ils fréquentent.

Sur le plan statistique, cela se traduit par l'existence d'une variance σ_0^2 statistiquement significative (i.e non nulle) (voir équation (3)). Graphiquement, on se situe alors dans le cas de droites de régression toutes parallèles entre elles (elles ne diffèrent que par la valeur de leurs constantes) (Figure 10).

Dans cette dernière étape de l'analyse, on va tenter d'expliquer pourquoi les établissements sont diversement efficaces en introduisant dans la modélisation leurs caractéristiques qui ont été décrites précédemment (voir 4.5 Les caractéristiques des établissements). C'est donc ici que l'on en saura davantage sur la nature de l'effet établissement et plus particulièrement s'il s'agit d'un effet de composition sociale ou académique ou d'un effet d'une autre nature (caractéristiques «organisationnelles» ou éléments relatifs au climat d'établissement).

En pratique, nous avons mentionné dans ce qui précède que si la taille des échantillons dont disposent les cantons du Tessin et de Genève est globalement suffisante pour mettre en œuvre une analyse multiniveau, on se situe tout de même dans des volumes qui sont minimalement requis pour ce type d'analyse, en particulier du point de vue du nombre d'établissements disponibles (35 au Tessin et 19 à Genève). Dès lors, cela va avoir des répercussions sur la spécification des modèles.

Dans l'analyse de régression multiple, il existe une règle communément admise qui consiste à devoir disposer d'au moins 10 observations pour chaque variable explicative introduite dans la modélisation (Bressoux, 2010). Si on transpose cette règle pour le deuxième niveau des analyses multiniveaux, alors pour une variable qui caractérise l'établissement que l'on intègre dans le modèle, il faut disposer d'au moins 10 établissements.

Certains auteurs avancent même qu'il serait nécessaire d'en avoir le double (*Centre for Multilevel modelling*, 2007). Cette règle empirique ne constitue toutefois pas une contrainte qu'il convient de respecter au sens strict mais elle nous rend attentif au fait qu'il faudra conserver dans la modélisation finale un nombre de variables de niveau 2 très limité, compte tenu des données dont nous disposons.

La stratégie de spécification des modèles retenue sera celle préconisée par Bressoux (2010) qui précise que «dans le cas d'une approche exploratoire, le nombre de variables à inclure dans le modèle n'est pas fixé a priori. Une possibilité consiste alors à diviser le nombre de variables explicatives de niveau 2 en plusieurs groupes de variables conceptuellement distinctes et à réaliser un modèle séparé pour chacun de ces groupes. Ensuite, il reste à construire un modèle général unique ne retenant que les variables qui exercent les effets les plus forts» (p. 325).

Dans la présente étude, on pourra considérer que l'on dispose de quatre groupes de variables qui caractérisent les établissements (voir 4.5 Les caractéristiques des établissements), à savoir :

- des caractéristiques sociodémographiques;
- des caractéristiques scolaires;
- des caractéristiques «organisationnelles»;
- des variables relatives au climat d'établissement.

On construira donc des modèles successifs avec les variables de chacun de ces groupes afin de déterminer quelles sont celles qui ont le plus de poids dans l'explication des différences de résultats. Ensuite, un modèle final qui intègre l'ensemble de ces variables sera élaboré (seul ce modèle sera présenté dans les résultats).

Par ailleurs, dans le traitement des informations relatives à l'établissement, nous nous sommes assez fréquemment placés dans une perspective catégorielle qui a consisté à transformer une variable, le plus souvent exprimée en pourcentages, en plusieurs variables catégorielles. Cette stratégie génère mécaniquement un nombre de variables plus important à introduire dans la modélisation ($n-1$ variables dichotomiques à introduire pour une typologie en n modalités). Dès lors il faudra privilégier, dans la mesure du possible, les autres méthodes de traitement de l'information (variables «brutes», indices, etc.).

6. Présentation des résultats

6.1 Des effets établissement relativement modestes, au Tessin comme à Genève

La réalisation des modèles «vides» qui correspond simplement à une décomposition de la variance totale des scores obtenus à PISA en une part de variance inter-groupes et une part de variance intra-groupes (voir 5.2 Spécification des modèles et stratégie de mise en œuvre) permet de constater qu'au Tessin comme à Genève, une part des différences de résultats est imputable à l'établissement fréquenté. Autrement dit, il existe bel et bien un effet statistiquement significatif de l'établissement sur les performances des élèves, mais ce dernier semble toutefois de taille relativement modeste.

Ainsi on estime qu'en mathématiques, dans chacun de ces deux cantons, environ 7% des différences de résultats sont liées à l'établissement fréquenté (Figure 12). Cela signifie que les différences entre élèves d'un même établissement sont beaucoup plus marquées que les différences entre les élèves d'établissements différents. Les variations du score obtenu à PISA sont donc davantage imputables à des différences individuelles qu'à des différences d'établissements.

On pourra faire le même constat dans le domaine des sciences, où les ordres de grandeur obtenus sont sensiblement les mêmes qu'en mathématiques. En effet, au Tessin la variance inter-établissements représente environ 5% de la variance totale des scores et à Genève elle avoisine les 7%. Seul le domaine de la lecture se distingue pour le Tessin où près de 13% des différences de résultats sont imputables à l'établissement, alors qu'à Genève ce chiffre n'est que de 5%.

La revue de la littérature sur les «effets établissement» proposée par Duru-Bellat et al. (2004) semble montrer que, de façon générale, ces résultats sont assez conformes à ce qui a été observé dans d'autres contextes où ces effets ont une importance modérée. Ainsi, selon les recherches considérées, entre 8% et 15% de la variance des scores des élèves s'expliquent par les différences entre écoles (Elliott, 1996), la plupart des estimations se situant néanmoins dans le bas de cette fourchette. Notons qu'Origoni (2007) avait trouvé une part de variance inter-établissements pour le test PISA 2003 en mathématiques de 7% pour Genève et le Tessin, ce qui, une nouvelle fois, est assez cohérent avec ce qui précède.

On notera enfin que pour le canton de Genève, l'analyse des résultats des élèves aux épreuves de référence de français et de mathématiques en 11^e CO pour l'année 2012, et plus particulièrement de la partie «tronc commun» de ces épreuves qui est commune à toutes les sections²⁰, semble valider l'existence d'un effet établissement de taille relativement modeste. En effet, dans ces épreuves qui ont été passées à peu près au même moment que PISA et par les mêmes élèves (PISA correspond plus exactement à un échantillon de ces élèves), on estime que 4% des différences de résultats de français sont liées à l'établissement et que ce chiffre est légèrement supérieur à 5% pour les mathématiques.

²⁰ On dispose d'un score pour l'ensemble des élèves de 11^e exprimé sur une échelle de 50 points pour le français et de 78 points pour les mathématiques. Par ailleurs, dans l'analyse des résultats aux Evacom réalisée ici, le Collège de Cayla a été exclu du champ afin que le périmètre retenu soit identique à celui de l'enquête PISA.

Figure 12 – Modèles «vides» décomposant les parts de variance inter- et intra-établissements des scores de mathématiques, lecture et sciences selon le canton

	Mathématiques				Lecture				Sciences			
	Tessin		Genève		Tessin		Genève		Tessin		Genève	
	Valeur (s.e)	P-value										
Effets fixes												
Constante	514.31 (4.23)	<0.001	500.58 (5.89)	<0.001	484.47 (5.88)	<0.001	499.37 (5.33)	<0.001	490.05 (3.79)	<0.001	487.25 (5.73)	<0.001
Effets aléatoires												
Niveau 2 (établissement): variance des constantes	433.68	<0.001	465.89	<0.001	946.56	<0.001	340.98	<0.001	317.97	<0.001	430.77	<0.001
Niveau 1 (élève): variance inter-élèves	5626.14		6247.92		6464.31		6641.00		5924.35		5924.70	
Part de variance inter- élèves (en %)	92.8%		93.1%		87.2%		95.1%		94.9%		93.2%	
Part de variance inter- établissements (en %)	7.2%		6.9%		12.8%		4.9%		5.1%		6.8%	
Total	100%		100%		100%		100%		100%		100%	

Source : SRED et CIRSE

6.1.1 Des effets probablement un peu surestimés sur la base des données PISA

La mise en œuvre du modèle «vide», première étape d'une démarche analytique multiniveau, nous a permis de constater qu'au Tessin comme à Genève, il existe bien un effet de l'établissement sur les performances des élèves mais que ce dernier est vraisemblablement de taille modeste. En effet, quel que soit le domaine de compétences considéré, la part de variance des résultats imputable à l'établissement, autrement dit la variance inter-établissements, ne représente que quelques pourcents de la variance totale des scores de mathématiques, sciences ou lecture (Figure 12).

À ce stade de l'analyse, il nous paraît toutefois important de mentionner que les valeurs fournies dans ce qui précède sont potentiellement un peu surestimées sans que l'on soit toutefois en mesure de savoir dans quelle proportion. Ceci s'explique principalement par la structure des données PISA qui ne comportent que deux niveaux, à savoir des élèves scolarisés dans des établissements. Or, les travaux méthodologiques d'Opdenakker et Van Damme (2000) ont montré que dans la modélisation multiniveau, le nombre de niveaux retenus dans l'analyse peut influencer de façon significative les résultats d'une étude et l'estimation des paramètres.

Dans la présente recherche, il existe un niveau intermédiaire, correspondant à la classe, qui n'a pas pu être pris en compte dans l'analyse car, dans PISA, les élèves ne sont échantillonnés qu'en fonction de l'établissement fréquenté. En pratique, la prise en compte d'un tel niveau pourrait se traduire par une diminution de la taille de l'effet établissement puisqu'une partie de la variance inter-établissements que l'on a estimée correspondrait, en réalité, à un effet de la classe dans un modèle à trois niveaux.

Pour illustrer le propos nous allons nous appuyer une nouvelle fois sur l'analyse des résultats des élèves genevois aux épreuves cantonales de références (Evacom) de fin de 11^e pour l'année 2012. Il s'agira plus précisément de comparer la taille de l'effet établissement sur les performances des élèves dans un modèle comportant deux ou trois niveaux (respectivement des élèves et des établissements ou des élèves, des classes et des établissements). Seul le cas des mathématiques sera traité²¹. Par ailleurs, étant donné que l'enseignement secondaire I genevois est organisé en filières dans lesquelles la taille des classes varie de façon conséquente en fonction du niveau des élèves (effectifs ordinaires dans le regroupement A et effectifs réduits dans le regroupement B) (voir 3. Rappel de quelques éléments contextuels), il conviendra de construire un modèle pour une filière de formation donnée afin d'éviter que le niveau classe ne prenne artificiellement trop d'importance. En effet, dans un modèle qui agrégerait les élèves issus de tous les regroupements, le niveau classe traduirait en réalité l'appartenance à une filière. Ainsi seuls les élèves du regroupement A seront retenus ici, soit 2'288 élèves répartis dans 100 classes et 16 établissements.

Sur la base de ces données, la réalisation de modèles «vides» permet de constater qu'entre 2% et 3.5% des différences de résultats des élèves sont imputables à l'établissement fréquenté selon que l'on construit un modèle à trois ou à deux niveaux (Figure 13), l'estimation de la taille de l'effet établissement se révélant plus importante dans un modèle à deux niveaux. Dans les deux cas, les résultats nous conduisent à conclure à l'existence d'un effet établissement de taille plutôt modeste sur les performances des élèves, mais l'ampleur de cet effet se révèle toutefois un peu différente selon que l'on considère l'un ou l'autre des modèles.

Compte tenu de ce qui précède, on peut légitimement supposer qu'il en va de même avec l'estimation de la taille des effets établissement sur les compétences des élèves mesurées par l'enquête PISA dans les cantons du Tessin et de Genève.

²¹ Il s'agit de la partie «tronc commun» de l'épreuve administrée à l'ensemble des élèves de 11^e. Elle correspond à un score exprimé sur une échelle de 78 points.

Figure 13 – Modèles «vides» à deux et trois niveaux décomposant la variance des scores de mathématiques à l'épreuve cantonale de référence (Evacom) de 11^e CO, canton de Genève 2012

	Mathématiques - Evacom 11e CO Genève (échelle de 78 points)	
	Modèle 2 niveaux	Modèle 3 niveaux
Effets aléatoires		
Variance niveau élèves	85.1	79.1
Variance niveau classes		7.1
Variance niveau établissements	3.1	1.8
Variance totale des scores	88.2	87.9
Part de variance élèves (en %)	96.5%	89.9%
Part de variance classes (en %)		8.1%
Part de variance établissements (en %)	3.5%	2.0%
Total	100%	100%

Source : SRED et CIRSE

6.2 Analyse des scores obtenus en mathématiques

Le modèle «vide», première étape de la démarche analytique, a permis de constater qu'à Genève comme au Tessin, environ 7% des différences de résultats obtenus en mathématiques sont imputables à l'établissement fréquenté.

À ce stade de l'analyse, les différences de compétences en mathématiques ne sont toutefois pas expliquées puisque le modèle «vide», comme son nom l'indique, ne comporte aucune variable explicative. L'étape suivante consiste donc à introduire dans la modélisation les caractéristiques individuelles des élèves afin de déterminer quelles sont celles qui possèdent un effet, «toutes choses égales par ailleurs», sur les performances.

Il convient de rappeler ici que l'expression «toutes choses égales par ailleurs» traduit simplement le fait qu'à l'exception de la variable considérée, toutes les autres variables explicatives introduites dans la modélisation sont maintenues constantes. Par ailleurs, le terme d'effet ne doit pas s'interpréter nécessairement comme un lien de causalité: il s'agit plutôt de formaliser la relation qui existe entre une caractéristique de l'élève et son résultat au test PISA en mathématiques.

D'une manière générale, les effets des caractéristiques individuelles des élèves sont assez semblables à Genève et au Tessin. Ainsi, on relève un effet positif statistiquement significatif du genre qui traduit le fait qu'être un garçon constitue dans ces deux cantons, «toutes choses égales par ailleurs», un avantage pour les performances obtenues en mathématiques (Figures 14 et 15).

Par ailleurs, toujours à caractéristiques individuelles identiques, l'âge est associé à un effet négatif. Ainsi, en moyenne, plus l'élève est âgé et moins il obtient des scores élevés à PISA, ce qui semble assez cohérent, puisqu'un âge plus important est bien souvent le signe de difficultés dans le parcours scolaire antérieur.

Concernant le niveau initial de mathématiques, mesuré de façon analogue dans les deux cantons par la note de mathématiques et le profil en fin de 10^e CO, là encore les résultats convergent. D'une part, on relève un effet positif associé à la note que l'on peut interpréter, en quelque sorte, comme le fait que plus les résultats initiaux en mathématiques sont bons et mieux on réussit in fine au test PISA. D'autre part, le profil de fin de 10^e vient nuancer ce constat puisqu'à Genève, le fait d'être initialement scolarisé dans la filière à exigences moyennes et élémentaires constitue, «toutes choses égales par ailleurs», un élément qui tend à faire diminuer les scores obtenus par rapport à ceux des élèves scolarisés dans la filière à exigences élevées (coefficient négatif) (Figure 14).

Figure 14 - Résultats des analyses multiniveaux pour le score de mathématiques à PISA 2012 – Canton de Genève

Résultats des analyses multiniveaux pour le score de mathématiques à PISA 2012 (Canton de Genève)									
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)			Modèle 3 (variables niveaux 1 et 2)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes									
Constante	500.58	(5.898)	<0.001	507.45	(6.299)	<0.001	509.00	(6.369)	<0.001
Niveau initial fin 10 ^e CO									
Note de mathématiques				50.57	(2.980)	<0.001	50.65	(2.978)	<0.001
Profil de 10 ^e CO									
<i>(réf. Regroupement A ou assimilé)</i>									
Regroupement B ou assimilé				-54.53	(7.511)	<0.001	-54.51	(7.543)	<0.001
Sexe									
<i>(réf. Fille)</i>									
Garçon				19.17	(4.301)	<0.001	19.08	(4.302)	<0.001
Age				-17.51	(4.723)	<0.001	-17.45	(4.714)	<0.001
Statut socio-économique et culturel (ESCS)				9.35	(3.230)	0.004	9.4	(3.200)	0.004
Statut migratoire									
<i>(réf. Natif)</i>									
2 ^{nde} /1 ^{ère} génération				-4.75	(6.463)	0.463	-4.53	(6.471)	0.484
Langue									
<i>(réf. Francophone)</i>									
Allophone				1.68	(5.688)	0.768	1.99	(5.620)	0.723
Ancienneté moyenne du corps enseignant de mathématiques									
Etablissement avec faible ancienneté moyenne des enseignants							-18.22	(6.581)	0.013
Effets aléatoires									
Niveau 2 (établissements)									
Variance des constantes	465.89		<0.001	209.88		<0.001	179.17		<0.001
Niveau 1 (élèves)									
Variance inter-élèves	6247.92			2918.69			2919.32		
Pouvoir explicatif du modèle									
Pseudo R ² Niveau 1 (élèves)				53.26%			53.27%		
Pseudo R ² Niveau 2 (établissements)				54.93%			61.54%		

Source : SRED et CIRSE

Note : En gras, la valeur de la p-value inférieure à 0.05. Les coefficients associés sont statistiquement significatifs à un seuil de 5%.

Exemple de lecture du tableau : à Genève, les filles natives et francophones (catégorie de référence), de niveau initial moyen et qui étaient scolarisées dans le Regroupement B (ou assimilé) en fin de 10^e, obtiennent en moyenne 54 points en moins au test PISA en mathématiques, par rapport au même groupe de filles qui fréquentaient le Regroupement A (ou assimilé).

Au Tessin, ce sont les élèves initialement de niveaux mixtes ou de base²² qui obtiennent des scores PISA plus faibles que ceux des élèves scolarisés dans les deux niveaux les plus exigeants en allemand et en mathématiques (deux niveaux A). On relèvera qu'il existe même une gradation des effets conforme à l'organisation du système de formation, la situation qui pèjore davantage les résultats obtenus à PISA étant celle d'un élève ayant deux niveaux de base dans les disciplines mentionnées précédemment (le coefficient négatif associé à la modalité *Niveaux BB* est de -73.2 points alors que celui associé à *Niveaux AB* n'est que de -47.1 points) (Figure 15).

Par ailleurs, le statut migratoire et la langue ne possèdent pas d'effets statistiquement significatifs, le lien entre ces variables et les résultats en mathématiques étant probablement plus modeste que pour d'autres domaines comme la lecture. En fait, sur le plan des caractéristiques individuelles, la seule différence notable entre les deux cantons concerne le statut socio-économique et culturel qui conserve, à Genève, un effet statistiquement significatif sur les performances (plus ce statut est élevé et meilleurs sont les résultats obtenus à PISA) alors qu'il n'en est rien au Tessin.

À ce stade de la démarche analytique, d'autres éléments intéressants peuvent être déduits de l'observation du pouvoir explicatif des modèles mis en œuvre.

En effet, toute la complexité du comportement des élèves et des mécanismes d'acquisition de compétences ne peut évidemment pas être résumée dans les quelques caractéristiques individuelles que nous venons de prendre en compte.

Ces caractéristiques, bien que couramment mentionnées dans la littérature pour expliquer des différences de résultats, ne constituent qu'une partie des facteurs explicatifs. Dès lors il est intéressant de quantifier le pouvoir explicatif de chacun de ces modèles pour déterminer quelle est la part du phénomène que nous parvenons à expliquer dans chaque canton. Dans la modélisation multiniveau, le pouvoir explicatif se calcule pour chacun des niveaux qui structurent les données²³. On rappelle également ici que les deux modèles analysés ont été spécifiés en utilisant le même jeu de variables de premier niveau, ce qui les rend directement comparables sur le plan de leur capacité à expliquer les différences de performances en mathématiques.

En outre, comme nous l'avons mentionné précédemment, la décomposition de la variance totale des scores (modèle «vide») s'est révélée très similaire dans les deux cantons, ce qui facilite encore davantage la comparaison puisque, dans les deux cas, environ 93% des différences de résultats sont imputables à des différences de caractéristiques individuelles et seuls 7% de ces mêmes différences sont liés à l'établissement fréquenté. La prise en compte des caractéristiques sociodémographiques (statut socio-économique et culturel, âge, sexe, statut migratoire et langue maternelle) et scolaires (niveau initial de mathématiques) dans l'analyse permet d'expliquer 53% de la variance inter-élèves (1^{er} niveau) à Genève contre 42% au Tessin.

Par ailleurs, ces mêmes caractéristiques expliquent aussi une partie de la variance inter-établissements (2^e niveau) dans les deux cantons. En effet, à Genève, plus de la moitié des différences de performances entre écoles sont en réalité liées aux caractéristiques individuelles du public scolarisé dans chaque établissement (55%). Au Tessin, ce chiffre n'est que de 20% (Figure 15).

Ainsi, il semble que le poids des caractéristiques individuelles dans l'explication des différences de résultats des élèves soit plus faible au Tessin qu'à Genève et ce aussi bien au premier niveau de l'analyse que dans l'explication des différences liées à l'établissement. Comment interpréter le fait que l'intensité du lien entre ces caractéristiques et le score des élèves en mathématiques soit différente d'un canton à l'autre ? Dans le cas présent, faut-il y voir le signe d'une plus grande équité du système éducatif tessinois et, si oui, quelle en serait l'origine ? Felouzis et Charmillot (2017), qui se sont intéressés à l'évolution des inégalités scolaires en Suisse en exploitant les données des enquêtes PISA 2003 et 2012, avancent l'hypothèse que cette plus grande équité serait le fruit d'une organisation différente du système scolaire, les systèmes intégrés comme celui en vigueur au Tessin ayant tendance à être plus équitables que les systèmes segmentés comme celui qui prévaut à Genève.

²² L'élève dont les niveaux sont dits "mixtes" est inscrit dans un cours de niveau de base en allemand (B) et un niveau plus exigeant en mathématiques (A) ou inversement. L'élève dont les niveaux sont de base (deux niveaux B) est quant à lui inscrit à la fois dans des cours de niveaux de base en allemand et en mathématiques. Cette modalité correspond aux élèves dont le niveau scolaire est le moins élevé.

²³ Dans le cas présent on calcule un «Pseudo R²» au niveau des élèves et un au niveau des établissements (voir Bressoux (2010) pour plus de précisions sur la méthodologie de calcul). Comme l'indiquent Bressoux et Pansu (2003), l'intérêt de ces modèles est de séparer les niveaux d'explication du phénomène, les pourcentages de variance expliquée à chacun des niveaux ne pouvant être directement sommés pour rendre compte d'une part de variance expliquée globale du phénomène.

Figure 15 - Résultats des analyses multiniveaux pour le score de mathématiques à PISA 2012 – Canton du Tessin

Résultats des analyses multiniveaux pour le score de mathématique à Pisa 2012 (Canton du Tessin)									
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)			Modèle 3 (variables niveaux 1 et 2)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes									
Constante	514.31	(4.238)	<0.001	534.29	(4.253)	<0.001	533.78	(4.323)	<0.001
Niveau initial fin 10 ^e									
Note en math				34.9	(3.263)	<0.001	35.02	(3.264)	<0.001
Niveaux en math et allemand <i>(réf. Deux niveaux A)</i>									
Niveau AB				-47.09	(6.238)	<0.001	-47.11	(6.241)	<0.001
Niveau BB				-73.21	(4.270)	<0.001	-73.19	(4.274)	<0.001
Sexe									
<i>(réf. Fille)</i>									
Garçon				14.21	(4.234)	0.001	14.22	(4.234)	0.001
Age				-9.37	(4.512)	0.039	-9.46	(4.505)	0.037
Statut socio-économique et culturel (ESCS)				4.68	(2.634)	0.076	4.73	(2.633)	0.073
Statut migratoire									
<i>(réf. Natif)</i>									
2 nd /1 ^{ère} génération				-8.3	(4.762)	0.092	-8.24	(4.768)	0.095
Langue									
<i>(réf. Italophone)</i>									
Allophone				5.6	(5.051)	0.269	5.51	(5.046)	0.276
Taille moyenne des classes									
Etablissements avec des classes de taille inférieur à la norme							19.75	(8.282)	0.023
Effets aléatoires									
Niveau 2 (établissements)									
Variance des constantes	433.68		<0.001	344.95		<0.001	334.55		<0.001
Niveau 1 (élèves)									
Variance inter-élèves	5626.14			3259.33			3259.36		
Pouvoir explicatif du modèle									
Pseudo R ² Niveau 1 (élèves)				42.07%			42.07%		
Pseudo R ² Niveau 2 (établissements)				20.46%			22.86%		

Source : SRED et CIRSE

Note : En gras, la valeur de la p-value inférieure à 0.05. Les coefficients associés sont statistiquement significatifs à un seuil de 5%.

Exemple de lecture du tableau : pour les mathématiques, au Tessin les filles natives et italophones avec deux niveaux A (en mathématiques et en allemand) obtiennent en moyenne 534 points au test PISA. En outre, plus la note en mathématiques en fin de 10^e est élevée et plus le score obtenu à PISA l'est également.

L'introduction des caractéristiques individuelles dans la modélisation a permis d'expliquer une part non négligeable des différences de résultats en mathématiques mais le phénomène n'est toutefois pas totalement déterministe puisque, à ce stade, une part encore importante de la variance reste inexpliquée.

Il subsiste notamment, au 2^e niveau de l'analyse, une variance résiduelle statistiquement significative (variance inter-établissements) que l'on va tenter d'expliquer davantage en introduisant les caractéristiques des établissements dans la modélisation. Dans cette ultime étape de la démarche analytique, il s'agit d'apporter des éléments de réponse sur la nature de l'effet établissement.

Quelles sont les caractéristiques des écoles qui pourraient expliquer que deux élèves semblables en début de période de formation obtiennent, in fine, des résultats en mathématiques qui divergent ? Autrement dit, pourquoi certains établissements sont-ils plus efficaces que d'autres ?

L'hypothèse principale que nous souhaitons tester était celle de l'existence d'un effet de la composition sociale et de la composition académique de l'établissement sur les performances des élèves. Il apparaît qu'une telle hypothèse est infondée puisqu'aucun effet de cette nature ne peut être mis en évidence sur les données de l'enquête PISA 2012 dans les cantons de Genève et du Tessin.

En revanche, d'autres caractéristiques propres à l'établissement semblent jouer un rôle sur les performances des élèves en mathématiques.

À Genève, l'expérience des enseignants possède par exemple un effet statistiquement significatif. Ainsi, à caractéristiques individuelles identiques (notamment le niveau initial), un élève obtient de moins bons résultats à PISA lorsqu'il est scolarisé dans un établissement où les enseignants ont une ancienneté plus faible que celle qui existe en moyenne dans l'ensemble du canton. Ce résultat est assez conforme à ce qu'ont mis en évidence certains auteurs, à savoir que l'expérience des enseignants est une caractéristique assez fortement corrélée aux compétences des élèves (Greenwald, Hedges et Laine, 1996; Hedges et Greenwald, 1996; Gustafsson, 2003). On pourrait également émettre l'hypothèse que ceci s'explique aussi en partie par la formation dispensée aux enseignants, les plus expérimentés n'ayant pas nécessairement été formés de la même manière que leurs collègues plus récemment entrés dans le métier.

Quoi qu'il en soit, il convient toutefois de relever ici le caractère non systématique de ce résultat qui ne se retrouve pas pour le Tessin et ne se vérifie pas non plus dans les autres domaines évalués par PISA, rejoignant ainsi le constat fait par Hanushek (2002, 2003) d'une absence d'effet systématique évident de ce type de variable sur les performances des élèves.

Au Tessin, pour les mathématiques, la seule caractéristique de l'établissement qui possède un effet sur les résultats des élèves est une variable de type organisationnel. En effet, «toutes choses égales par ailleurs», un élève obtient de meilleurs résultats à PISA lorsqu'il est scolarisé dans un établissement qui constitue des classes dont la taille est plutôt inférieure à ce qui se fait dans le reste du canton (entre 16 et 17 élèves en moyenne contre un peu plus de 20 élèves dans l'ensemble du canton). Ce constat pourrait en premier lieu s'interpréter à la lumière des travaux qui se sont intéressés à l'impact de la taille des classes sur la réussite des élèves. Bien que la question reste encore ouverte, il apparaît en effet dans bon nombre d'entre eux que la réduction de la taille des classes se traduit par de meilleures acquisitions des élèves (Piketty et Valdenaire, 2006; Meuret, 2001).

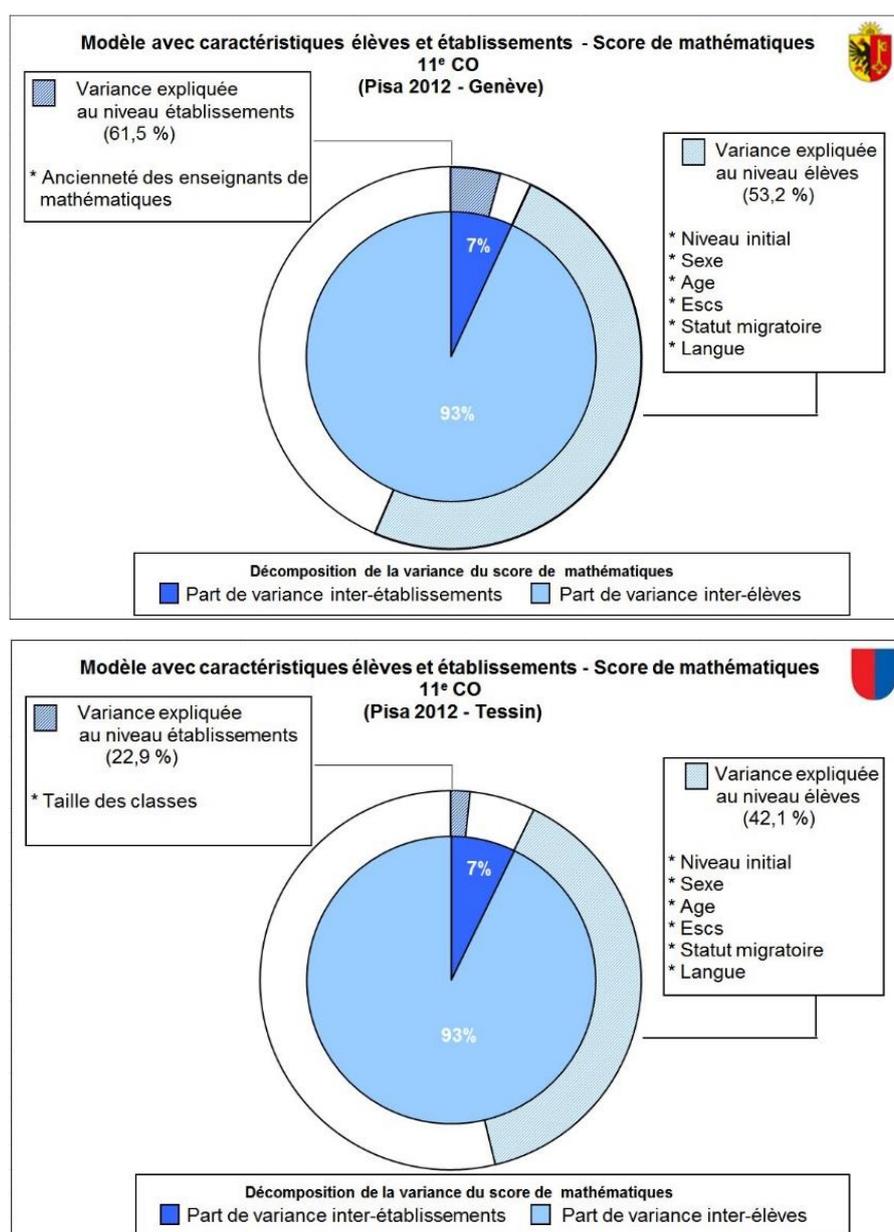
L'interprétation de ce résultat n'est toutefois pas nécessairement aussi aisée qu'il y paraît, dans la mesure où les établissements concernés appartiennent en fait à une même zone géographique. Dès lors on peut supposer que les élèves y bénéficient de conditions d'enseignement un peu particulières, peut-être plus favorables à certains égards que celles en vigueur dans le reste du canton. Ces dernières pourraient d'ailleurs être davantage le fruit d'une situation géographique atypique que le résultat d'un choix de politique éducative différent. On ne peut dès lors exclure le fait que l'impact positif sur les performances en mathématiques d'une scolarisation dans l'une de ces écoles ne soit pas véritablement lié à la taille des classes mais plutôt à une autre caractéristique inobservée de ces mêmes établissements un peu hors norme dans le paysage cantonal.

Enfin, on relèvera qu'à Genève comme au Tessin, les seules variables qui caractérisent l'établissement et pour lesquelles nous sommes parvenus à mettre en évidence un effet statistiquement significatif sur les résultats des élèves, à savoir l'expérience des enseignants de mathématiques et la taille des classes, renvoient plus ou moins directement à un contexte qui correspond davantage à la classe qu'à l'établissement.

Est-ce le signe d'une plus grande pertinence de la classe en tant qu'unité d'analyse des variations d'acquisitions de compétences ? Dans une revue de la littérature consacrée à la distinction entre effet «classe» et effet «établissement», Duru-Bellat (2003) relevait que «les effets du contexte sont vraisemblablement plus forts au niveau de la classe pour ce qui est des progressions scolaires et plus marqués au niveau de l'établissement en ce qui concerne le déroulement des carrières et peut-être aussi la socialisation des élèves» (p. 191).

Il aurait vraisemblablement été intéressant de traiter simultanément dans une même analyse ces deux niveaux hiérarchisés de contexte, ce que permet d'ailleurs le modèle multiniveau que nous avons mis en œuvre, mais cela ne pouvait malheureusement pas s'envisager dans le cas des données PISA puisque les élèves ne sont échantillonnés qu'en fonction de l'établissement fréquenté.

Figure 16 – Pouvoir explicatif des modèles multiniveaux pour les mathématiques selon le canton



Source : SRED et CIRSE

6.3 La lecture

Dans le domaine de la lecture, la réalisation du modèle «vide» nous permet de constater qu'à Genève la part des différences de résultats imputables à l'établissement fréquenté s'élève à environ 5% et la grande majorité de la variance des scores des élèves (soit environ 95%) est à mettre au crédit de différences de caractéristiques individuelles (Figure 12). Au Tessin, ces chiffres sont relativement différents puisque, bien que les caractéristiques individuelles soient toujours prépondérantes dans l'analyse des variations de performances (environ 87%), la variance inter-établissements représente tout de même 13% de la variance totale des scores.

On notera que la lecture est le seul des trois domaines évalués dans lequel on observe une différence importante entre les deux cantons en ce qui concerne l'estimation de la part de variance inter-établissements fournie par les modèles «vides». Comme pour les mathématiques, à ce stade de l'analyse il n'y a pas encore de variables explicatives introduites dans la modélisation pour fournir des pistes d'explication de ces différents pourcentages.

Nous allons donc introduire successivement dans le modèle les variables qui caractérisent l'élève (niveau 1) puis celles qui caractérisent l'établissement (niveau 2) afin de déterminer lesquelles possèdent un effet sur les performances.

Premièrement, on constate qu'au Tessin comme à Genève, le niveau initial de l'élève a un effet positif sur les performances, ce qui nous paraît assez raisonnable. Ainsi, à caractéristiques individuelles identiques, plus la moyenne annuelle de français ou d'italien est élevée en fin de 10^e et plus on réussit au test PISA en lecture en fin de 11^e (Figures 17 et 18). Le profil de l'élève en fin de 10^e, lui aussi utilisé à des fins de contrôle du niveau initial, joue également un rôle significatif sur les performances «toutes choses égales par ailleurs». Ainsi, à Genève un élève issu du regroupement B ou assimilé réussit moins bien que son camarade qui était scolarisé dans le regroupement A. De même, dans le canton du Tessin, ce sont les élèves initialement de niveaux «mixtes» (un niveau A et un niveau B) ou de base (deux niveaux B) en mathématiques et allemand qui obtiennent des scores PISA plus faibles que ceux des élèves scolarisés dans les deux niveaux les plus exigeants (deux niveaux A), les premiers réussissant d'ailleurs un peu mieux que les seconds. En effet, à caractéristiques individuelles et note d'italien de 10^e identiques, un élève tessinois de niveau «mixte» obtient environ 35 points de moins en lecture qu'un élève de profil «AA». Pour un élève de profil «BB» cette différence s'élève à près de 52 points (Figure 18). Ces résultats sont conformes à ce qui s'observe pour les mathématiques et les sciences et se révèlent tout à fait cohérents dans la mesure où le profil de 10^e est une variable qui permet d'affiner le contrôle du niveau initial de l'élève.

Par ailleurs, dans les deux cantons concernés, on constate que le genre possède lui aussi un effet significatif sur les performances, les filles obtenant de meilleurs résultats que les garçons. On notera que ce résultat avait déjà été observé lors de la précédente enquête PISA 2009 dans une analyse de régression effectuée sur les scores de lecture des élèves tessinois et genevois (Consortium PISA.ch, 2011).

On peut relever encore d'autres similitudes entre les deux cantons en ce qui concerne l'effet du statut migratoire et du statut socio-économique et culturel sur les résultats en lecture des élèves.

Ainsi, «toutes choses égales par ailleurs», à Genève comme au Tessin, les natifs réussissent en moyenne mieux que les élèves migrants de première et deuxième génération. En outre, on peut noter que la différence de performances entre natifs et élèves migrants est du même ordre de grandeur dans les deux cantons (un peu plus d'une dizaine de points).

Le statut socio-économique et culturel ne possède quant à lui pas d'effet statistiquement significatif dans les deux cantons concernés, cette absence d'effet s'expliquant principalement par le contrôle du niveau initial des élèves dans la modélisation²⁴.

²⁴ Lorsqu'on ne tient pas compte des notes et du profil de 10^e dans la modélisation, on observe un effet statistiquement significatif et positif de la variable ESCS.

Figure 17 - Résultats des analyses multiniveaux pour le score de lecture à PISA 2012 – Canton de Genève

Résultats des analyses multiniveaux pour le score de lecture à Pisa 2012						
(Canton de Genève)						
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes						
Constante	499.37	(5.334)	<0.001	525.84	(5.480)	<0.001
Niveau initial fin 10 ^e CO						
Note de français				44.23	(4.314)	<0.001
Profil de 10e CO						
(réf. Regroupement A ou assimilé)						
Regroupement B ou assimilé				-48.25	(7.300)	<0.001
Sexe						
(réf. Fille)						
Garçon				-14.88	(3.976)	<0.001
Age				-22.32	(5.028)	<0.001
Statut socio-économique et culturel (ESCS)				8.46	(4.571)	0.065
Statut migratoire						
(réf. Natif)						
2 ^{nde} /1 ^{ère} génération				-13.1	(6.049)	0.032
Langue						
(réf. Francophone)						
Allophone				-2.33	(6.368)	0.716
Effets aléatoires						
Niveau 2 (établissements)						
Variance des constantes	340.98		<0.001	163.2		<0.001
Niveau 1 (élèves)						
Variance inter-élèves	6641			3813.65		
Pouvoir explicatif du modèle						
Pseudo R ² Niveau 1 (élèves)				42.57%		
Pseudo R ² Niveau 2 (établissements)				52.13%		

Source : SRED et CIRSE

Note : En gras, la valeur de la p-value inférieure à 5%. Les coefficients associés sont statistiquement significatifs à un seuil de 0.05.

Exemple de lecture du tableau : en lecture, l'âge a un effet négatif sur le score obtenu à PISA. Ainsi, l'élève de référence (fille, native, francophone et scolarisée dans le regroupement A en fin de 10^e) plus âgée obtient en moyenne des résultats inférieurs à ceux d'une camarade qui possède les mêmes caractéristiques individuelles mais qui est plus jeune.

Au niveau des autres caractéristiques individuelles prises en considération dans l'explication des différences de résultats, on notera que l'âge et la première langue parlée ne possèdent pas les mêmes effets à Genève et au Tessin. En effet, alors qu'au Tessin l'âge ne possède aucun effet significatif sur les résultats des élèves, à Genève les plus âgés obtiennent des résultats plutôt inférieurs aux autres (Figures 17 et 18). Comme pour les sciences et les mathématiques, on peut émettre l'hypothèse que les élèves plus âgés représentent principalement les cas de redoublements et correspondent donc généralement à ceux qui ont eu des difficultés scolaires dans leur parcours antérieur.

En ce qui concerne la première langue parlée à la maison, alors qu'à Genève on ne relève pas de différence significative entre francophones et allophones (vraisemblablement en raison de la prise en compte simultanée du statut migratoire dans l'analyse), au Tessin on constate en revanche que les allophones réussissent mieux que les italophones. L'interprétation de ce dernier résultat n'est pas aisée. Une explication possible réside peut-être dans les politiques d'insertion des élèves allophones mises en place par le canton et dans la façon dont chacun aborde la compréhension de l'écrit à l'école²⁵.

À l'issue de cette première étape de la démarche analytique, l'introduction des caractéristiques individuelles dans la modélisation a permis d'expliquer plus de la moitié des différences de résultats imputables au niveau élève pour le canton de Genève (soit $0.52 \times 0.95 = 49\%$ de la variance totale des scores de lecture) (Figure 17). Au Tessin, le même jeu de variables explique un peu moins de 40% de ces mêmes différences (soit $0.39 \times 0.87 = 34\%$ de la variance totale des scores de lecture).

On retrouve donc ici le constat effectué précédemment pour les mathématiques, à savoir que le poids des caractéristiques individuelles dans l'explication des différences de résultats semble plus important à Genève qu'au Tessin, ce qui pose une fois de plus la question de la moins grande équité du système éducatif genevois et de son origine. La prise en compte des caractéristiques sociodémographiques et scolaires des élèves dans l'analyse a également permis d'expliquer une partie de la variance inter-établissements (respectivement 42.5% et 16.11% à Genève et au Tessin) (Figures 17 et 18) mais la majeure partie de cette dernière demeure encore inexpliquée. Dès lors on peut introduire dans le modèle les caractéristiques des établissements afin de voir dans quelle mesure on parvient à mieux appréhender le phénomène.

À Genève, aucune des caractéristiques de l'établissement dont nous disposons ne possède d'effet statistiquement significatif sur les performances des élèves. Cela signifie que l'explication des 5% de variance des scores imputables à l'établissement passe par l'investigation d'autres pistes que celles de la composition, de l'organisation ou de la gestion des écoles.

²⁵ Pour plus de détails, un document avec les directives à suivre pour l'accueil des élèves ayant ces spécificités a été rédigé par le DECS (http://www4.ti.ch/fileadmin/DECS/DS/UIM/Cosa_facciamo/Diversi/Linee_guida_allievi_alloglotti_scuola_dell_obbligo.pdf). Il est dès lors possible d'y trouver les différentes mesures et instruments mis en place pour l'insertion des élèves allophones.

Figure 18 - Résultats des analyses multiniveaux pour le score de lecture à PISA 2012 – Canton du Tessin

Résultats des analyses multiniveaux pour le score en lecture à Pisa 2012 (Canton du Tessin)									
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)			Modèle 3 (variables niveaux 1 et 2)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes									
Constante	484.47	(5.887)	<0.001	517.15	(6.761)	<0.001	519.36	(6.704)	<0.001
Niveau initial fin 10 ^e									
Note en lecture (italien)				37.72	(5.560)	<0.001	37.71	(5.868)	<0.001
Niveaux en math et allemand <i>(réf. Deux niveaux A)</i>									
Niveau AB				-34.83	(7.454)	<0.001	-34.9	(7.446)	<0.001
Niveau BB				-51.71	(7.865)	<0.001	-51.63	(7.878)	<0.001
Sexe									
<i>(réf. Fille)</i>									
Garçon				-27.62	(5.560)	<0.001	-27.67	(5.558)	<0.001
Age				-5.714	(4.891)	0.245	-5.604	(4.839)	0.249
Statut socio-économique et culturel (ESCS)				5.251	(3.225)	0.106	5.343	(3.236)	0.101
Statut migratoire									
<i>(réf. Natif)</i>									
2 nd e/1 ^{ère} génération				-11.09	(4.687)	0.02	-10.87	(4.719)	0.024
Langue									
<i>(réf. Italophone)</i>									
Allophone				14.325	(5.025)	0.006	14.19	(5.018)	0.006
Taille moyenne des classes									
Etablissements avec des classes de taille inférieur à la norme							22.07	(10.203)	0.038
Taux de non-promotion									
Etablissements avec un taux de non-promotion supérieur à la norme							-59.32	(15.738)	<0.001
Effets aléatoires									
Niveau 2 (établissements)									
Variance des constantes	946.56		<0.001	794.1		<0.001	620.4		<0.001
Niveau 1 (élèves)									
Variance inter-élèves	6464.37			3912.84			3912.58		
Pouvoir explicatif du modèle									
Pseudo R ² Niveau 1 (élèves)				39.47%			39.47%		
Pseudo R ² Niveau 2 (établissements)				16.11%			34.46%		

Source : SRED et CIRSE 2017

Note : En gras, la valeur de la p-value inférieure à 0.05. Les coefficients associés sont statistiquement significatifs à un seuil de 5%. Exemple de lecture du tableau : au Tessin, les garçons réussissent en moyenne moins bien par rapport aux filles, toutes choses égales par ailleurs. Ainsi, en moyenne, ils obtiennent 27.6 points de moins que les filles au test PISA de lecture.

Peut-être ce résultat est-il aussi le fruit du caractère un peu atypique de la lecture par rapport aux mathématiques ou aux sciences puisqu'il s'agit d'un domaine qui est à la fois une compétence transversale et un objet de savoir. Les compétences acquises dans ce domaine seraient-elles davantage liées à ce qui se passe en dehors de l'école (aspects culturels par exemple), rendant de fait plus difficiles les questionnements autour d'un effet établissement et de sa nature ?

Pour le canton du Tessin, par contre, l'introduction de variables de niveau 2 a produit des résultats intéressants. Comme pour les mathématiques, il semblerait que les performances varient en fonction de la politique de constitution des classes, à savoir qu'à caractéristiques individuelles identiques, on obtient de meilleurs résultats lorsqu'on est scolarisé dans un établissement qui crée des classes plus petites que la moyenne cantonale.

Plusieurs auteurs ont souligné l'impact positif d'une réduction de la taille des classes sur les performances des élèves (Piketty et Valdenaire, 2006). Dans ce cadre, les enseignants seraient plus attentifs aux élèves, consacraient davantage de temps à revenir sur les éléments appris, à vérifier les acquis et mettraient en œuvre un enseignement plus individualisé (Betts et Shkolnik, 1999). D'après Meuret (2001), qui s'appuie sur des recherches américaines réalisées dans le cadre du projet STAR, les élèves auraient également un comportement davantage centré sur la tâche et feraient preuve de plus d'engagement.

Cependant, dans le cas du Tessin, comme nous l'avons déjà mentionné précédemment, il faut rester relativement prudent quant à l'interprétation de ce résultat dans la mesure où on ne peut faire abstraction d'un certain contexte. En effet, on ne dénombre que deux établissements qui constituent des classes plus petites que la norme et tous deux peuvent apparaître comme atypiques sur le territoire cantonal. En effet, positionnés dans un contexte périphérique, ces établissements sont aussi les plus petits de l'échantillon et connaissent, de fait, des conditions d'enseignement un peu particulières.

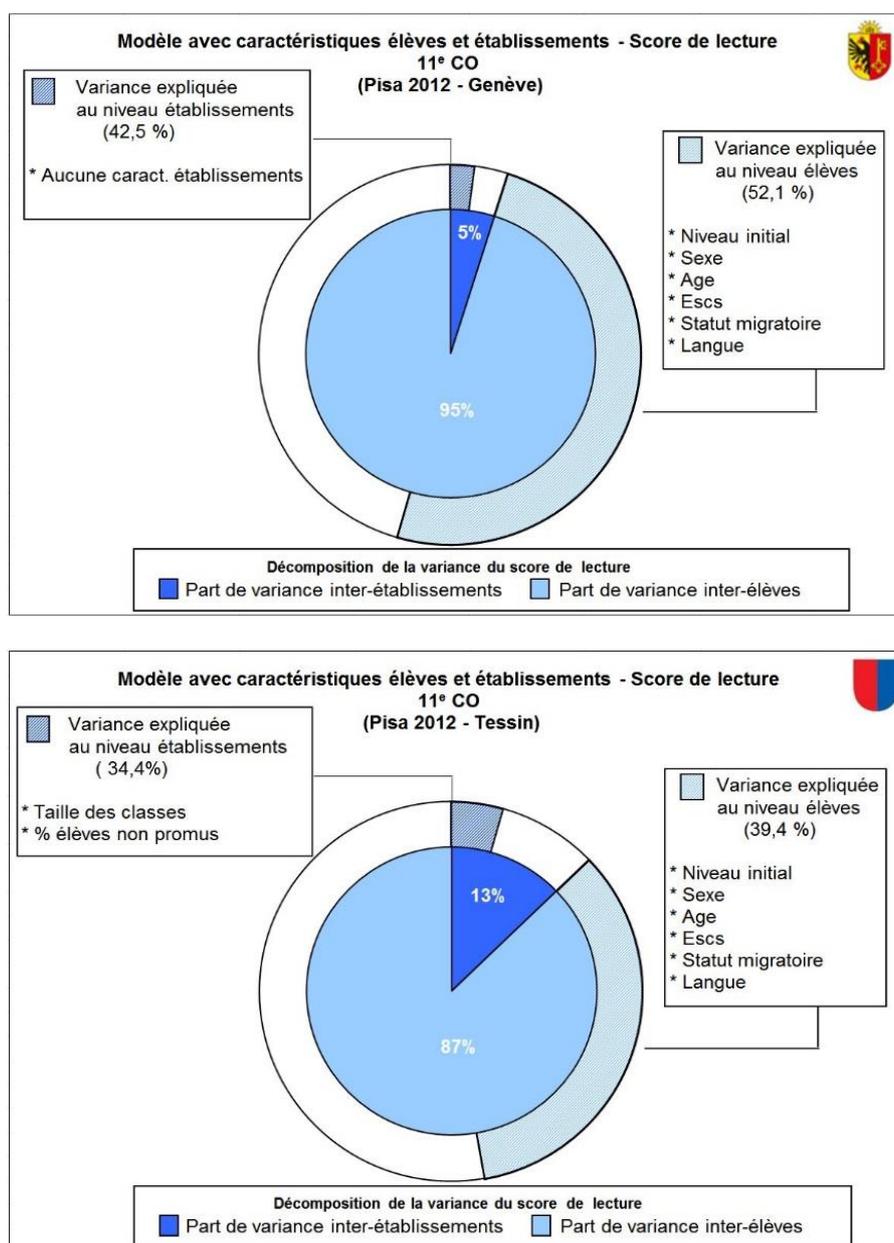
Enfin, on notera encore qu'au Tessin le fait d'être scolarisé dans une école où le pourcentage de redoublements est supérieur à la moyenne (6% contre 2% pour l'ensemble de la Scuola Media) est associé à une réussite significativement plus faible dans le domaine de la lecture.

Ce résultat pourrait en quelque sorte faire écho à ce qui a été démontré à une plus grande échelle, à savoir que les élèves scolarisés dans des pays où le redoublement est nul ou faible réussissent mieux que les élèves appartenant à des pays où cette pratique est plus fréquemment mise en œuvre (Rocher, 2008).

L'interprétation reste cependant délicate. En effet, les établissements ont une large marge de manœuvre dans la mise en œuvre du redoublement et la signification que chaque institution lui accorde peut se révéler assez différente. La note finale obtenue par un élève dans une discipline ne résulte pas uniquement de son niveau de performance aux différentes épreuves auxquelles il a été soumis mais dépend également de nombreux autres paramètres pris en compte par l'enseignant (Allal et Mottier Lopez, 2008 citées par Mottier Lopez et Crahay, 2009). Ainsi, on peut se demander ce qui se cache derrière la décision de non-promotion d'un élève : est-ce que les enseignants sont plus exigeants ? Est-ce que la direction de l'établissement a une conception particulière du redoublement (considéré comme plus ou moins efficace selon les cas) ? Le résultat que nous avons mis en évidence ici ne nous permet malheureusement pas d'aller au-delà du seul constat d'une performance moyenne plus faibles des élèves dans les écoles concernées (seulement deux sur tout l'échantillon). Les mécanismes sous-jacents restent cependant encore à identifier. On notera tout de même, en guise de conclusion, que l'introduction des variables de niveau école dans la modélisation a permis de parvenir à un modèle qui explique davantage la variance²⁶ des performances individuelles.

²⁶ Le Pseudo R² passe de 16% à plus de 34% (Figure 18).

Figure 19 - Pouvoir explicatif des modèles multiniveaux pour la lecture selon le canton



Source : SRED et CIRSE 2017

6.4 Les sciences

Pour les sciences, comme pour les mathématiques et la lecture, la démarche analytique reste la même, à savoir qu'on introduit dans un premier temps dans le modèle multiniveau les variables qui caractérisent l'élève puis, dans un second temps, celles qui caractérisent l'établissement. Le but est, encore une fois, de parvenir à déterminer quels sont les facteurs qui expliquent, au moins en partie, les différences de résultats des élèves au test PISA dans ce domaine.

Dans les deux cantons on retrouve, sans grande surprise, un effet du niveau initial sur l'acquisition des compétences en sciences. Ainsi on observe un effet positif associé aux moyennes annuelles de fin de 10^e et un effet négatif associé au profil des élèves (Figures 20 et 21).

Comme précédemment, les élèves tessinois initialement de niveaux «mixtes» (un niveau A et un niveau B) ou de base (deux niveaux B) en mathématiques et allemand obtiennent, «toutes choses égales par ailleurs», des scores plus faibles en fin de 11^e que ceux de leurs camarades ayant été scolarisés dans les deux niveaux les plus exigeants («AA»).

Pour Genève, ce sont les élèves initialement scolarisés dans le regroupement B en 10^e qui obtiennent des résultats inférieurs à ceux du regroupement A, filière la plus exigeante. Il est important de souligner ici que pour le canton de Genève, la note de fin de 10^e utilisée à des fins de contrôle du niveau initial est la moyenne annuelle de biologie, discipline qui ne concerne qu'une partie des items de sciences proposés dans PISA (l'autre partie porte sur des items qui relèvent davantage de la physique).

Ce choix, imposé par l'organisation de système de formation genevois (en 10^e la physique n'est enseignée qu'à une partie des élèves alors que tous se voient dispenser un enseignement de biologie), rend la comparaison stricte avec le Tessin un peu plus difficile qu'à l'accoutumée dans la mesure où, dans ce dernier canton, la moyenne annuelle de sciences en fin de 10^e correspond davantage à ce qui est évalué dans PISA (à la fois biologie et physique). On a cependant pu vérifier que, sur le plan statistique, les corrélations entre moyennes annuelles de 10^e en sciences ou biologie et scores de sciences à PISA en 11^e étaient du même ordre de grandeur dans chacun des cantons, ce qui nous laisse penser que la comparabilité des résultats reste relativement robuste (Figure 5).

En ce qui concerne les effets des caractéristiques sociodémographiques, on relève tout d'abord un certain nombre de similitudes entre les deux cantons. Ainsi, à Genève comme au Tessin, les natifs obtiennent, «toutes choses égales par ailleurs», de meilleurs résultats que ceux des élèves issus de la migration (1^{re} ou 2^e génération).

De même, plus le statut socio-économique et culturel est élevé et plus le score obtenu en sciences est important. Enfin, on ne relève aucune différence de résultats statistiquement significative entre les allophones et ceux parlant la langue du test, ceci s'expliquant vraisemblablement par le contrôle simultané du statut migratoire dans la modélisation.

D'autres variables ont, en revanche, un effet différencié dans les deux cantons. On notera qu'à Genève les élèves plus âgés, qui sont vraisemblablement ceux ayant rencontré des difficultés dans leur parcours scolaire antérieur, réussissent moins bien que leurs camarades plus jeunes alors qu'il n'en est rien au Tessin (Figures 20 et 21). Par ailleurs alors qu'à Genève le fait d'être un garçon constitue un avantage en matière de performances en sciences, au Tessin garçons et filles réussissent de la même manière.

À l'issue de cette première phase de la démarche analytique, l'observation du pouvoir explicatif des différents modèles nous permet de constater une nouvelle fois que le poids des caractéristiques individuelles dans l'explication des différences de résultats semble un peu plus important à Genève. Ainsi, au Tessin, la part de variance des scores expliquée par ces variables de premier niveau est de 40.6 % (soit $0.406 \times 0.94 = 38\%$ de la variance totale) (Figure 21). À Genève, ce chiffre s'élève à 44.8 % (soit $0.448 \times 0.93 = 41.7\%$ de la variance totale).

Figure 20 - Résultats des analyses multiniveaux pour le score de sciences à PISA 2012 – Canton de Genève

Résultats des analyses multiniveaux pour le score de sciences à PISA 2012 (Canton de Genève)									
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)			Modèle 3 (variables niveaux 1 et 2)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes									
Constante	487.25	(5.736)	<0.001	496.36	(5.539)	<0.001	516.88	(7.298)	<0.001
Niveau initial fin 10 ^e CO									
Note de biologie				44.31	(4.182)	<0.001	44.7	(4.250)	<0.001
Profil de 10 ^e CO <i>(réf. Regroupement A ou assimilé)</i>									
Regroupement B ou assimilé				-45.97	(7.415)	<0.001	-45.75	(7.294)	<0.001
Sexe									
<i>(réf. Fille)</i>									
Garçon				15.51	(3.818)	<0.001	15.83	(3.756)	<0.001
Age				-20.48	(5.844)	0.003	-20.44	(5.817)	0.003
Statut socio-économique et culturel (ESCS)				11.85	(4.373)	0.007	12.34	(4.338)	0.005
Statut migratoire									
<i>(réf. Natif)</i>									
2 nd e/1 ^{ère} génération				-11.71	(4.896)	0.018	-12.71	(4.707)	0.008
Langue									
<i>(réf. Francophone)</i>									
Allophone				-2.27	(5.564)	0.683	-3.36	(5.594)	0.549
Taille de l'établissement scolaire									
Etablissement de taille plutôt supérieure à la norme							-30.03	(7.510)	0.001
Indice de participation des enseignants à la gestion de l'établissement							9.48	(2.672)	0.003
Effets aléatoires									
Niveau 2 (établissements)									
Variance des constantes	430.77		<0.001	274.79		<0.001	80.91		<0.001
Niveau 1 (élèves)									
Variance inter-élèves	5924.7			3268.21			3266.85		
Pouvoir explicatif du modèle									
Pseudo R ² Niveau 1 (élèves)				44.83%			44.86%		
Pseudo R ² Niveau 2 (établissements)				36.20%			81.21%		

Source : SRED et CIRSE

Note : En gras, la valeur de la p-value inférieure à 5%. Les coefficients associés sont statistiquement significatifs à un seuil de 0.05.

Exemple de lecture du tableau : à Genève, le statu socio-économique a un effet positif sur les performances en sciences évaluées dans PISA. Plus précisément, les élèves provenant de milieux plus favorisés obtiennent en moyenne des résultats meilleurs, toutes choses égales par ailleurs.

On ne peut exclure ici l'hypothèse que cet écart soit en partie lié à la différence de variable utilisée pour contrôler le niveau initial des élèves (note de biologie à Genève vs note de sciences au Tessin), mais ce résultat, analogue à ceux observés pour les mathématiques et la lecture, repose tout de même la question de la plus grande équité du système éducatif tessinois et de ses causes potentielles (organisation du système éducatif, pratiques d'apprentissage, politiques publiques en matière d'éducation).

Par ailleurs, on constate qu'au 2^e niveau de l'analyse, l'introduction des caractéristiques individuelles dans la modélisation a permis d'expliquer une partie de la variance inter-établissements des scores genevois (36.2 %), attestant du fait qu'une partie des différences de résultats entre établissements du Cycle d'orientation s'explique en réalité par les différences de structure de la population accueillie au sein de ces entités (Figure 20).

Au Tessin, les résultats sont assez différents puisqu'il semble y avoir, pour le domaine des sciences uniquement, une déconnexion complète entre caractéristiques sociodémographiques et scolaires des élèves et résultats des établissements de la Scuola Media. Autrement dit, dans ce canton l'introduction des caractéristiques individuelles dans l'analyse multiniveau ne permet pas d'expliquer, même partiellement, la variance inter-établissements²⁷ (Figure 21).

Il s'agit là d'un phénomène documenté dans la littérature sur l'analyse multiniveau. Bressoux (2010) explique que «ce phénomène est le plus susceptible de se produire quand la plus grande partie, voire la quasi-totalité de la variance résiduelle se répartit sur l'un des niveaux impliqués dans l'analyse (Singer et Willett, 2003)» (p. 314).

On rappelle que la part de variance inter-établissements dans le domaine des sciences n'est au Tessin que de 5% (il s'agit là du chiffre le plus faible pour les trois domaines évalués dans les deux cantons concernés). Ce résultat peut s'interpréter comme le fait que, pour la Scuola Media, les différences de résultats imputables à l'école ne sont pas liées aux différences de caractéristiques des élèves accueillis, comme c'est bien souvent le cas, mais davantage à l'établissement lui-même. Par comparaison, ce résultat nous interroge également sur le statut des sciences dans le canton de Genève, où tous les élèves ne reçoivent pas le même enseignement en fonction de la filière qu'ils fréquentent²⁸ alors qu'au Tessin tous se voient dispenser les mêmes savoirs.

À ce stade de l'analyse, toute la variance inter-établissements des scores de sciences au Tessin, tout comme une partie de cette même variance pour le canton de Genève, reste donc à expliquer. L'introduction de caractéristiques des écoles dans la modélisation devrait nous permettre de progresser dans ce domaine.

Comme pour les mathématiques et la lecture, l'hypothèse d'un effet de composition sociale et académique de l'établissement sur les performances des élèves n'est pas confirmée. En revanche, d'autres caractéristiques des établissements ont bien un effet sur les performances des élèves dans ce domaine.

Bien que les variables de deuxième niveau qui possèdent un effet statistiquement significatif ne soient pas les mêmes dans les deux cantons, on peut tout de même les regrouper en deux dimensions, l'une relative à l'infrastructure (taille de l'établissement pour Genève et taille des classes pour le Tessin) et l'autre relative, au sens large, au corps enseignant (participation à la gestion de l'établissement pour Genève et niveau de qualification pour le Tessin).

En ce qui concerne la dimension relative au corps enseignant, on relève à Genève un effet du degré de participation des enseignants à la gestion de l'établissement. Ainsi, là où les enseignants participent davantage à la gestion, les élèves réussissent mieux *ceteris paribus* (Figure 21).

Il s'agit là d'un résultat qui rappelle ce que Goddard, Goddard et Tshane-Moran (2007) ont démontré, à savoir que dans les établissements où les enseignants ont la perception d'avoir plus d'impact sur les décisions de gestion, les résultats en lecture et mathématiques des élèves sont plus élevés, bien que l'effet mis en évidence soit d'une ampleur modeste.

²⁷ Sur le plan statistique, ceci se traduit par un Pseudo R² de niveau 2 négatif qui n'a pas réellement de sens.

²⁸ En 10^e seuls les élèves de la section LS, la plus exigeante, avec un profil S (Sciences) suivent un cours de physique.

Figure 21 - Résultats des analyses multiniveaux pour le score de sciences à PISA 2012 – Canton du Tessin

Résultats des analyses multiniveaux pour le score de sciences à Pisa 2012 (Canton du Tessin)									
	Modèle 1 (vide)			Modèle 2 (variables niveau 1)			Modèle 3 (variables niveaux 1 et 2)		
	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value	Valeur	(s.e.)	P-value
Effets fixes									
Constante	490.05	(3.800)	<0.001	502.42	(5.464)	<0.001	495.30	(5.350)	<0.001
Niveau initial fin 10 ^e									
Note de science				41.55	(3.89)	<0.001	41.63	(3.946)	<0.001
Niveaux en math et allemand <i>(réf. Deux niveaux A)</i>									
Niveau AB				-28.797	(6.950)	<0.001	-29.28	(6.975)	<0.001
Niveau BB				-41.71	(6.401)	<0.001	-41.69	(6.404)	<0.001
Sexe									
<i>(réf. Fille)</i>									
Garçon				7.698	(4.425)	0.091	7.80	(4.411)	0.086
Age				-6.809	(4.390)	0.121	-6.91	(4.300)	0.108
Statut socio-économique et culturel (ESCS)				9.554	(3.023)	0.003	9.595	(3.033)	0.003
Statut migratoire									
<i>(réf. Natif)</i>									
2 ^{nde} /1 ^{ère} génération				-13.296	(4.040)	0.001	-12.88	(3.959)	0.001
Langue									
<i>(réf. Italophone)</i>									
Allophone				5.24	(5.525)	0.347	5.054	(5.498)	0.362
Taille moyenne des classes									
Etablissements avec des classes de taille inférieur à la norme							35.042	(11.872)	0.006
Proportion d'enseignants sans habilitation à enseigner									
Etablissements avec une proportion d'enseignants sans habilitation inférieure à la norme							21.88	(6.390)	0.002
Effets aléatoires									
Niveau 2 (établissements)									
Variance des constantes	317.97		<0.001	397.84		<0.001	257.77		<0.001
Niveau 1 (élèves)									
Variance inter-élèves	5924.36			3519.39			3519.07		
Pouvoir explicatif du modèle									
Pseudo R ² Niveau 1 (élèves)				40.59%			40.60%		
Pseudo R ² Niveau 2 (établissements)				-25.12%			18.93%		

Source : SRED et CIRSE

Note : En gras, la valeur de la p-value inférieure à 0.05. Les coefficients associés sont statistiquement significatifs à un seuil de 5%.

Exemple de lecture du tableau : dans le canton du Tessin, les élèves issus de la migration (de première ou deuxième génération) réussissent en moyenne moins bien au test PISA. Ainsi, une fille avec deux niveaux A, italophone (catégorie de référence) mais issue de la migration obtient en moyenne 13 points de moins en sciences qu'une autre fille native ayant les mêmes caractéristiques.

Il convient toutefois de relever ici qu'il existe un biais potentiel dans l'interprétation de ce résultat car l'indice construit dans PISA repose sur les réponses fournies par les chefs d'établissements et non pas par les enseignants eux-mêmes. En outre, les questions posées portent davantage sur les activités mises en œuvre par le chef d'établissement pour encourager la participation des enseignants à la gestion de l'école mais ne permettent pas forcément d'appréhender la participation réelle de ces derniers.

On ne retrouve pas le même résultat au Tessin, mais une autre variable liée au corps enseignant présente un effet statistiquement significatif. En effet, il apparaît qu'à caractéristiques individuelles identiques, les élèves tessinois obtiennent de meilleurs résultats lorsqu'ils sont scolarisés dans des établissements où le corps enseignant est plus qualifié, c'est-à-dire plus exactement dans des établissements où la proportion d'enseignants ne disposant pas de l'habilitation à enseigner est plus faible que ce qui s'observe dans l'ensemble du canton²⁹ (Figure 21).

C'est un résultat qui est conforme à ce que l'on trouve dans la littérature puisque Woessmann (2011) ou Meunier (2011) ont déjà montré que les élèves des écoles ayant un grand nombre d'enseignants sans habilitation à enseigner réussissent en moyenne moins bien.

En ce qui concerne les variables qui portent davantage sur l'infrastructure, à Genève on peut mettre en évidence un effet de la taille de l'établissement sur les performances en sciences. Ainsi les élèves des établissements de taille supérieure à la norme réussissent, *ceteris paribus*, moins bien que les autres (Figure 20).

L'interprétation de ce résultat s'avère relativement délicate. Dans un premier temps, on pourrait penser qu'il s'agit là d'une illustration de ce qu'ont montré certains auteurs, à savoir que les structures de plus petite taille auraient de meilleurs résultats et inversement (Spielhofer, Benton et Schagen, 2007). Toutefois il existe encore, à l'heure actuelle, une controverse scientifique autour de ces questions et le débat ne peut être considéré comme clos.

Par ailleurs, il convient de relever ici que seuls deux établissements du Cycle d'orientation sont dans cette catégorie et qu'il pourrait s'agir de cas particuliers (ils scolarisent respectivement 743 et 751 élèves alors que la moyenne cantonale est de 650 élèves). L'analyse détaillée du questionnaire établissement de PISA nous a en effet permis de constater qu'ils peuvent apparaître atypiques à certains égards. Ainsi, un des deux établissements concernés est le seul de tout l'échantillon genevois à déclarer avoir une forte pénurie d'enseignants qualifiés en sciences et en mathématiques.

Cet élément pourrait en partie expliquer de moins bons résultats en sciences et confirmerait en outre également le résultat trouvé pour le Tessin concernant la qualification des enseignants. Le second établissement concerné relate quant à lui une pénurie ou une inadéquation importante en matière d'ordinateurs pour le travail en classe (il ne s'agit toutefois pas du seul établissement du canton dans cette situation puisque quatre autres directeurs ont déclaré la même chose).

Dès lors, il nous semble important de souligner ici que le résultat qui précède doit être interprété avec une certaine prudence et qu'il ne peut, en aucun cas, encourager à la constitution d'établissements de taille plus modeste.

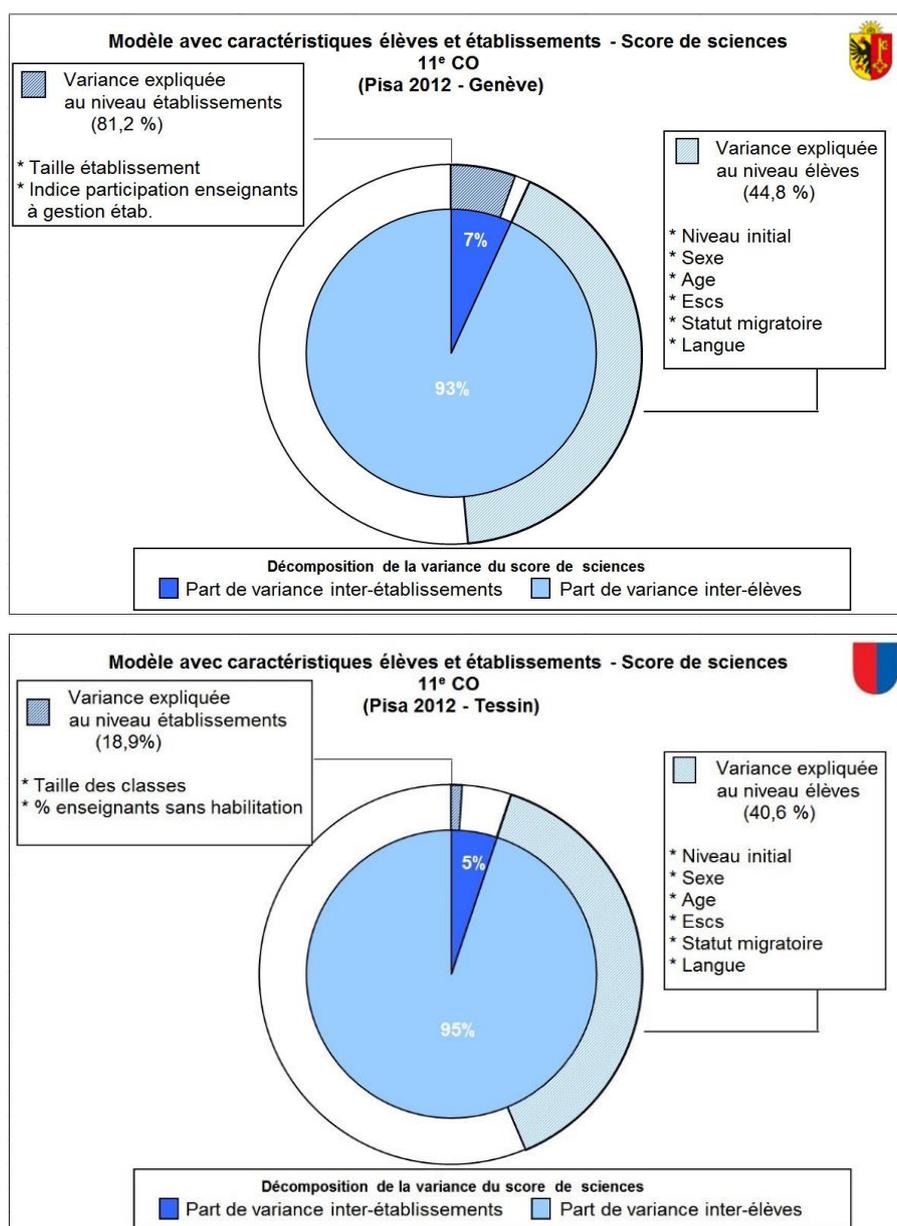
Cette même prudence est également de mise en ce qui concerne le Tessin où, une fois de plus, on retrouve l'effet positif associé au fait d'être scolarisé dans des établissements qui constituent des classes plus petites que la moyenne cantonale.

Comme nous l'avons mentionné précédemment, l'interprétation de ce résultat est elle aussi ambiguë dans la mesure où on peut y voir aussi bien l'impact d'une réduction de la taille des classes sur les performances des élèves que le reflet d'une scolarisation dans des établissements de la Scuola Media qui bénéficient vraisemblablement de conditions d'enseignement un peu particulières en raison de leur localisation géographique.

Pour conclure, on relèvera que l'introduction des caractéristiques des établissements dans l'analyse donne, en matière de pouvoir explicatif, des résultats sensiblement différents d'un canton à l'autre puisqu'elle permet de parvenir à des parts de variance inter-établissements expliquées de 81% pour Genève et de 19% pour le Tessin (Figure 22).

²⁹ Il s'agit d'enseignants qui ont seulement obtenu un diplôme universitaire de niveau bachelor ou master (niveau 5A de la CITE) mais ne disposent pas, en revanche, de l'habilitation à enseigner.

Figure 22 - Pouvoir explicatif des modèles multiniveaux pour les sciences selon le canton



Source : SRED et CIRSE

7. Eléments de synthèse et discussion

La question de recherche à laquelle le CIRSE et le SRED ont tenté d'apporter des éléments de réponse en mettant en œuvre la modélisation multiniveau est celle, encore peu traitée à ce jour en Suisse, des relations qui existent entre acquisitions des élèves et contexte de scolarisation, notamment au niveau de l'établissement.

De façon plus précise, on pourrait résumer la problématique abordée dans la présente recherche aux deux grandes questions suivantes :

- Existe-t-il, au Tessin et à Genève, un effet de l'établissement fréquenté sur l'acquisition des compétences des élèves en mathématiques, lecture et sciences et, le cas échéant, quelle est la taille de cet effet ?
- Si un tel effet existe, de quelle nature est-il ? De façon plus précise, s'agit-il d'un effet de composition, notamment sociale et académique au sens où l'ont défini Dumay et Dupriez (2009), à savoir comme l'impact des caractéristiques agrégées des élèves après que leur effet sur le plan individuel ait été pris en compte, ou s'agit-il plutôt d'un effet d'une autre nature ?

En préambule, relevons que la richesse des données PISA couplée à celle des informations contenues dans les bases de données scolaires cantonales a offert un cadre extrêmement propice à la réalisation de ces travaux et à la mise en œuvre d'un design de recherche particulièrement bien adapté au traitement de la mesure des effets de composition.

En effet, les échantillons cantonaux de 11^e année constitués dans le cadre de l'enquête PISA nous ont fourni des mesures de compétences robustes et comparables d'un canton à l'autre en mathématiques, lecture et sciences à la fin de la scolarité obligatoire.

Ils nous ont également permis de disposer des principales caractéristiques sociodémographiques des élèves et de certains éléments sur leur contexte de scolarisation, les directeurs de tous les établissements de la Scuola Media et du Cycle d'orientation ayant été interrogés sur le climat et la gestion de leur établissement.

Ces informations sur le rôle du directeur d'établissement se sont d'ailleurs révélées extrêmement précieuses dans nos travaux car comme l'a indiqué Bressoux (1994), «les chercheurs qui se sont intéressés au rôle que joue le directeur dans l'école tombent pratiquement tous d'accord sur son importance dans la construction du système social qu'est l'école, dans sa capacité à le faire fonctionner comme un tout cohérent et en particulier dans son rôle pour la constitution d'un climat favorable à la réussite des élèves» (p. 118).

Les bases de données scolaires cantonales ont quant à elles permis de fournir une mesure du niveau initial des élèves, élément indispensable à l'identification d'un véritable effet établissement. En l'absence d'une telle information, il est de fait impossible de conclure, au sens strict, que c'est l'exposition à un contexte différent (en l'occurrence un établissement dont la tonalité sociale ou académique varierait) qui génère des différences de résultats chez des élèves initialement semblables.

Par ailleurs, l'utilisation d'autres informations extraites de ces mêmes bases de données a également permis de parvenir à une description plus judicieuse des établissements en s'appuyant sur les caractéristiques de toute la population d'élèves scolarisée et non des seuls jeunes échantillonnés dans PISA.

En premier lieu, il apparaît qu'on peut mettre en évidence, au Tessin comme à Genève, un effet statistiquement significatif de l'établissement fréquenté sur les performances des élèves dans les trois domaines évalués par PISA.

Toutefois la taille de cet effet est relativement modeste puisque, selon le canton et le domaine considérés, on estime qu'entre 5% et 13% des différences de résultats sont liées à l'établissement.

On notera qu'Orioni (2007) avait déjà mis en évidence, avec les données PISA 2003, une variance inter-établissements de 7% pour les mathématiques, au Tessin et à Genève, chiffres identiques à ceux mesurés dans la présente étude.

La variabilité des scores obtenus en mathématiques, lecture et sciences est donc davantage le fruit de différences individuelles que de différences liées aux établissements.

On retrouve d'ailleurs, sans grande surprise, les effets usuels des caractéristiques sociodémographiques des élèves maintes fois décrits dans la littérature et sur lesquels nous ne reviendrons pas ici (Salvisberg et Zampieri, 2015; Nidegger (éd.), 2014; Consortium PISA.ch, 2011).

À ce sujet, on relèvera simplement que le poids de ces caractéristiques dans l'explication des différences de résultats semble plus faible au Tessin qu'à Genève. Se pose alors la question d'une plus grande équité du système éducatif tessinois et de son origine qui, si l'on s'en remet aux hypothèses formulées par certains auteurs (Felouzis et Charmillot, 2017), serait liée à une organisation intégrée tendancielle plus équitable que celle articulée autour de filières prévalant à Genève.

Une autre façon d'appréhender la relative faiblesse de l'effet établissement dans ces deux cantons consiste à dire que les différences entre élèves au sein d'un même établissement y sont beaucoup plus marquées que les différences entre établissements. Il s'agit là d'un constat qui n'est pas propre à ces deux entités géographiques puisque Bressoux (1994) indique que dans les pays industrialisés, l'environnement scolaire s'avère beaucoup moins explicatif pour la réussite des élèves que leur environnement non scolaire et qu'en règle générale on y observe beaucoup plus de variabilité au sein des écoles qu'il n'y en a d'une école à l'autre.

L'existence d'un effet modeste de l'école fréquentée sur les performances des élèves tessinois et genevois apparaît donc comme un résultat tout à fait conforme à ce que l'on retrouve dans la littérature et illustre parfaitement ce que Duru-Bellat et al. (2004) ont relevé, à savoir que dans bon nombre de travaux la part de la variance des résultats des élèves imputable à des différences entre écoles oscille entre 8% et 15%, la plupart des estimations se situant plutôt dans le bas de cette fourchette.

À quoi faut-il attribuer cette relative faiblesse ? Crahay et Monseur (2008) avancent l'hypothèse que l'effet établissement est, en réalité, tributaire des caractéristiques des systèmes éducatifs. Certains d'entre eux seraient plus enclins que d'autres à faire émerger de tels effets.

Pour illustrer le propos, on peut reprendre l'exemple de Bressoux (1994) qui a tenté de fournir une interprétation à la faiblesse de l'effet établissement dans le contexte français : « dans les pays anglo-saxons, les écoles construisent au moins pour partie leur programme d'enseignement, peuvent aussi définir leurs propres objectifs et peuvent elles-mêmes, sous l'autorité du directeur, recruter leurs enseignants, ou s'en séparer. Tel n'est pas le cas en France, où les programmes sont nationaux, où la décision d'affectation des enseignants n'appartient pas à l'école, ce qui peut donner à penser a priori que les différences de fonctionnement d'une école à l'autre sont faibles » (p. 92).

Par analogie avec ce qui précède, on peut supposer que la faiblesse de l'effet établissement au Tessin et à Genève ainsi que la très grande similitude des résultats obtenus dans les deux cantons trouvent leur origine dans un fonctionnement relativement homogène des établissements de la Scuola Media ou du Cycle d'orientation.

En effet, dans les deux cantons étudiés, bien que les établissements disposent d'un certain degré d'autonomie, bon nombre des principaux facteurs qui régissent leur fonctionnement (plans d'études, niveau de formation des enseignants, taille des classes, ressources, etc.) sont en réalité fixés par voie réglementaire et s'appliquent de façon identique sur l'ensemble du territoire cantonal. Dès lors la variabilité entre établissements s'amenuise et les possibilités de voir émerger un effet contextuel relativement important s'en trouvent réduites.

Par ailleurs, rappelons enfin que dans la présente étude la taille des effets établissement, bien que modeste, est vraisemblablement un peu surestimée, sans que l'on soit toutefois en mesure de savoir dans quelle proportion. Ceci s'explique principalement par la structure des données PISA qui, en raison du plan d'échantillonnage retenu au niveau international, ne comportent que deux niveaux, à savoir des élèves scolarisés dans des établissements.

Les travaux méthodologiques d'Opdenakker et Van Damme (2000) ont montré que dans pareil cas, la prise en compte d'un troisième niveau intermédiaire correspondant à la classe était susceptible d'influencer de façon significative les résultats et l'estimation des paramètres.

Nous avons d'ailleurs pu le vérifier dans le contexte genevois en analysant les scores obtenus par les élèves du regroupement A aux épreuves cantonales de références (Évacom) de mathématiques en fin de 11^e pour l'année 2012, soit des épreuves passées à peu près au même moment que PISA et par les mêmes élèves. La réalisation de modèles «vides» à deux et trois niveaux nous a en effet permis de constater que la taille de l'effet établissement peut presque varier du simple au double selon que

l'on tient compte ou non du niveau classe (de 2% à 3,5% de la variance des scores avec et sans la classe).

Ceci ne remet pas en cause l'effet établissement de taille relativement modeste que nous avons mis en évidence au Tessin et à Genève, mais doit toutefois nous conduire à considérer avec une certaine précaution les estimations fournies dans ce qui précède.

En ce qui concerne la nature de l'effet établissement, l'hypothèse principale que nous souhaitions tester était celle d'un effet de composition sociale et académique conforme aux théories développées par le courant de recherche dit du *school mix* qui a émergé dans les années 80. D'après ces théories, ce sont principalement le niveau scolaire des élèves (*academic mix*) et leur profil sociologique (*social mix*) qui affectent les progressions académiques en modifiant à la fois la quantité et la qualité de l'instruction délivrée et les attitudes et comportements des élèves.

Certains auteurs sont d'ailleurs parvenus à des conclusions de ce type en se basant sur des analyses de données PISA antérieures aux nôtres. Citons par exemple Monseur et Crahay (2006) qui, suite à une comparaison internationale réalisée sur les données de l'enquête 2006, ont indiqué que l'effet école, qui modifie les performances des élèves en lecture et en sciences dans la plupart des pays, était avant tout un effet d'agrégation sociale et académique.

Dans la présente étude, il s'avère en revanche que l'effet établissement identifié au Tessin et à Genève n'est pas de cette nature.

Quel que soit le domaine de compétences considéré, il n'a jamais été possible de montrer que la composition sociale ou académique d'un établissement de la Scuola Media ou du Cycle d'orientation expliquait une partie des différences de performances des élèves. Autrement dit, des élèves semblables du point de vue de leurs caractéristiques sociodémographiques (notamment leur statut socio-économique et culturel) et de leur niveau scolaire initial n'obtiennent pas des résultats qui diffèrent lorsqu'ils sont scolarisés dans des établissements dont la tonalité sociale ou académique varie.

Ce résultat peut, en un certain sens, être interprété comme un signe relativement positif du fonctionnement des systèmes éducatifs tessinois et genevois. En effet, ces derniers ne semblent pas renforcer la production d'inégalités sociales de réussite scolaire en mettant en œuvre des modes de regroupement des élèves au sein des établissements qui génèrent une ségrégation des publics telle qu'elle impacte les performances individuelles.

On peut avancer au moins deux explications à ce résultat, qui diffère quelque peu de ce que d'autres ont trouvé en utilisant des données et des méthodologies pourtant très similaires.

La première de ces explications réside très certainement dans la mise en œuvre d'un design de recherche particulièrement bien adapté à la mesure des effets de composition. Rappelons que l'une des recommandations émises par Thrupp et al. (2002) et reprise par Dumay et al. (2009) stipule que le modèle d'estimation doit incorporer un ensemble complet de prédicteurs individuels et au minimum une mesure de l'état des connaissances antérieures des élèves.

Ce contrôle du niveau initial des élèves, rendu possible par la mobilisation des informations contenues dans les bases de données scolaires cantonales, n'a pu être effectué dans les travaux qui reposent sur les seules données PISA. Or certains auteurs (Willms, 1992; Lauder et Hughes, 1990) ont émis l'hypothèse que les effets de composition mis en évidence dans de telles recherches pouvaient, en réalité, être dus à un mauvais contrôle de variables caractérisant les élèves, et tout particulièrement à une prise en compte insuffisante de leur niveau initial.

Nous avons d'ailleurs pu vérifier que dans le cas genevois, la non-prise en compte des variables permettant de contrôler le niveau initial (profil et moyenne annuelle en fin de de 10^e) dans l'analyse multiniveau des performances en mathématiques nous aurait conduit à conclure, de façon erronée, à l'existence d'un effet de composition sociale de l'établissement.

En l'occurrence, on aurait pu affirmer que «toutes choses égales par ailleurs», un élève obtient de moins bons résultats en mathématiques lorsqu'il est scolarisé dans un établissement de tonalité sociale plutôt défavorisée, ce qui n'est en réalité pas le cas lorsque l'on intègre dans la modélisation une mesure de l'état des connaissances antérieures des élèves.

Par ailleurs, nous avons également pu constater qu'à l'échelle relativement réduite des cantons, la caractérisation de la composition sociale ou académique des établissements, qui demeure toujours une opération délicate, aboutit parfois à des résultats assez différents selon que l'on utilise les

données relatives aux seuls élèves de l'échantillon PISA ou les données exhaustives extraites des bases de données scolaires cantonales. Cet élément est, lui aussi, susceptible d'expliquer en partie l'absence d'effet de composition sociale ou académique des établissements.

La production de connaissances sur les effets de contexte passe bien par l'élaboration de designs de recherche adaptés et la mise en œuvre de méthodologies dédiées, autant de contraintes auxquelles le CIRSE et le SRED ont essayé de répondre dans la présente recherche.

La seconde explication qui permet d'interpréter l'absence d'effet de composition sociale et académique des établissements au Tessin et à Genève n'est pas liée à des facteurs méthodologiques. Dumay, Dupriez et Maroy (2010) ont indiqué que «la présence et la taille d'effets de composition au sein de différents systèmes éducatifs sont conditionnées par les modes de régulation institutionnels des systèmes scolaires et leurs conséquences sur l'environnement des établissements» (p. 479). Contrairement à la situation de quasi-marché analysée par ces auteurs, il semble que les systèmes scolaires tessinois et genevois ne favorisent pas l'émergence d'un effet de composition pour différentes raisons.

Tout d'abord, ils ne permettent pas de réel développement de la ségrégation entre écoles dans la mesure où ces dernières n'ont pas la possibilité de sélectionner leurs élèves (seul le lieu de résidence détermine le lieu de scolarisation). Par ailleurs, comme l'ont relevé Dumay et al. (2010), «la ségrégation entre écoles risque davantage de produire des effets sur les élèves si les écoles bénéficient d'une forte autonomie pédagogique, susceptible d'entraîner une forme d'adaptation des écoles à leur public et une différenciation des objectifs en fonction des compétences (présumées) des élèves fréquentant un établissement» (p. 479).

Or au Tessin et à Genève, en ce qui concerne les écoles publiques, il n'existe rien de tel dans la mesure où ces cantons disposent d'un cadre relativement contraignant en matière de plans d'études et de moyens d'enseignement. Ce dernier contribue à rendre les choses assez homogènes dans l'ensemble des établissements, même s'il existe une certaine liberté académique des enseignants.

Dès lors, il apparaît que les systèmes éducatifs des deux cantons étudiés, bien que sensiblement différents dans leur organisation, sont en réalité des contextes peu propices à l'émergence d'effets de composition.

Puisqu'au Tessin et à Genève l'effet établissement n'est pas un effet de composition sociale ou académique, la question de sa nature reste ouverte.

Dans ces deux cantons, quelles sont les autres caractéristiques des écoles qui pourraient expliquer que deux élèves semblables en début de période de formation obtiennent, in fine, des résultats en mathématiques, lecture et sciences qui divergent ? Pourquoi certains établissements y sont-ils plus efficaces que d'autres ?

Nous avons pu identifier différents facteurs qui expliquent, en partie, les différences de résultats imputables à l'établissement, mais on relèvera d'emblée qu'ils diffèrent à la fois d'un domaine de compétences à l'autre et d'un canton à l'autre, ce qui nous empêche de dégager des conclusions univoques et de généraliser.

Ainsi, au Tessin, on a pu mettre en évidence qu'un élève obtient, «toutes choses égales par ailleurs», de meilleurs résultats dans les trois domaines de compétences lorsqu'il est scolarisé dans un établissement qui a tendance à constituer des classes dont la taille est plutôt inférieure à ce qui se fait dans le reste du canton.

Comme nous l'avons mentionné précédemment, ce résultat pourrait certes s'interpréter comme un impact positif de la réduction de la taille des classes sur les acquisitions des élèves (Piketty et Valdenaire, 2006; Meuret, 2001) mais on ne peut toutefois exclure qu'il s'agisse aussi d'un effet d'une toute autre nature, dans la mesure où tous les établissements concernés appartiennent à une même zone dans laquelle les conditions d'enseignement sont un peu particulières, ces dernières étant d'ailleurs plus le fruit d'une situation géographique atypique que le résultat d'un véritable choix de politique éducative. Dans cette zone géographique, le rapport à l'institution scolaire serait-il un peu différent de ce qui prévaut dans le reste du canton ? Y aurait-il une plus grande proximité de cette dernière ?

Dans le domaine des sciences, une autre variable explique également une partie des différences de résultats liées à l'établissement. Il apparaît en effet que dans les écoles où le corps enseignant est plus qualifié (proportion plus importante d'enseignants pleinement qualifiés disposant de l'habilitation à enseigner), les élèves tessinois obtiennent de meilleurs résultats.

Ceci fait écho à ce que l'on a pu observer dans le canton de Genève pour les mathématiques où, à caractéristiques individuelles identiques, les élèves obtiennent de moins bons résultats lorsqu'ils sont scolarisés dans un établissement où les enseignants de mathématiques ont une ancienneté plus faible que celle qui existe en moyenne dans l'ensemble du canton.

Ces deux résultats, qui semblent montrer que le niveau de qualification et l'expérience des enseignants sont des caractéristiques assez fortement corrélées aux compétences des élèves, sont assez conformes à ce que l'on trouve dans la littérature et n'appellent pas de commentaires particuliers.

Ce qui surprend en revanche davantage à leur sujet c'est qu'ils renvoient, tout comme le résultat sur la taille des classes au Tessin, à un contexte de scolarisation qui correspond finalement plus à la classe qu'à l'établissement. Dès lors, faut-il en déduire que la classe serait une unité d'analyse des variations d'acquisitions de compétences plus pertinente que l'école ? C'est en tous cas ce que semblent penser certains auteurs comme Duru-Bellat (2003) qui relève que les effets du contexte sont, en général, plus forts au niveau de la classe pour ce qui est des progressions scolaires.

Ces éléments pourraient à l'avenir nous guider vers d'autres recherches prometteuses, mais ils constituent malheureusement à l'heure actuelle des pistes qui n'ont pu être explorées dans le cadre de cette recherche puisque le niveau classe est absent des données PISA en raison d'un plan d'échantillonnage qui ne comporte que deux niveaux (des élèves dans des établissements).

Concernant le domaine des sciences pour le canton de Genève, deux autres caractéristiques des établissements ayant un effet sur les performances des élèves ont pu encore être identifiées. La première d'entre elles correspond à un certain type de gestion d'établissement. En effet, il apparaît que les élèves réussissent mieux, «toutes choses égales par ailleurs», lorsque les enseignants participent davantage à la gestion de l'établissement.

Il s'agit là d'un résultat qui est assez conforme à ce que d'autres ont mis en évidence. Comme le rappelle Bressoux (1994), tracer le portrait du bon directeur d'établissement (celui qui est efficace) est une tâche difficile mais les travaux de recherche ont tout de même permis d'identifier certains traits distinctifs. Parmi eux, on peut citer la capacité de faire partager certaines valeurs à l'ensemble de l'équipe éducative, un grand sens de la communauté et l'implication de l'ensemble des maîtres dans les décisions à prendre, autant de facteurs qui renvoient assez directement au résultat que nous avons mis en évidence.

La seconde caractéristique qui possède un effet statistiquement significatif sur les performances des élèves genevois en sciences est d'ordre organisationnel puisqu'il s'agit de la taille des établissements.

À caractéristiques individuelles identiques, les élèves scolarisés dans des établissements de taille supérieure à la norme, soit environ une centaine d'élèves de plus que la moyenne cantonale, obtiennent de moins bonnes performances.

Encore une fois, l'interprétation de ce résultat n'est pas vraiment aisée. Bien que la question du lien entre la taille des établissements et la réussite des élèves doive encore être considérée comme ouverte, on pourrait penser qu'il s'agit là d'une illustration de ce que relève Afsa (2014), sur la base des revues de la littérature produites par Slate et Jones (2005) et Leithwood et Jantzi (2009), à savoir que «les structures de petites tailles produisent de meilleurs résultats (réussite plus élevée aux tests ou aux examens, taux d'abandon au cours de la scolarité plus faibles). Autrement dit, la taille a un effet négatif sur la performance scolaire de l'établissement».

On notera cependant que, comme le souligne l'auteur, il ne s'agit pas là d'un effet causal qui impliquerait un impact direct de la taille de l'établissement sur les résultats des élèves mais plutôt d'un effet indirect qui passe par de multiples canaux.

Toutefois, comme pour la taille des classes dans le canton du Tessin, on ne peut exclure l'hypothèse que cet effet observé à Genève soit en réalité d'une toute autre nature et qu'il soit davantage lié à une caractéristique inobservée d'établissements un peu hors norme dans le paysage cantonal.

Ainsi, on a par exemple pu relever dans l'analyse détaillée des items du questionnaire établissement PISA que l'un des directeurs des deux établissements concernés a fait état d'un manque sévère d'enseignants qualifiés en sciences et en mathématiques³⁰. Le second établissement concerné déclare, quant à lui, une pénurie ou une inadéquation importante en matière d'ordinateurs pour le

³⁰ C'est le seul à avoir fourni cette réponse pour le domaine des sciences sur la vingtaine de directeurs interrogés et seul un autre directeur a fait état de la même situation pour les mathématiques.

travail en classe. On notera toutefois que ce n'est pas le seul établissement du canton dans cette situation (au total 5 directeurs sur 19 ont fourni cette réponse).

Enfin, pour le domaine de la lecture, on peut encore mentionner qu'au Tessin les élèves tendent, «toutes choses égales par ailleurs», à obtenir des résultats plus faibles lorsqu'ils sont scolarisés dans des établissements où le taux de non-promotion est supérieur à la norme cantonale. Ce résultat est difficile à interpréter car, comme nous l'avons indiqué précédemment, la note finale qui permet de déterminer la promotion ou non d'un élève ne dépend pas des seules compétences de ce dernier mais aussi d'autres facteurs pris en considération par l'enseignant (attitude de l'élève, intérêt pour la matière, moyennes des notes des épreuves écrites, etc.).

Par ailleurs, on peut également se demander si dans les établissements où le taux de non-promotion est supérieur à la moyenne, les enseignants sont plus sévères ou si le conseil de direction attribue un rôle plus important, d'un point de vue pédagogique, à la non-promotion, par rapport à ce qui se fait dans d'autres écoles.

Dans le canton de Genève, bien que l'on ait aussi pu mettre en évidence un effet de l'établissement sur les performances des élèves dans ce domaine (environ 5% des différences de résultats sont liées à l'établissement fréquenté), il n'a malheureusement pas été possible d'identifier des caractéristiques propres aux écoles qui expliquent de façon statistiquement significative ces différences.

Quelles pourraient être les raisons de ce constat ? Peut-être la lecture est-elle un domaine de compétences un peu particulier par rapport aux mathématiques ou aux sciences, à la fois compétence transversale et objet de savoir. Les compétences acquises dans ce domaine seraient-elles davantage liées à des connaissances extra-scolaires, à ce qui se passe en dehors de l'école (aspects culturels par exemple), rendant de fait plus difficiles les questionnements autour d'un effet établissement et de sa nature ?

En résumé, la présente étude nous a permis de mesurer tout l'intérêt de la modélisation multiniveau dans le traitement de la problématique des effets de contexte.

La mise en œuvre de cette méthodologie et d'un design de recherche approprié nous ont conduits à un certain nombre de réponses sur l'existence et la taille d'un effet de l'établissement sur les compétences des élèves à la fin de la scolarité obligatoire dans les cantons du Tessin et de Genève.

La question de la nature de cet effet reste, en revanche, largement ouverte. Si nous avons pu rejeter l'hypothèse d'un effet de composition sociale et académique sur les performances des élèves, nous n'avons pas été pour autant en mesure de dégager d'autres résultats qui auraient une portée relativement générale.

Nous avons certes pu identifier quelques facteurs relatifs au corps enseignant, à l'organisation, la gestion ou le fonctionnement des établissements qui semblent expliquer en partie pourquoi certains d'entre eux sont plus efficaces que d'autres, mais l'interprétation de ces résultats reste bien souvent délicate et l'absence d'effet systématique évident de ces facteurs dans tous les domaines ou tous les cantons nous oblige à reconnaître que la nature de l'effet établissement reste encore insaisissable.

La présente recherche nous a, en réalité, permis d'identifier des établissements un peu atypiques dans chacun des cantons concernés, des contextes de scolarisation particuliers qui manifestement produisent des effets sur les performances des élèves. Mais par quels canaux, quels processus ou quels mécanismes ? Tout ceci reste encore à explorer et nécessitera vraisemblablement le recours à des méthodologies complémentaires, plus qualitatives, et des données plus fines que celles contenues dans PISA ou dans des bases de données scolaires cantonales. Bien qu'extrêmement riches et précieuses pour la production de connaissances sur le fonctionnement de nos systèmes éducatifs, ces dernières semblent ici se révéler trop frustrées pour appréhender le phénomène dans toute sa complexité.

Par ailleurs, certains des résultats obtenus dans cette recherche nous laissent penser que d'autres unités d'analyse telles que la classe sont au moins aussi pertinentes que l'établissement pour comprendre les variations d'acquisitions de compétences.

Il faudrait donc, à l'avenir, être en mesure de conduire des recherches permettant de considérer simultanément tous ces niveaux pour mieux comprendre comment les différents facteurs s'articulent entre eux.

Ceci irait d'ailleurs dans le sens de ce que recommandent certains auteurs. Ainsi Duru-Bellat (2003) relève, sur la base de travaux anglo-saxons (Dreeben et Barr, 1998; Gamoran et Mare, 1989), qu'«on

admet aujourd'hui qu'il faut penser ensemble ce qui se joue au sein de ces niveaux hiérarchisés de contexte (les anglophones parlent de *nested hierarchical layers*) que sont l'établissement et la classe» (p. 191). L'auteur rappelle que l'idée principale de tels travaux est que «le niveau organisationnel supérieur joue sur l'efficacité du niveau inférieur en facilitant les formes d'organisation et de fonctionnement qui s'avèrent les plus porteuses d'efficacité» (ibid.).

En quelque sorte, il s'agira de rendre compte du fait que ce qui se passe au niveau de l'établissement en termes d'organisation, de gestion ou de leadership est susceptible d'affecter les processus qui prennent place au sein des classes.

La présente recherche ne constitue, au Tessin et à Genève, qu'un pas modeste dans l'exploration de la problématique des effets de contexte, qui demeure à ce jour un vaste champ d'études à investiguer.

8. Références

- AA. VV. (2017). *Linee guida per l'accoglienza e la frequenza scolastica degli allievi allogliotti nella scuola dell'obbligo*. Bellinzona: DECS. Repéré dans http://www4.ti.ch/fileadmin/DECS/DS/UIM/Cosa_facciamo/Diversi/Linee_guida_allievi_allogliotti_scuola_dell_obbligo.pdf
- Afsa, C. (2014). Une question de taille. Dans Transformations des parcours des élèves; implication des parents; performance des établissements. *Éducation et formations* (85). Paris: DEPP.
- Betts, J.R et Shkolnik, J.L. (1999). The Behavioral Effects of Variations in Class Size: The case of Math Teachers. *Educational Evaluation and Policy Analysis*, 21(2), 193-213.
- Blatchford, P., Goldstein, H., Martin, C. et Browne, W. (2010). A study of class size effects in English school reception year classes. *British educational research journal*, 28(2), 169-185.
- Bradshaw, C.P, Waasdorp, T.E. et Leaf, P.J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Official journal of the American academy of pediatrics*, 130(5), e1136-e1145.
- Bressoux, P. (1994). Note de synthèse. Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*, 108, 91-137.
- Bressoux, P. (2010). *Modélisation statistique appliquée aux sciences sociales*. Bruxelles : De Boeck.
- Bressoux, P., Coustère P. et Leroy-Audouin, C. (1997). Les modèles multiniveau dans l'analyse écologique: le cas de la recherche en éducation. *Revue française de sociologie*, 38(1), 67-96.
- Bressoux, P. et Pansu, P. (2003). *Quand les enseignants jugent leurs élèves*. Education et Formation, PUF.
- Caldas, S.J. et Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of educational research*, 90(5), 269-277.
- Centre for Multilevel Modelling. (2007). *Multilevel structures and classifications*. Repéré dans : <http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/structures.pdf>
- Consortium PISA.ch (2011). *PISA 2009 : Résultats régionaux et cantonaux*. Berne et Neuchâtel : OFFT/CDIP et Consortium PISA.ch
- Crahay, M. et Monseur, C. (2006). Différences individuelles et effets d'agrégation en ce qui concerne les performances en lecture. Analyse secondaire des données PISA 2000. In C. Houssemand, R. Martin et P. Dickes, *Perspectives de psychologie différentielle* (p. 23-34). Rennes : Presses universitaires de Rennes.
- Davezies, L. (2005). Influence des caractéristiques du groupe des pairs sur la scolarité élémentaire. *Education et formations*, 72, 171-199.
- Dreeben, R. et Barr R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, (61), 129-142.
- Dumay, X. (2004). Effet établissement : effet de composition et/ou effet des pratiques managériales et pédagogiques ? Un état du débat. *Les cahiers de recherche en éducation et formation*, 34.
- Dumay, X. et Dupriez, V. (2009). Contexte d'établissement et apprentissage des élèves. In Dumay, X. et Dupriez, V. (Eds.), *L'efficacité dans l'enseignement. Promesses et zones d'ombre* (p. 103-122). Bruxelles: De Boeck.
- Dumay, X., Dupriez, V. et Maroy, C. (2009). Ségrégation, effets de composition et inégalités de résultats (Penser les marchés scolaires Working Paper). Rappe : Université de Genève. Repéré dans <http://www.unige.ch/fapse/ggape/files/3114/1570/7332/DumayDupriezMaroy2.pdf>
- Duru-Bellat, M. (2003). Les apprentissages des élèves dans leur contexte: les effets de la composition de l'environnement scolaire. *Carrefours de l'éducation*, 16, 182-206.
- Duru-Bellat M., Danner M., Le Bastard-Landrier S., et Piquée C. (2004). *Les effets de la composition scolaire et sociale du public d'élèves sur leur réussite et leurs attitudes: évaluation externe et explorations qualitatives*. Dijon: IREDU.

Duru-Bellat, M., Le Bastard-Landrier, S., Piquée, C. et Suchaut B. (2004). Tonalité sociale du contexte et expérience scolaire des élèves au lycée et à l'école primaire. *Revue française de sociologie*, 45(3), 441-468.

Duru-Bellat, M. et Mingat, A. (1997). La constitution de classes de niveau par les collègues: les effets pervers d'une pratique à visée égalisatrice. *Revue française de sociologie*, 38, 759-790.

Elliott, J. (1996). School Effectiveness Research and its Critics: alternative visions of schooling. *Cambridge Journal of Education*, 26(2), 199-224.

Felouzis, G. et Charmillot, S. (2017). Les inégalités scolaires en Suisse. *Social Change in Switzerland*, 8. Repéré dans <http://www.socialchangeswitzerland.ch/?p=1094>

Gamoran, A. et Mare, R.D. (1989). Secondary school tracking and educational inequality: compensation, reinforcement, or neutrality? *American Journal of Sociology*, 94, 1146-1183.

Goddard, Y. L., Goddard, R. D., et Tschannen-Moran, M. (2007) A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers college record*, 109(4), 877-896.

Gorard, S. (2006). Is there a school mix effect ? *Educational Review*, 58(1), 87-94.

Greenwald, R., L. Hedges et R. Laine (1996). The Effect of School Resources on Student Achievement. *Review of Educational Research*, 66(3), 361-396.

Gustafsson, J-E. (2003). What Do We Know About Effects of School Resources on Educational Results?. *Swedish Economic Policy Review*, 10, 77-110.

Hanushek, E. (2002). Publicly Provided Education. Dans A. Auerbach et M. Feldstein (éd.), *Handbook of Public Economics*. Amsterdam: Elsevier.

Hanushek, E. (2003). The Failure of Input-Based Schooling Policies. *The Economic Journal*, 113, F64-F98.

Harker, R. et Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177-199.

Hedges, L. et Greenwald, R. (1996). Have times changed? The Relation Between School Resource and Student Performance. Dans G. Burtless (éd.), *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*. Washington D.C.: Brookings Institution Press.

Kreft, I.G. et de Leeuw, J. (1998). *Introducing Multilevel Modeling*. Thousand Oaks, CA: Sage.

Lauder, H. et Hughes, D. (1990). Social inequalities and differences in school outcome. *New Zealand Journal of Educational Studies*, 25, 37-60.

Leclercq, D, Nicaise, J. et Demeuse, M. (2004). Docimologie critique: des difficultés de noter des copies et d'attribuer des notes aux élèves. Dans Marc Demeuse, *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation* (273-292). Liège : Les éditions de l'Université de Liège.

Leithwood, K. et Jantzi, D. (2009). A Review of Empirical Evidence About School Size Effects: A Policy Perspective. *Review of Educational Research*, 79(1), 464-490.

Merle, P. (2007). *Les notes. Secrets de fabrication*. Paris: Presses Universitaires de France.

Meunier, M. (2011). Immigration and student achievement: evidence from Switzerland. *Economics of education review*, 30(1), 16-38.

Meuret, D. (2001). *Les recherches sur la réduction de la taille des classes*. Repéré dans <http://www.ladocumentationfrançaise.fr/var/storage/rapports-publics/024000197.pdf>

Monseur, C. et Crahay, M. (2008). Composition académique et sociale des établissements, efficacité et inégalités scolaires: une comparaison internationale. *Revue française de pédagogie*, 164, 55-66.

Moreau, J. (2007). *Compétences et contexte des élèves vaudois lors de l'enquête PISA 2003. Comparaisons entre cantons, filières et types d'élèves*. Lausanne: Unité de recherche pour le pilotage des systèmes pédagogiques.

Mottier Lopez, L. et Crahay, M. (Ed.). (2009). *Evolutions en tension. Entre la régulation des apprentissages et le pilotage des systèmes*. Bruxelles : De Boeck.

- Nidegger, C. (éd.). (2014). *PISA 2012 : compétences des jeunes Romands : résultats de la cinquième enquête PISA auprès des élèves de fin de scolarité obligatoire*. Neuchâtel : IRDP.
- OCDE. (2014). *Résultats du PISA 2012: Les clés de la réussite des établissements d'enseignement: Ressources, politiques et pratiques (Volume IV)*. Paris : Editions OCDE.
- Opdenakker, M.-C. et Van Damme, J. (2000). The Importance of Identifying Levels in Multilevel Analysis : An Illustration of the Effects of Ignoring the Top or Intermediate Levels in School Effectiveness Research. *School Effectiveness and School Improvement : An International Journal of Research, Policy and Practice*, 11(1), 103-130.
- Origoni P. (a cura di). (2007). *Equi non per caso. I risultati dell'indagine PISA 2003 in Ticino*. Ricerche in educazione. Ufficio studi e ricerche.
- Paterson, L. et Goldstein, H. (1991). New statistical methods for analysing social structures: an introduction to multilevel models. *British educational research journal*, 17(4), 387-393.
- Petrucchi, F., Soussi, A., Rastoldo, F., Guilley, E. et Nidegger, C. (2015). *Evolution des moyennes trimestrielles en 9^{ème} du cycle d'orientation: quels en sont les déterminants ?* Genève: Service de la recherche en éducation.
- Piketty, T. et Valdenaire, M. (2006). L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. *Les Dossiers Evaluations et Statistique*, 173.
- Regolamento della scuola media (del 18 settembre 1996). Repéré dans https://m3.ti.ch/CAN/RLeggi/public/index.php/raccolta-leggi/legge/vid/05_06
- Rocher, T. (2008). Que nous apprennent les évaluations internationales sur le fonctionnement des systèmes éducatifs? Une illustration avec la question du redoublement, *Education et formation*, 78, 63-68.
- Salvisberg, M., et Zampieri, S. (2015). *Valutazioni sotto esame. Piste esplorative per un confronto tra PISA e note scolastiche 2009 e 2012*. Locarno: Centro Innovazione e Ricerca sui Sistemi Educativi.
- Singe, J. D. et Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. USA: Oxford University Press.
- Slate J.R. et Jones C.H. (2005). Effects of School Size: A Review of the Literature with Recommendations. *Essays in Education*, 13.
- Snijders, T. et Bosker, R. (1999). *Multilevel Analysis: an introduction to basic and advanced multilevel modeling*. London: Sage.
- Spielhofer, T., Benton, T. et Schagen, S. (2007). A study of the effects of school size and single-sex education in English schools. *Research papers in education*, 19(2), 133-159.
- Suchaut, B. (2013). *PISA à l'épreuve des notes. Une autre lecture de la comparaison entre pays*. URSP, IREDU. Repéré dans https://iredu.u-bourgogne.fr/images/stories/Documents/Publications_chercheurs/2013/13030.pdf
- Thrupp, M. (1999). *Schools making a difference: Let's be realistic! School mix, school effectiveness and the social limits of reform*. Buckingham: Open University Press.
- Thrupp, M., Lauder, H. et Robinson, T. (2002). School composition and peer effects. *International journal of educational research*, 37, 483-504.
- Willms, J.D. (1992). *Monitoring school performance*. London: Falmer Press.
- Woessman, L. (2011). Cross-country evidence on teacher performance pay. *Economics of education review*, 30(3), 404-418.

