

Les effets du redoublement sur la réussite scolaire des élèves au Québec : une évaluation économétrique¹

Catherine Haeck²
Guy Lacroix³
Gino Santarossa⁴

Février 2022

Résumé

- Le *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire* prévoit qu'un élève qui ne satisfait pas aux exigences minimales de réussite de l'un des trois cycles du primaire ou de l'un des deux cycles du secondaire peut être amené à prolonger ce cycle d'une année.
- Cette étude visait à évaluer les effets de cette mesure sur le cheminement et la réussite scolaires des élèves au secondaire.
- Les données statistiques recueillies montrent que les élèves qui reprennent une année d'enseignement ont de moins bons résultats aux épreuves uniques du ministère de l'Éducation et, pour la plupart, sont moins susceptibles d'obtenir un diplôme du secondaire que les autres élèves.
- Ces observations sont appuyées par des écrits scientifiques qui tendent à décrire les effets délétères du redoublement sur le cheminement scolaire, bien que plusieurs recherches montrent aussi une contribution de cette mesure au succès des élèves.
- Dans cette étude qui repose sur plusieurs méthodes économétriques, les développements récents du domaine de l'apprentissage machine ont été exploités dans le but de valider la robustesse des effets estimés du redoublement scolaire.
- Les résultats sont sans équivoque. Le redoublement scolaire :
 - est lié à une baisse des résultats aux épreuves uniques du ministère de l'Éducation et des scores aux tests du Programme international pour le suivi des acquis des élèves (PISA);
 - est associé à une plus faible diplomation aux études secondaires;

¹ Les auteurs sont très reconnaissants envers Philippe Chabot, étudiant au département des sciences économiques de l'Université du Québec à Montréal pour sa collaboration à la recension des écrits. Ils remercient aussi la Direction générale des communications du ministère de l'Éducation pour la révision linguistique de ce rapport. Ils sont aussi très reconnaissants envers Valérie Saisset pour la relecture du rapport. Enfin, ils remercient Mylène Jetté et Caroline Bégin du Ministère pour leurs suggestions utiles.

² Département des sciences économiques, Université du Québec à Montréal.

³ Chaire en évaluation économique des programmes publics, Département d'économie, Université Laval.

⁴ Direction de la recherche et de l'analyse économique, ministère de l'Éducation.

- entraîne une hausse des taux de sortie sans diplôme ni qualification à la formation générale des jeunes;
 - contribue à l'obtention d'un diplôme chez les élèves qui optent pour une formation menant à un métier semi-spécialisé;
 - contribue à l'obtention d'un diplôme chez les élèves les plus jeunes de leur cohorte qui n'ont redoublé qu'au primaire.
- Une analyse coût-avantage indique que le redoublement scolaire, pour une seule cohorte, engendre des pertes socioéconomiques de 3,4 milliards de dollars à long terme.

Sommaire

Cette étude visait à évaluer les effets du redoublement au Québec sur quatre indicateurs de la réussite scolaire, à savoir les notes aux épreuves uniques du ministère de l'Éducation (MEQ), la diplomation en ce qui a trait au secondaire et à la formation qualifiante, le décrochage scolaire et les scores aux tests du Programme international pour le suivi des acquis des élèves (PISA) de 2000 à 2018. Le redoublement scolaire est une mesure prévue par la *Loi sur l'instruction publique* et le *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire* qui vise à soutenir la réussite scolaire des élèves qui éprouvent des difficultés d'apprentissage inhabituelles ou répétées. Une analyse descriptive montre que les élèves qui redoublent une année d'enseignement ont de faibles notes aux épreuves uniques et présentent un faible taux d'obtention du diplôme à la formation générale des jeunes (FGJ) tout en risquant davantage que les autres de quitter l'école sans diplôme ni qualification. Ces élèves ont également moins de succès aux tests du PISA. Les doubleurs sont aussi moins nantis sur le plan socioéconomique, ont davantage des parents sous-scolarisés et adhèrent en moins grand nombre à des projets éducatifs que les autres élèves.

Cette étude s'est toutefois heurtée aux difficultés inhérentes à toute évaluation des effets de mesures publiques. Le redoublement n'est pas aléatoire; il touche les élèves ayant de plus grandes difficultés que les autres. Le redoublement est une mesure de soutien à la réussite scolaire qui s'explique donc par plusieurs facteurs, notamment les aptitudes scolaires des élèves, le revenu de leur famille, la scolarité de leurs parents et le statut de défavorisation socioéconomique des établissements scolaires qu'ils fréquentent. La reprise d'une année et le succès scolaire peuvent aussi s'expliquer par les relations avec les camarades de classe et certains événements inattendus qui sont jugés défavorables à la réussite éducative. Plusieurs de ces facteurs peuvent s'avérer difficilement mesurables ou observables. Ainsi, les estimations des effets du redoublement scolaire pourraient être entachées d'erreurs statistiques si ces facteurs ne sont pas pris en considération.

Pour remédier à ces erreurs ou contribuer à les réduire, cette étude s'appuie sur deux sources de données. La première fait référence aux bases de données administratives du MEQ, qui regroupent des renseignements historiques sur le parcours scolaire des élèves. Ces bases de données totalisent un nombre important d'observations. En revanche, elles rassemblent très peu de renseignements socioéconomiques sur les élèves. La seconde porte sur les enquêtes du PISA, qui, bien que riches sur le plan des données socioéconomiques, ne concernent que les élèves âgés de 15 ans et comptent un nombre beaucoup moins important d'observations que les bases de données administratives.

Plusieurs méthodes économétriques ont été utilisées pour cette recherche, qui a profité des récents développements du domaine de l'apprentissage machine pour la vérification de la robustesse statistique des estimations. Les modèles classiques de la sélection de l'échantillon, de l'appariement par les scores de propension, de l'approche des variables instrumentales, du modèle de probit ordonné à traitement endogène et des forêts aléatoires causales sont autant de méthodes mises à profit pour prendre en considération les particularités statistiques de la reprise scolaire. Par ailleurs, les écrits sur le redoublement scolaire rapportent que les élèves les moins âgés de leur cohorte ont une propension plus importante que les autres à reprendre une année. Les analyses présentées dans cette étude montrent un phénomène identique, en particulier pour les élèves qui redoublent plus d'une fois. Les méthodes économétriques utilisées pour l'étude tirent parti de cette information pour contribuer à la fiabilité statistique des estimations.

Les résultats de l'étude sont sans équivoque : le redoublement ne favorise pas la réussite scolaire des élèves en difficulté et contribue même à compromettre leur succès ainsi que leurs chances d'obtenir un diplôme d'études secondaires à la FGJ. L'étude relève toutefois une exception : la reprise d'une année d'enseignement accroît en moyenne la diplomation chez les élèves qui envisagent de poursuivre ou qui poursuivent une formation préparatoire au marché du travail et qui souhaitent obtenir une qualification en vue de l'exercice d'un métier semi-spécialisé. Il est également estimé que la reprise d'une ou de plusieurs années de scolarité est liée à une baisse des notes aux épreuves uniques d'environ 5 à 12 points de pourcentage selon le moment de la reprise ou la discipline scolaire et selon que des filles ou des garçons sont assujettis à la mesure. En outre, il est estimé que le redoublement scolaire affecte davantage les filles que les garçons en ce qui concerne les résultats aux épreuves uniques du MEQ.

Les estimations indiquent aussi que cette mesure entraîne une baisse de 43 points de pourcentage pour la propension de l'élève à obtenir un diplôme d'études secondaires en 7 ans, et ce, davantage chez les garçons. De plus, le redoublement fait augmenter de 25 points de pourcentage la probabilité qu'un élève mette fin à son parcours scolaire sans diplôme ni qualification, peu importe la formation poursuivie. L'augmentation est de 8 à 15 points de pourcentage pour les départs de la formation générale des jeunes sans diplôme. Exceptionnellement, le redoublement scolaire est lié à une hausse de 18 points de pourcentage quant à la propension de l'élève à obtenir un certificat de formation à un métier semi-spécialisé.

Par ailleurs, les résultats obtenus à partir des données du PISA confirment les effets délétères de la reprise d'une année sur les résultats aux épreuves scolaires. Les scores aux épreuves de mathématiques, de lecture et de science sont inférieurs de 50 à 180 points en

moyenne si l'élève a repris une année d'enseignement avant l'âge de 15 ans. Enfin, une analyse coût-avantage montre que le redoublement scolaire est une mesure coûteuse pour la société québécoise, du moins en ce qui a trait aux élèves touchés négativement par ses effets. Pour une seule cohorte, le redoublement entraîne une baisse à long terme de 2 milliards de dollars des revenus anticipés sur le marché du travail, une réduction de 655 millions de dollars des recettes fiscales et une augmentation estimée à 753 millions de dollars des dépenses gouvernementales pour l'aide financière de dernier recours et la gestion de la criminalité au Québec.

1 Introduction

Au Québec, un élève du primaire ou du secondaire qui ne répond pas aux exigences de son grade scolaire est dans l'obligation de reprendre sa formation si la décision prise par les intervenants impliqués (enseignants, conseillers pédagogiques, etc.) et ses parents y est favorable. La reprise d'une année scolaire est motivée par l'idée qu'un élève qui ne termine pas avec succès un degré de scolarité est vraisemblablement sous-qualifié pour répondre aux exigences du palier suivant de sa scolarisation. Le *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire* présente le redoublement comme une prolongation de l'apprentissage scolaire et non comme une reprise officielle de cours du primaire ou du secondaire. Dans tous les cas, le redoublement vise à soutenir le cheminement éducatif d'un élève qui éprouve des difficultés importantes au regard de sa réussite scolaire.

En 2019-2020, 1,8 % des effectifs du primaire, soit 9 979 élèves, ont repris une année scolaire. Au secondaire, le taux de redoublement pour la même année se chiffre à 7,1 % des effectifs, soit 30 528 élèves. Ces taux ont diminué constamment depuis 2008-2009, année au cours de laquelle le gouvernement a restreint le nombre attendu de redoublements à un seul par élève. Malgré tout, le nombre de jeunes qui reprennent au moins une année scolaire demeure élevé et le coût financier de la mesure reste important. Le coût annuel moyen de la scolarité d'un élève du primaire ou du secondaire est estimé à 13 485 \$ pour 2017-2018 selon les plus récentes données de Statistique Canada. Ainsi, le coût financier du redoublement totalise plus ou moins 546,2 millions de dollars pour la seule année 2019-2020. L'évaluation de l'efficacité et de l'efficience budgétaire de la mesure s'avère donc à propos compte tenu de la méconnaissance de ses effets sur le cheminement et la réussite scolaires des élèves au Québec de même que des sommes qui y sont consacrées.

La pertinence de la reprise scolaire comme mesure de soutien au cheminement éducatif d'un élève est largement débattue aussi bien dans les milieux de l'éducation que dans la communauté scientifique. Ses défenseurs soutiennent que le redoublement offre la chance aux élèves en difficulté d'acquérir les compétences scolaires minimales et de développer leur sens de l'effort tout en contribuant à leur développement psychologique et social (Range *et al.*, 2012; Cham *et al.*, 2015). Les opposants à la mesure soutiennent en revanche que la reprise d'une année scolaire stigmatise l'élève et, du coup, ajoute un fardeau psychique et émotionnel difficile à supporter (Mathys, Véronneau et Lecocq, 2019). L'élève doit aussi s'adapter à de nouveaux camarades de classe dont la maturité physique et psychologique peut différer de façon importante de la sienne. Dans certains cas, les

relations perturbées avec les pairs expliqueraient les difficultés scolaires (Goos, 2013). De façon générale, le redoublement scolaire aurait des effets délétères sur le développement comportemental et psychosocial des élèves (Pagani *et al.*, 2001) et leurs résultats scolaires (McCoy et Reynolds, 1998). Plusieurs chercheurs ont évalué les effets du redoublement sur les notes et le décrochage (Roderick, 1994; Jimerson *et al.*, 2002; Stearns *et al.*, 2007). D'autres se sont penchés sur le statut économique de l'élève une fois atteint l'âge adulte (Eide et Showalter, 2001; Babcock et Bedard, 2011). Enfin, l'impact de la reprise scolaire sur le développement psychologique, émotionnel et social de l'élève a aussi été largement étudié.

Dans la plupart des recherches, l'estimation économétrique des effets du redoublement scolaire se heurte à l'existence de biais statistiques potentiels causés par la participation sélective des élèves à cette mesure de soutien éducatif. Les décisions en matière de rétention scolaire sont typiquement associées aux caractéristiques individuelles, familiales et socioéconomiques de l'élève, à savoir ses aptitudes et habiletés personnelles, sa maturité affective et sa motivation scolaire, ses problèmes relationnels avec ses pairs, des difficultés familiales inattendues ou certains comportements sociaux propres à l'adolescence (Allen *et al.*, 2009; Babcock et Bedard, 2011; Fruehwirth, Navaroo et Takahashi, 2016; Nagin *et al.*, 2003). Écarter ou négliger ces facteurs déterminants du redoublement scolaire au cours de l'analyse pourrait sévèrement fausser les effets estimés de cette mesure.

Cette étude visait à mesurer l'impact du redoublement sur quatre indicateurs du cheminement et de la réussite scolaires au Québec : les notes aux épreuves uniques de quatrième et de cinquième secondaire, la diplomation, le décrochage scolaire et les scores obtenus aux tests du Programme international pour le suivi des acquis des élèves (PISA) de 2000 à 2018. L'approche avait pour but de prendre en considération la disparité des effets du redoublement rapportée dans la littérature en opposant les résultats d'une variété de techniques d'évaluation. L'étude visait aussi à déterminer les effets hétérogènes du redoublement scolaire. En effet, il est possible que le redoublement s'avère délétère pour un groupe d'élèves en particulier et les amène à quitter hâtivement l'école sans diplôme ni qualification, mais profite par ailleurs à d'autres élèves sur le plan de l'amélioration des compétences scolaires ou de l'obtention d'un diplôme. Puisque le phénomène du décrochage scolaire touche un nombre plus ou moins important d'élèves, il pourrait s'avérer peu avisé de renoncer entièrement au redoublement comme mesure de soutien à la réussite éducative si certains en bénéficient.

Deux sources de données ont été employées pour évaluer l'impact du redoublement scolaire au Québec : les données administratives du ministère de l'Éducation (MEQ) et les données des enquêtes du PISA. L'utilisation de ces deux sources a permis de profiter de

leurs avantages respectifs. Les données administratives comptent un grand nombre d'observations et offrent des renseignements détaillés sur le parcours scolaire des élèves. Cependant, elles contiennent très peu de renseignements socioéconomiques sur eux. En revanche, les enquêtes du PISA, de type transversal, réunissent plusieurs déterminants de la réussite scolaire. Elles regroupent toutefois un nombre beaucoup moins élevé d'observations que les données administratives et ne concernent que les élèves âgés de 15 ans.

Les données administratives et d'enquête utilisées dans cette étude ont donc un potentiel limité pour corriger les problèmes statistiques pouvant être liés à la sélection selon les facteurs confondants inobservables. C'est pourquoi nous avons eu recours à plusieurs méthodes économétriques et avons exploité les forces de chacune d'elles. Cela nous a permis de vérifier la sensibilité des résultats dérivés des différentes données et approches statistiques. Les méthodes classiques de sélection de l'échantillon à un et à deux régimes ont notamment été privilégiées pour l'analyse des effets du redoublement sur les notes aux épreuves uniques. Ces notes présentent généralement une distribution normale, ce qui justifie l'emploi de ces méthodes. De plus, la méthode de l'appariement par les scores de propension a été mise à profit. Enfin, le modèle de probit ordonné à traitement endogène a été utilisé dans le cas de la modélisation des choix de diplôme. Toutes ces méthodes tirent avantage de la prise en considération du trimestre de naissance de l'élève comme renseignement supplémentaire permettant de mieux déterminer les effets du redoublement scolaire.

Cette étude a également profité des récents développements méthodologiques fondés sur l'apprentissage machine, qui permettent de mieux capter les interactions entre les diverses variables observées et, donc, d'augmenter la qualité du groupe de comparaison. En particulier, la méthode des forêts aléatoires causales a été utilisée en raison de sa capacité à désigner les meilleurs prédicteurs de la réussite éducative et à permettre un appariement optimal des doubleurs et des non-doubleurs. Cette méthode permet aussi de bien prendre en compte les données administratives de grande dimension, ce qui n'est pas toujours possible dans les modèles classiques.

En outre, une attention particulière a été portée aux effets du redoublement selon le cycle au cours duquel il se produit dans le cheminement scolaire. Cette distinction découle de résultats rapportés dans la littérature et selon lesquels la reprise scolaire peut ne pas avoir les mêmes effets au primaire qu'au secondaire. L'analyse distingue également les effets du redoublement selon le sexe de l'élève, compte tenu des écarts largement consignés en ce qui concerne la réussite scolaire des filles et des garçons.

Cette étude débute, à la section 2, par une description du système d'enseignement du Québec et des modalités qui régissent la reprise d'une année scolaire, à la suite de laquelle est présentée, à la section 3, une brève revue de la littérature. La section 4 montre les données utilisées et un portrait descriptif du redoublement scolaire. La section 5 consiste en un bref rappel de la problématique fondamentale de l'évaluation des effets d'un programme public et porte spécifiquement sur les problèmes liés au redoublement scolaire. Les méthodes économétriques y sont également décrites de façon succincte, en particulier les modèles classiques de la sélection de l'échantillon, de l'appariement par les scores de propension, de l'approche des variables instrumentales, du modèle de probit ordonné et des forêts aléatoires causales. La section 6 présente une analyse des résultats obtenus au moyen de ces méthodes. Elle débute par l'analyse des facteurs qui expliquent la propension d'un élève à redoubler une année scolaire et se poursuit avec celle des effets estimés du redoublement sur l'abandon scolaire, les notes aux épreuves uniques et la diplomation. Une analyse coût-avantage est ensuite proposée à la section 7 dans le but d'évaluer l'efficacité budgétaire et la rentabilité socioéconomique du redoublement scolaire. Enfin, la section 8 présente les conclusions de l'étude.

2 Redoublement scolaire au Québec

Au Québec, la *Loi sur l'instruction publique* (RLRQ, c. I-13.3) définit les responsabilités et les modalités du cheminement et de la réussite scolaires des élèves. Le *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire* (RLRQ, c. I-13.3, r. 8) prévoit trois divisions de l'enseignement, soit l'éducation préscolaire, l'enseignement primaire et le secondaire. « Les services d'éducation préscolaire ont pour but de favoriser le développement intégral de l'élève par l'acquisition d'attitudes et de compétences qui faciliteront la réussite de ses parcours scolaire et personnel et de lui permettre de s'intégrer graduellement dans la société. Les services d'enseignement primaire ont pour but de permettre le développement intégral de l'élève et son insertion dans la société par des apprentissages fondamentaux qui contribueront au développement progressif de son autonomie et qui lui permettront d'accéder aux savoirs proposés à l'enseignement secondaire. Les services d'enseignement secondaire ont pour but de poursuivre le développement intégral de l'élève, de favoriser son insertion sociale et de faciliter son orientation personnelle et professionnelle. Ils complètent et consolident la formation de base de l'élève en vue d'obtenir un diplôme d'études secondaires ou une autre qualification et, le cas échéant, de poursuivre des études supérieures » (*Régime pédagogique*, art. 2).

L'enseignement primaire s'organise sur trois cycles de deux ans chacun (*Régime pédagogique*, art. 15; *Programme de formation de l'école québécoise : éducation préscolaire, enseignement primaire*, p. 4). Le cycle est une période d'apprentissage au cours de laquelle les élèves acquièrent un ensemble de compétences disciplinaires et transversales leur permettant d'accéder aux apprentissages ultérieurs. Au primaire, c'est le directeur d'école, sur la proposition des enseignants, qui a la responsabilité de préciser les règles de passage et de classement qui font qu'un élève accède au cycle suivant ou continue ses apprentissages dans le même cycle une année additionnelle (*Loi sur l'instruction publique*, art. 96.15). La décision du passage d'un élève d'un cycle à l'autre s'appuie sur son dernier bulletin (*Régime pédagogique*, art. 28.).

Les règles de passage du primaire au secondaire ainsi que du premier au deuxième cycle du secondaire relèvent du centre de services scolaire ou de la commission scolaire. Il est à noter qu'au secondaire, l'enseignement est réparti sur deux cycles : le premier s'étend sur deux années scolaires et le deuxième, sur trois années scolaires (*Régime pédagogique*, art. 15). Pour ce dernier cycle, « le passage de l'élève d'une année à l'autre s'effectue par matière s'il s'agit d'un élève du parcours de formation générale ou du parcours de formation générale appliquée » (*Régime pédagogique*, art. 28). Le seuil de réussite est fixé à 60 % pour chaque matière, aussi bien au primaire qu'au secondaire (*Régime pédagogique*, art. 28.1).

« Au primaire, comme au premier cycle du secondaire, si le bilan des apprentissages révèle que l'élève ne satisfait pas aux exigences minimales de réussite du cycle, deux décisions peuvent être prises :

- Permettre la poursuite des apprentissages au cycle suivant en mettant en place des mesures de soutien qui tiennent compte des besoins de l'élève;
- Permettre à l'élève de poursuivre ses apprentissages dans le même cycle en mettant aussi en place des mesures de soutien qui tiennent compte de ses besoins. Cependant, l'école ne peut décider qu'une fois de prolonger le cycle par une année additionnelle pour un élève du primaire.

C'est dans le cadre d'un plan d'intervention, réalisé en concertation avec les intervenants, les parents et l'élève, que doivent être prises les décisions relatives au cheminement de ce dernier ainsi que celles portant sur les mesures d'appui à lui offrir pour favoriser sa progression⁵ ».

⁵ Ministère de l'Éducation, du Loisir et du Sport (2006). *Le cheminement scolaire de l'élève : questions et éléments de réponse; principales références dans les encadrements ministériels*. Québec : Gouvernement du Québec, p. 4.

Ainsi, « [l]e redoublement, compris comme la reprise intégrale et dans les mêmes conditions d'une année, que ce soit la première année du cycle ou la deuxième, ne constitue pas un moyen efficace de faire progresser l'élève parce que cela ne permet pas de tenir compte des apprentissages qu'il a déjà réalisés »⁶. En revanche, « [l]e redoublement ou la prolongation du cycle, compris comme la poursuite des apprentissages dans une logique de continuité, est possible, selon les encadrements légaux en vigueur, au terme de la durée prévue du Programme de formation, soit deux ans »⁷. Cependant, « [à] l'enseignement primaire et à la fin de la première année du secondaire, le directeur de l'école peut, exceptionnellement, dans l'intérêt d'un élève, lui permettre de rester une seconde année dans la même classe s'il appert de son plan d'intervention que cette mesure est celle qui, parmi celles possibles, est davantage susceptible de faciliter son cheminement scolaire » (*Régime pédagogique*, art. 13.1).

Dans le cas spécifique du passage du primaire au secondaire, « [l]e directeur de l'école peut exceptionnellement, dans l'intérêt d'un élève qui n'a pas atteint les objectifs et maîtrisé les contenus notionnels obligatoires de l'enseignement primaire au terme de la période fixée par le régime pédagogique pour le passage obligatoire à l'enseignement secondaire, avec le consentement des parents, après consultation de l'enseignant et selon les modalités déterminées par les règlements du ministre, admettre cet élève à l'enseignement primaire pour une année additionnelle, s'il existe des motifs raisonnables de croire que cette mesure est nécessaire pour faciliter son cheminement scolaire » (*Loi sur l'instruction publique*, art. 96.18).

Au secondaire, le cheminement de l'élève peut varier selon son degré de réussite ou ses champs d'intérêt scolaires et professionnels. Le passage de l'élève du premier au deuxième cycle du secondaire est permis par l'école si les conditions établies par le centre de services scolaire ou la commission scolaire ont été remplies. L'élève qui réussit la transition au deuxième cycle choisit, chaque année, le parcours de formation générale ou de formation générale appliquée (*Régime pédagogique*, art. 23.1). Lorsque les exigences du premier cycle du secondaire ne sont pas remplies, quelques possibilités s'offrent :

- « Permettre à l'élève de passer au 2^e cycle du secondaire avec des mesures de soutien en fonction de ses besoins [...] »;

⁶ Ministère de l'Éducation, du Loisir et du Sport (2006). *Le cheminement scolaire de l'élève : questions et éléments de réponse; principales références dans les encadrements ministériels*. Québec : Gouvernement du Québec, p. 3.

⁷ Ministère de l'Éducation, du Loisir et du Sport (2006). *Le cheminement scolaire de l'élève : questions et éléments de réponse; principales références dans les encadrements ministériels*. Québec : Gouvernement du Québec, p. 3.

- Permettre à l'élève de poursuivre une année de plus ses apprentissages au 1^{er} cycle du secondaire en mettant en place des mesures de soutien qui tiennent compte de ses besoins [...];
- S'il est âgé d'au moins 15 ans au 30 septembre de l'année scolaire au cours de laquelle il commence sa formation, il lui est possible de s'inscrire dans la formation préparatoire au travail ou dans celle menant à l'exercice d'un métier semi-spécialisé [...] »⁸.

De plus, « [l]'élève peut être admis à la formation préparatoire au travail s'il n'a pas atteint les objectifs des programmes d'études de l'enseignement primaire dans les matières langue d'enseignement et mathématique » (*Régime pédagogique*, art. 23.4). Cette formation prévoit notamment, en sus de la préparation au marché du travail, un enseignement en langue d'enseignement, en mathématiques et en langue seconde. À sa troisième année de formation, l'élève peut être admis au cours de préparation à l'exercice d'un métier semi-spécialisé s'il satisfait les conditions à cet égard⁹.

En dépit des directives de la *Loi sur l'instruction publique* et du *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire*, plusieurs élèves du Québec font le choix de quitter l'école sans diplôme ni qualification. Le redoublement, perçu comme une mesure de soutien à la réussite scolaire, vise donc, entre autres choses, à pallier l'abandon scolaire et à favoriser la persévérance des élèves dans leur cheminement scolaire.

3 Recension sommaire des écrits

Les effets du redoublement peuvent varier selon le degré de scolarité de l'élève, ses attributs psychologiques et comportementaux ainsi que son environnement familial et social, et s'échelonnent tout au long du parcours scolaire (Wu, West et Hughes, 2010). La perspective d'un redoublement pour les élèves en difficulté au secondaire peut amener certains à abandonner leur scolarité avant même le moment prévu de la reprise. Le retard occasionné par le redoublement peut engendrer des impacts psychologiques et

⁸ Ministère de l'Éducation, du Loisir et du Sport (2006). *Le cheminement scolaire de l'élève : questions et éléments de réponse; principales références dans les encadrements ministériels*. Québec : Gouvernement du Québec, p. 5.

⁹ Voir l'article 23.4 du *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire*.

comportementaux qui varient en fonction de l'âge et de la maturité de l'élève de même que du rattrapage nécessaire des connaissances. Ainsi, le redoublement et le « retard » social qu'il peut susciter sont un processus naturellement dynamique dont les effets psychologiques et comportementaux, positifs ou non, peuvent être ressentis en ce qui regarde la réussite de l'élève à différents moments de son parcours scolaire.

Nous aborderons d'abord les effets psychologiques du redoublement à court et à long terme, puis nous résumerons l'état des connaissances concernant ses effets sur le décrochage et le rendement scolaires.

3.1 Effets psychologiques du redoublement

La littérature sur les effets psychologiques et comportementaux du redoublement présente des résultats mitigés. D'un côté, le redoublement peut être souhaitable selon la théorie de la comparaison sociale (Festinger, 1954; Wu, West et Hughes, 2010). Celle-ci soutient que les individus définissent leurs habiletés selon les signaux qu'ils reçoivent de leur environnement en se comparant, par exemple, à leurs camarades de classe. D'un autre côté, les opposants à la mesure soutiennent que le redoublement peut conduire à une hausse des troubles du comportement (Byrd, Weitzman et Auinger, 1997) et du stress chez l'élève (Anderson, Jimerson et Whipple, 2005).

Plusieurs recherches s'appuient sur le concept de « Gros poisson-petit bassin », selon lequel les individus ont tendance à développer une meilleure image de soi en se comparant à leurs pairs lorsque ceux-ci sont moins performants qu'eux (Marsh et Parker, 1984). La reprise d'une année scolaire pourrait ainsi permettre aux doubleurs de se comparer à des élèves plus jeunes qu'eux et d'améliorer leur image de soi (Mariano et Martorell, 2017). Afin d'étudier l'hypothèse du BFLPE, Marsh et Parker (1984) de même que Marsh et ses collaborateurs (2017) utilisent des données recueillies au moyen d'un panel et provenant d'une enquête allemande sur la réussite en mathématiques des élèves du secondaire. Ils abordent les effets du redoublement sur un ensemble d'indicateurs psychologiques et scolaires liés à l'élève, de la première à la quatrième année suivant la reprise. Selon leurs résultats, le redoublement a des effets bénéfiques, au cours de l'année de la reprise, sur l'image de soi relativement aux compétences en mathématiques et sur l'anxiété par rapport aux épreuves de mathématiques. Toutefois, ces effets ne sont pas significatifs dans les années suivant le redoublement. Les résultats de cette étude sont discutables étant donné le faible nombre de doubleurs dans les deux premières vagues de l'enquête.

Par rapport à Marsh et à ses collaborateurs (2017), Hwang et Cappella (2018) considèrent seulement les effets à long terme du redoublement. Leur analyse, fondée sur la

méthode de l'appariement par les scores de propension, n'a montré aucun effet significatif du redoublement effectué en première ou en deuxième année du primaire sur un ensemble d'indicateurs psychosociaux une fois l'élève rendu en deuxième secondaire. Le fait que le redoublement au primaire n'ait pas d'effet positif sur l'image de soi lorsque l'élève arrive au secondaire rejoint les résultats mentionnés précédemment (Marsh *et al.*, 2017) et remet en question l'hypothèse du « big fish little pond » avancée par Marsh et Parker (1984), du moins si on considère un horizon à long terme.

Lamote et ses collègues (2014) abordent, pour leur part, les effets à court et à long terme du redoublement sur l'image de soi à l'aide de données longitudinales relatives à la performance scolaire au secondaire. Au moyen de la méthode de l'appariement par les scores de propension, ces auteurs ont constaté que l'image de soi des doubleurs est significativement positive à court terme, soit dans l'année de reprise au huitième échelon scolaire (deuxième secondaire). Toutefois, cette image de soi n'est pas différente de celle des non-doubleurs rendus à la dixième ou douzième année, ce que les auteurs désignent comme étant un effet à long terme. Ces résultats soutiennent l'hypothèse du « big fish little pond », mais cet effet est encore une fois seulement présent dans l'année de reprise. D'autres études ont également distingué l'effet sur l'image de soi générale de celui produit sur l'image de soi scolaire et en ont corroboré la nature temporaire (Ehmke, Drechsel et Carstensen, 2010; Bonvin, Bless et Schuepbach, 2008).

Les opposants au redoublement soutiennent qu'une telle mesure est susceptible de conduire à une hausse des troubles du comportement (Byrd, Weitzman et Auinger, 1997) et du stress chez l'élève (Anderson, Jimerson et Whipple, 2005). Allant dans le même sens, Paganini et ses collaborateurs (2001) ont constaté, à l'aide de données québécoises sur la performance scolaire, que le redoublement est associé, jusqu'à la fin du primaire, à un plus grand nombre de troubles de l'attention, de troubles de l'anxiété et de comportements perturbateurs en classe. Ces effets négatifs du redoublement doivent être pris en compte, particulièrement à la lumière des résultats de Segal (2013), qui indiquent un lien entre les problèmes de comportement à l'école et les salaires peu élevés à l'âge adulte.

Par ailleurs, bien que les résultats des méta-analyses de Holmes (1984) et de Jimerson (2001) montrent un lien négatif entre le redoublement et divers indicateurs de l'état socioémotionnel et du comportement des élèves, la littérature récente sur le sujet a depuis critiqué la stratégie d'identification des études considérées dans ces méta-analyses, en particulier le fait qu'elles ne permettent pas de déterminer l'effet causal du redoublement (Alexander, Entwisle et Dauber, 2003; Allen *et al.*, 2009; Hong et Raudenbush, 2005; Lorence *et al.*, 2002; Mariano et Martorell, 2013). La causalité de l'effet du redoublement peut

difficilement être déterminée, puisque les doubleurs ont plus de chances d'éprouver des difficultés scolaires et de manifester des troubles du comportement que les autres élèves.

Des études récentes ont permis d'estimer l'effet du redoublement sur le comportement de l'élève en tenant compte de ce biais de sélection (Mariano et Martorell, 2017; Hu et Hannum, 2020). Par exemple, Mariano et Martorell (2017) ont estimé, à l'aide d'une méthode de régression par discontinuité (logique « floue ou fuzzy »), l'effet causal du redoublement sur la fréquentation scolaire et la suspension des élèves. Leurs résultats montrent que cet effet varie de façon sporadique selon le nombre d'années depuis l'année de reprise et l'échelon scolaire en ayant fait l'objet. Ainsi, un redoublement qui a lieu entre la troisième et la cinquième année du primaire tend à améliorer la fréquentation scolaire, à réduire la probabilité d'être suspendu et à diminuer la durée des suspensions. Le redoublement qui se produit entre la sixième année du primaire et la deuxième secondaire tend plutôt à entraîner des effets mixtes sur la fréquentation scolaire et à augmenter la probabilité d'être suspendu ainsi que la durée des suspensions.

De leur côté, Hu et Hannum (2020) ont étudié l'effet causal du redoublement sur un indicateur de problèmes de comportement à partir de données longitudinales recueillies dans des écoles primaires de la province rurale du Gansu en Chine. Les résultats de cette étude montrent que le redoublement est lié à une hausse du nombre de problèmes de comportement, mais ne permettent pas d'établir un lien de causalité. Plus spécifiquement, l'estimation effectuée à l'aide de la méthode de différence en différence indique que le comportement des doubleurs s'est amélioré après le redoublement par rapport à leurs camarades de classe. Cependant, lorsque les auteurs tiennent compte de différences préexistantes entre les doubleurs et les non-doubleurs par la méthode de l'appariement, la différence entre les deux groupes en matière de comportement n'est plus significative.

En plus des répercussions sur le comportement de l'élève, le redoublement implique une adaptation sociale de la part de celui-ci. À ce sujet, Demanet et Van Houtte (2016) ont utilisé un modèle hiérarchique pour étudier l'effet du redoublement sur le nombre d'amis du même niveau scolaire en tenant compte de la proportion de doubleurs dans l'école. Cette étude souligne que le redoublement au secondaire est associé à un moins grand nombre d'amis pour l'élève concerné et que les jeunes qui fréquentent des écoles présentant des taux de redoublement élevés ont aussi tendance à avoir moins d'amis. Toutefois, le fait d'avoir redoublé au primaire n'influence pas le nombre d'amis une fois l'élève rendu au secondaire.

Bien que le redoublement soit associé seulement temporairement à un changement de comportement ou d'état psychologique, l'étude de Mathys, Véronneau et Lecocq (2019) indique que cette mesure est aussi liée à une capacité d'adaptation plus faible chez les

doubleurs que chez les autres élèves. Cette étude traite de l'ajustement psychosocial des doubleurs à l'arrivée au secondaire par rapport à leurs camarades de classe. D'une part, ces auteurs ont constaté que les doubleurs ont davantage de problèmes d'agressivité et de délinquance. D'autre part, ils manifestent également une baisse de l'estime de soi, de leur perception du soutien parental et de leur implication dans leurs relations avec leurs parents.

Peixoto et ses collaborateurs (2016) ont, pour leur part, observé que les effets du redoublement sur l'image de soi scolaire perdurent même lorsque l'on constate un rattrapage des apprentissages. En effet, ils ont relevé que, dans l'année suivant l'année de reprise, l'image de soi scolaire des doubleurs qui ne sont pas en voie de reprendre une année de nouveau demeure inférieure à celle de leurs camarades de classe qui n'ont jamais redoublé. L'étude dépeint également le profil des doubleurs quant à la motivation. Les élèves qui ont repris une année ou qui sont en voie de le faire se distinguent de leurs camarades de classe qui n'ont jamais redoublé par une moins grande importance accordée à la performance scolaire et une plus grande tendance à se fixer un objectif d'évitement, soit d'en faire le moins possible à l'école.

Enfin, Kretschmann et ses collaborateurs (2019) mettent en lumière l'importance du moment de l'appariement des élèves dans l'utilisation de méthodes d'appariement par les scores de propension. En effet, à l'aide de données longitudinales, ces auteurs font l'appariement dans l'année précédant le redoublement et montrent que, s'il est effectué en sixième année, il a des effets négatifs sur l'image de soi scolaire, l'intérêt pour l'école et la motivation à apprendre de l'élève dans les derniers mois de l'année scolaire précédant la reprise. Ce résultat révèle que l'appréhension du redoublement produit des effets avant même la réalisation de celui-ci. Allant dans le même sens que les études précédemment mentionnées qui portent sur l'image de soi, Kretschmann et ses collègues (2019) concluent que les effets du redoublement ne sont pas significatifs dans l'année suivant la reprise.

En somme, il semble que le redoublement au secondaire soit associé à une meilleure image de soi dans l'année de reprise (Marsh *et al.*, 2017; Lamote *et al.*, 2014). Cependant, il est lié à court terme à un rétrécissement du réseau social ainsi qu'à une baisse de la motivation et implique davantage de troubles du comportement et de l'anxiété (Demagnet et Van Houtte, 2016; Mathys, Véronneau et Lecocq, 2019; Kretschmann *et al.*, 2019; Pagani *et al.*, 2001; Anderson, Jimerson et Whipple, 2005). Alors que les effets à long terme sur l'image de soi semblent inexistantes (Marsh *et al.*, 2017; Lamote *et al.*, 2014; Hwang et Cappella, 2018), la littérature ne permet pas de conclure à des effets à long terme eu égard à l'adaptation sociale, à l'anxiété, au comportement et à la motivation. D'un côté, Pagani et ses collaborateurs (2001) indiquent que le redoublement au primaire est associé à un plus grand nombre de troubles du comportement et de l'anxiété jusqu'à la fin du primaire.

Mathys, Véronneau et Lecocq (2019) soulignent, pour leur part, que les conséquences sur l'adaptation sociale suivent les doubleurs jusqu'au secondaire. D'un autre côté, Hwang et Cappella (2018) n'ont observé aucun effet significatif du redoublement effectué au début du primaire sur des indicateurs de l'état psychosocial de l'élève rendu en deuxième secondaire. Demanet et Van Houtte (2016) n'ont pas constaté non plus d'effets du redoublement du primaire sur le nombre d'amis au secondaire.

Du côté des tenants du redoublement, un argument peu étudié dans la littérature concerne le fait que, lorsqu'il est accompagné d'un soutien particulier dans les apprentissages (accès à un orthopédagogue, à un psychologue ou à un tuteur, récupération, etc.), le redoublement pourrait permettre de surmonter les effets négatifs sur le comportement et l'état psychologique des élèves, et leur donner l'occasion de rattraper leur retard (Valbuena *et al.*, 2020).

3.2 Études portant sur le décrochage et le rendement scolaires

Hughes et ses collaborateurs (2017) ont étudié les effets du redoublement sur le décrochage scolaire à l'aide d'un modèle de régression logistique multinomiale à pondération par des scores de propension. Leurs résultats indiquent que les élèves qui ont redoublé entre la première et la cinquième année du primaire ont une probabilité plus élevée d'abandonner leur scolarité que les autres jeunes. Bien que plusieurs facteurs socioéconomiques soient pris en considération, les estimations de ces auteurs dépendent d'un faible échantillonnage et reposent sur l'hypothèse d'indépendance conditionnelle (HIC). Pour pallier les restrictions de cette hypothèse, Uysal (2010) a utilisé la méthode de régression « doublement robuste » et a conclu également à une plus forte propension à quitter l'école sans diplôme ni qualification pour les élèves ayant repris une année scolaire. Enfin, d'autres études ont donné des résultats comparables (ex. : Mariano *et al.*).

Par ailleurs, García-Pérez et ses collaborateurs (2014) ont utilisé les modèles de la sélection de l'échantillon avec changement de régime pour estimer l'effet du redoublement scolaire sur les scores obtenus aux enquêtes du PISA. Leur étude indique un effet défavorable du redoublement beaucoup plus marqué chez les élèves du primaire que ceux du secondaire. D'autres se sont appuyés sur les méthodes de régression par discontinuité (RD) afin de contourner les problèmes statistiques liés aux facteurs confondants inobservables de la réussite et du redoublement scolaires. Certains trouvent que la reprise d'une année scolaire est profitable aux élèves du primaire (Schwerdt, West et Winters, 2017; Greene et Winters, 2007). En revanche, Jacob et Lefgren (2004) ont plutôt constaté que la mesure est sans effet pour ces élèves. Par ailleurs, Özek (2015) conclut que la reprise

d'une année entraîne à court terme une hausse des incidents disciplinaires et des suspensions, principalement chez les garçons désavantagés sur le plan socioéconomique. Enfin, Mariano et ses collègues (2018), par l'emploi de la méthode de RD, ont constaté que le redoublement accroît le risque d'abandon scolaire chez les élèves du secondaire. Cette méthode est applicable uniquement si des critères conventionnés dictent la rétention d'un élève. Puisque de tels critères sont inexistantes au Québec, il n'est pas possible d'utiliser cette méthode.

Il a également été observé que les élèves les moins âgés de leur cohorte ont une propension plus importante que les autres à reprendre une année. Diris (2017) utilise l'approche des variables instrumentales pour exploiter ce phénomène. Ainsi, il instrumente la probabilité de redoublement scolaire par la date de naissance de l'élève et estime son modèle en exploitant les données du PISA. Ses résultats indiquent que le redoublement scolaire est dommageable à la réussite des jeunes élèves. Pedraja-Chaparro et ses collègues (2015) ont obtenu des résultats comparables.

La communauté scientifique s'est aussi intéressée à la dynamique de la participation au redoublement scolaire ainsi qu'à ses effets selon certains facteurs d'intérêt. Cockx et ses collaborateurs (2019) ont ainsi étudié les effets du redoublement à l'aide d'un modèle à choix discret dynamique. Leur étude montre que le redoublement n'a pas d'effet à court terme, mais qu'à moyenne échéance il nuit à la réussite éducative de l'élève, en particulier par un risque plus élevé de décrochage scolaire. Ces auteurs notent également une forte disparité dans les effets du redoublement, qui cause des préjudices scolaires beaucoup plus importants aux élèves les moins doués. Gary-Bobo et ses collègues (2016) ont constaté le phénomène inverse à l'aide de méthodes comparables. Selon eux, le redoublement entraîne une hausse des résultats scolaires chez les élèves éprouvant d'importantes difficultés, alors qu'il est préjudiciable aux élèves plus doués. Fruehwirth et les membres de son équipe (2016) estiment, pour leur part, que le redoublement appliqué aux tout-petits de la maternelle compromet l'apprentissage scolaire dans l'année qui suit, mais que cet effet tend à se dissiper au primaire.

Valbuena et ses collaborateurs (2020) se sont exclusivement concentrés sur des études reconnaissant le problème de sélection inhérent au redoublement et ont utilisé différentes stratégies pour contourner ce problème. Au total, ces auteurs ont recensé 42 études publiées de 2001 à 2020. Parmi ces études, 27 consignent les effets du redoublement sur les résultats scolaires à court terme. De ces 27 études, 19 soulignent des effets positifs à court terme sur les résultats scolaires (dont 15 au primaire) et 6 indiquent plutôt des effets négatifs. Par contre, parmi les dix études portant sur les résultats scolaires à long terme, une seule présente un effet positif, alors que six révèlent des effets négatifs.

Toujours à long terme, mais cette fois concernant la probabilité d'obtenir un diplôme d'études secondaires, quatre études sur six attestent des effets négatifs, une seule ne mentionne aucun effet et une autre indique un effet positif du redoublement au primaire. Finalement, sur huit études portant sur la probabilité de décrocher, neuf soulignent des effets négatifs et une seule portant sur le redoublement au primaire ne montre aucun effet. Ces résultats amènent les auteurs à conclure que le redoublement scolaire est une pratique peu efficace qui présente peu d'effets positifs à long terme, mais qui engendre des coûts importants, tant pour le système scolaire que pour l'individu lui-même. La recherche sur le sujet recommande donc de revoir cette pratique en l'adaptant pour accroître son efficacité ou de réduire son usage.

En somme, la littérature suggère des effets néfastes du redoublement, mais parfois aussi des effets positifs. Il n'existe donc pas de consensus clair. Bien entendu, l'ensemble des résultats des études dépend de la manière dont le redoublement scolaire est effectué et de l'offre ou non de services complémentaires. Étant donné le coût important de cette mesure, des effets de substitution sont possibles. Les systèmes qui utilisent peu le redoublement peuvent offrir plus de ressources complémentaires ou de services adaptés que les autres systèmes. Il est ainsi possible que les effets du redoublement varient. Pour comprendre les répercussions de cette mesure au Québec, il est donc nécessaire de les quantifier en ayant recours à des données québécoises.

4 Données et statistiques descriptives

4.1 Description des données

Deux sources de données ont permis d'évaluer les effets du redoublement sur les quatre indicateurs de la réussite scolaire à savoir les notes aux épreuves uniques du ministère de l'Éducation, la diplomation aux études secondaires et à la formation qualifiante, le décrochage scolaire à la formation générale des jeunes et les scores aux tests de PISA de 2000 à 2018. La première est la base de données longitudinales sur les élèves du Québec (BDLEQ) du MEQ. Cette base de données réunit des renseignements administratifs pour chaque élève inscrit au système d'éducation québécois depuis les vingt dernières années. Elle inclut le secteur dans lequel l'élève est inscrit (formation générale des jeunes [FGJ], formation générale des adultes [FGA], etc.), le niveau du jeune pour chaque année scolaire ainsi que les résultats aux épreuves de 4^e et de 5^e secondaire pour les élèves qui ont poursuivi à la formation générale des jeunes. La BDLEQ contient également des informations sur l'âge de l'élève, son pays de naissance, sa participation à des programmes préscolaires et scolaires adaptés, les écoles qu'il a fréquentées et sa ville de résidence. À

ces renseignements s'ajoutent des indicateurs dérivés, dont la fréquence des changements d'établissement scolaire de l'élève au cours des cycles de formation du primaire et du secondaire.

De plus, un indicateur dichotomique de la reprise d'une année scolaire a été créé selon le procédé suivant : la variable est égale à 1 si, pour une année, l'ordre d'enseignement de l'élève inscrit dans les fichiers administratifs est identique à celui de l'année précédente. La variable est aussi égale à 1 si le palier de scolarité de l'élève se répète au-delà d'une année. Par exemple, la variable de la reprise scolaire est fixée à 1 si l'élève qui a commencé sa première année du primaire en 2006-2007 reprend ce palier de scolarité en 2008-2009¹⁰.

Toutefois, la BDLEQ, dans sa forme initiale, renferme très peu de renseignements socioéconomiques sur les élèves. Pour la bonifier, des données provenant des recensements de 2006 et de 2016 de Statistique Canada y ont été ajoutées¹¹. Ces renseignements sont accessibles par aire de diffusion, soit la plus petite agglomération pour laquelle l'organisme fédéral rend officiellement disponibles des renseignements sur les familles. Comme on ne compte en moyenne que de 400 à 700 personnes par aire de diffusion, les unités géographiques sont relativement petites. Il s'agit, entre autres choses, de la diplomation chez les parents, du statut familial (conjoints unis, famille monoparentale) et du revenu familial.

Des variables dichotomiques ont aussi été construites pour indiquer si une aire de diffusion donnée compte une majorité de parents ayant un statut socioéconomique particulier. Par exemple, la variable *Statut de diplomation de la mère*, introduite dans la BDLEQ, prend la valeur 1 si plus de 66 % des mères résidant dans une aire de diffusion donnée n'ont aucun diplôme. Ces renseignements contribuent à approximer la réalité socioéconomique d'un élève. En revanche, ils posent l'inconvénient que tous élèves partagent hypothétiquement le même statut socioéconomique pour une aire de diffusion donnée. La variable *Statut de diplomation du père* a été construite de la même façon ainsi que celle qui indique si la mère détient seule la responsabilité de la garde de ses enfants.

La base de données longitudinales sur les établissements scolaires (BDLES) du MEQ est également mise à profit dans cette étude et permet de prendre en considération l'effet de l'environnement scolaire sur la réussite éducative des élèves. L'indice de milieu socioéconomique (IMSE) du MEQ, le pourcentage d'élèves nés à l'étranger qui

¹⁰ Cette démarche a fait l'objet d'une validation additionnelle grâce à la numérotation des paliers de scolarité par l'approche du cycle d'enseignement.

¹¹ L'appariement des renseignements socioéconomiques a été effectué par l'entremise des fichiers de conversion des codes postaux de Statistique Canada. Cet appariement reste imparfait pour la prise en considération des renseignements socioéconomiques offerts par Statistique Canada étant donné qu'il s'appuie sur les codes postaux.

fréquentent un établissement et le revenu familial moyen de l'effectif scolaire sont autant de variables disponibles dans cette base de données.

Il est présumé que la stabilité de l'enseignement au primaire peut être profitable au cheminement éducatif des élèves. La BDLES regroupe donc aussi quelques variables portant sur le corps enseignant. Le pourcentage d'enseignants titulaires à temps plein d'un établissement scolaire a été calculé à partir des renseignements fournis par les systèmes de données en ce qui regarde la tâche principale des enseignants. De plus, le nombre moyen d'années d'enseignement est un facteur à ne pas négliger tout comme la qualité des infrastructures, indiquée par l'année de construction de l'immeuble scolaire et le nombre d'élèves qui fréquentent l'établissement.

Cependant, les renseignements longitudinaux disponibles dans ces bases de données ne sont pas directement utiles aux analyses économétriques de cette étude, puisqu'il n'existe pas de suivi longitudinal des indicateurs de la réussite scolaire des élèves à l'exception du suivi des sorties sans diplôme ni qualification sur une courte période. Les données doivent donc être disposées en coupes transversales pour les élèves. Par conséquent, quatre cohortes ont été constituées afin d'estimer les effets du redoublement scolaire. Il s'agit des cohortes des élèves inscrits à la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007. Une base de données en coupes transversales a été élaborée pour chacune de ces cohortes à partir des renseignements de la BDLEQ et de la BDLES.

Les bases de données pour chaque cohorte regroupent les variables de la BDLEQ et de la BDLES déjà existantes, notamment le sexe et le trimestre de naissance de l'élève, ou de nouvelles variables conçues pour aborder de façon adéquate les régressions économétriques avec des données en coupes transversales. Ainsi, la variable du revenu capte le revenu moyen de l'effectif scolaire de la première année du primaire à la dernière année de la scolarité pour chaque élève. La variable dichotomique relative au statut de garde de la mère est fixée à 1 si, pour une année ou l'autre du parcours scolaire de l'élève, le recensement de Statistique Canada rapporte un pourcentage de mères chefs de famille monoparentale dans une aire de diffusion supérieur à 66 %. Les statuts dichotomiques d'obtention du diplôme pour le père et la mère sont conçus de la même façon.

Par ailleurs, la construction des renseignements transversaux pour les variables du parcours et de l'environnement scolaires de l'élève varie selon les prémisses associées à leur effet ou la fréquence des observations pour les cycles du primaire et du secondaire. Ainsi, une variable dichotomique indique si un élève a fréquenté une école privée durant deux années ou plus au secondaire. Une autre précise si l'élève a participé à un projet éducatif particulier à au moins une reprise au secondaire également. Deux variables dichotomiques servent à prendre en considération le jeune qui a changé d'établissement au

cours de son cheminement scolaire, à l'exception du changement qui a eu lieu lors de la transition du primaire au secondaire. L'une indique si l'élève a changé d'établissement à au moins une reprise au primaire et l'autre, si cela s'est produit au secondaire.

De plus, les données longitudinales sur les indicateurs de l'environnement scolaire sont réduites à des moyennes selon les cycles scolaires privilégiés. Les renseignements transversaux par élève pour chacune des cohortes regroupent donc la moyenne des pourcentages d'enseignants titulaires à temps plein de la première à la sixième année du primaire, le nombre moyen d'années d'ancienneté de ces enseignants pour la même période ainsi qu'un indicateur de la date moyenne de construction des établissements scolaires fréquentés par l'élève. La valeur moyenne de l'indice de milieu socioéconomique est comptabilisée pour la durée de son cheminement scolaire. Deux autres variables correspondent aux pourcentages moyens d'élèves nés à l'étranger pour le primaire et le secondaire.

4.1.1 Base de données du PISA et Tendances de l'enquête internationale sur les mathématiques et les sciences

Les données des sept cycles d'enquête du PISA, soit ceux de 2000 à 2018, ont également été utilisées pour cette étude. Ce programme international permet de mesurer les connaissances et les compétences en mathématiques, en lecture et en science d'élèves âgés de 15 ans. Contrairement aux données administratives, les enquêtes du PISA fournissent un vaste éventail de renseignements scolaires et socioéconomiques sur les élèves. Elles offrent donc la possibilité de tenir compte d'un plus grand nombre de facteurs explicatifs de la réussite scolaire. En revanche, elles restreignent les analyses concernant le redoublement aux élèves ayant fait les tests du PISA. Toutefois, ce groupe représente une plus grande proportion des élèves d'une cohorte que celui formé des jeunes ayant passé les épreuves de quatrième et de cinquième secondaire dans la BDLEQ, comme nous le verrons plus bas.

Le PISA a été lancé par l'Organisation de coopération et de développement économiques (OCDE). En plus des résultats de l'évaluation, ce programme permet de recueillir des informations auprès des élèves et des directeurs d'école ainsi que des parents pour certains cycles de l'enquête. Ces renseignements servent, entre autres, à caractériser l'élève sur un plan individuel, son école, le soutien qu'il reçoit et son milieu familial. L'évaluation du PISA est réalisée sur une base trisannuelle depuis l'an 2000. Sept cycles d'enquête ont eu lieu depuis. Toutes les provinces canadiennes participent à ce programme depuis sa première édition. Au total, celui-ci a permis de récolter les réponses de 167 748 élèves au Canada de 2000 à 2018.

Les données du PISA sont accessibles en ligne sur le site de l'OCDE. Pour chaque cycle d'enquête, les données des élèves et des directeurs d'école sont séparées en deux fichiers. Pour la construction de la base de données, tous les fichiers contenant les réponses des élèves et des directeurs d'école ont été traités. Pour chaque cycle d'enquête, les variables d'intérêt sur les écoles ont été sélectionnées et seules les données associées au Canada ont été retenues. Une procédure comparable a été appliquée aux fichiers contenant les réponses des élèves.

Par la suite, l'identifiant des écoles a été utilisé pour fusionner la base de données des élèves à celle des écoles. En fusionnant les données relatives à chaque cycle et en sauvegardant la base de données en ayant décodé, les sept bases de données associées à chaque cycle ont été regroupées pour constituer une seule banque de renseignements. Celle-ci comprend les réponses données par les élèves et les directeurs d'école à travers les sept cycles d'enquête au Canada. Enfin, 400 observations qui ne sont pas associées à une province dans les données du PISA ont été retirées. La base de données finale comprend plus de 167 348 observations recueillies au Canada de 2000 à 2018.

L'analyse statistique du redoublement scolaire s'appuie également sur l'étude Tendances de l'enquête internationale sur les mathématiques et les sciences (TEIMS) pour les élèves de quatrième et de huitième année. Le programme TEIMS a été lancé en 1995 par l'International Association for the Evaluation of Educational Achievement (IEA). Comme le PISA, il permet de mesurer la performance scolaire des élèves en mathématiques et en science. Contrairement au PISA, où la sélection est faite selon l'âge, les élèves sont sélectionnés dans le programme TEIMS selon leur niveau scolaire. Ce programme permet d'évaluer la performance des élèves dans les quatrième et huitième années.

Dans le cadre de cette étude, la base de données du programme TEIMS a été employée afin de vérifier la fiabilité des statistiques obtenues avec les données du PISA. En ce sens, uniquement les données du programme TEIMS recueillies auprès des élèves de deuxième secondaire ont été utilisées, puisqu'elles sont comparables aux données du PISA qui portent sur les élèves de 15 ans. Depuis son lancement en 1995, le programme TEIMS a été reconduit tous les quatre ans. Ses données s'étendent aujourd'hui sur sept cycles d'enquête. Toutefois, les données de 1999 ont été exclues des analyses étant donné qu'il n'était pas possible d'y repérer la province de résidence. Au total, la base de données de ce programme contient les réponses de 55 522 élèves de huitième année résidant au Canada.

Les données du programme TEIMS sont accessibles sur le site du TIMSS & PIRLS International Study Center. Pour chaque cycle d'enquête, les données de chaque province sont divisées en plusieurs fichiers, dont un pour les écoles et un autre pour les élèves. Pour réunir ces fichiers en une seule base de données, il faut d'abord fusionner les fichiers des

écoles et des élèves pour chaque province et dans chaque cycle d'enquête à partir de l'identification des écoles. Les fichiers de données de chaque province sont ensuite regroupés. Pour chacun de ces six fichiers de données (un par cycle d'enquête), les variables d'intérêt ont été sélectionnées et leur mnémonique a été uniformisée. Il est enfin possible de réunir ces fichiers pour former une seule base de données. Étant donné que seulement le Québec et l'Ontario ont participé à tous les cycles d'enquête, seules les observations de ces deux provinces ont été conservées. La base de données contient ainsi 48 010 observations effectuées à travers les 6 cycles d'enquête de 1995 à 2019.

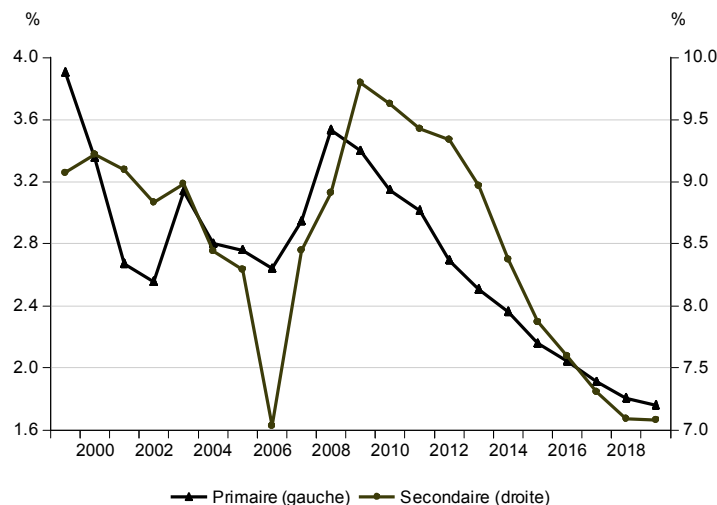
4.2 Profil statistique des élèves ayant redoublé une année

4.2.1 Profil statistique selon les données administratives du MEQ

Profil statistique de la reprise scolaire pour tous les élèves du primaire et du secondaire, de 1999-2000 à 2019-2020

Le pourcentage d'élèves en situation de reprise scolaire diminue au fil des ans, et ce, de façon particulièrement soutenue depuis 2008-2009, année au cours de laquelle le ministère de l'Éducation a restreint à un le nombre de redoublements par élève pour le primaire et le secondaire (graphique 1). Les élèves du secondaire sont, en proportion de leurs effectifs, plus nombreux que ceux du primaire à redoubler une année scolaire. En 2019-2020, 7,1 % des effectifs du secondaire ont repris une année, soit 30 528 élèves, alors qu'au primaire, le compte s'élevait à 9 979 élèves, soit 1,8 % des effectifs.

Graphique 1 — *Taux annuel de redoublement au primaire et au secondaire, de 1998-1999 à 2019-2020*

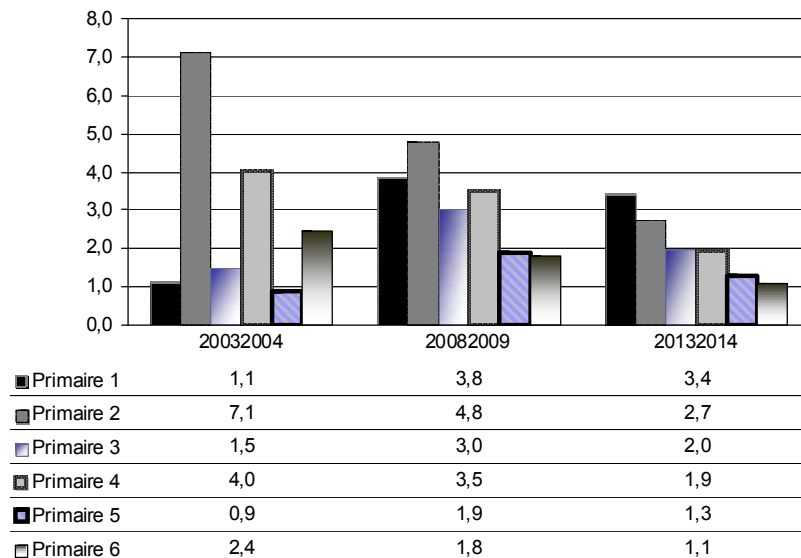


Source : MEQ.

Au primaire, les redoublements sont plus fréquents durant les premières années et, comme l'illustre le graphique 2, ce profil statistique change peu au fil des ans. De 2003-2004 à 2007-2008, les taux de redoublement étaient de 7,1 % pour la deuxième année du primaire, de 4,0 % pour la quatrième année et de 2,4 % pour la dernière année. Ces taux ont diminué de 2008-2009 à 2012-2013, mais le profil statistique des redoublements est

resté relativement inchangé à une exception près : les taux de redoublement à partir de la réforme de 2008-2009 sont semblables pour chacun des trois cycles du primaire. Par exemple, les pourcentages de reprise de la première et de la deuxième année du deuxième cycle du primaire se chiffrent respectivement à 3,0 % et à 3,5 %. Ceux de la première et de la deuxième année du troisième cycle du primaire sont respectivement de 1,9 % et de 1,8 %. Le graphique montre des écarts également peu importants pour les taux des années 2013-2014 à 2016-2017. Ainsi, bien que les directives du *Régime pédagogique* mentionnées à la section 2 prévoient une approche « cycle » pour l'acquisition des compétences, les données montrent que le redoublement est une mesure pédagogique utilisée à chacune des années de l'enseignement primaire.

Graphique 2 — Taux moyen de reprise d'une année scolaire au primaire, cohortes ayant fait leur entrée de 2003-2004 à 2007-2008, de 2008-2009 à 2012-2013 et de 2013-2014 à 2016-2017

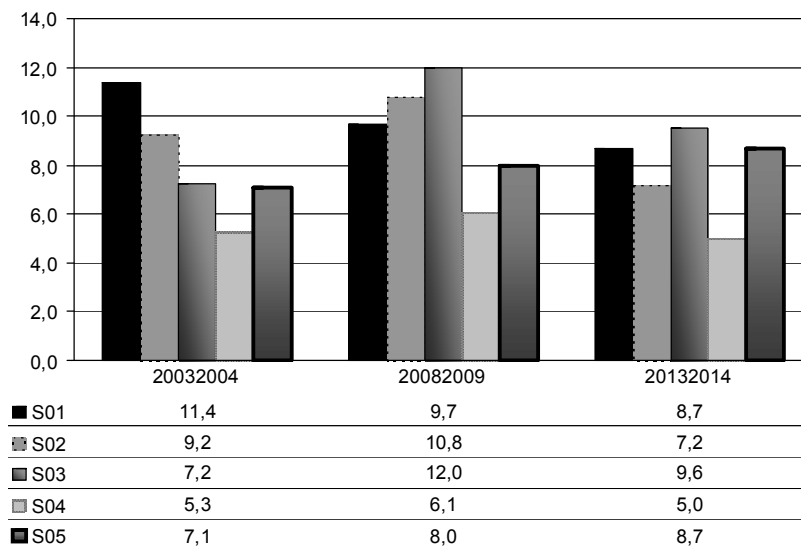


Source des données : MEQ

Au secondaire, le redoublement s'est avéré plus fréquent au cours des premières années, bien que ce profil statistique tende à être moins apparent à partir de 2013-2014 (graphique 3). De 2003-2004 à 2007-2008, les redoublements se sont produits essentiellement en première secondaire (11,4 %) et leur proportion a baissé aux paliers suivants, à l'exception de la cinquième secondaire, où ils ont remonté. Le profil des redoublements s'inverse toutefois depuis 2008-2009. Les taux de reprise d'une année scolaire sont tout aussi importants durant les trois premières années du secondaire, mais ils

augmentent plutôt que de diminuer d'un palier à l'autre. Ainsi, de 2008-2009 à 2012-2013, les taux moyens de redoublement se sont chiffrés à 9,7 % pour la première secondaire, à 10,8 % pour la deuxième secondaire et à 12,0 % pour la troisième secondaire. Le redoublement a ensuite chuté à la quatrième secondaire (6,1 %) pour remonter de nouveau à la dernière année du secondaire (8,0 %). Ce profil des redoublements est semblable pour les années 2013-2014 à 2016-2017. Le graphique montre par ailleurs que les taux de redoublement à la dernière année du secondaire tendent à s'amplifier au fil des ans.

Graphique 3 — Taux moyen de reprise d'une année scolaire au secondaire, de 2003-2004 à 2007-2008, de 2008-2009 à 2012-2013 et de 2013-2014 à 2016-2017



Source des données : MEQ

Les tableaux 1 à 4 montrent les fréquences de redoublement observées en 2019-2020 selon quelques caractéristiques sociodémographiques des élèves, ce qui permet de comparer le profil des élèves ayant repris une année avec celui des jeunes n'ayant jamais redoublé. Commençons tout d'abord par une comparaison selon le sexe de l'élève. Les garçons, tous ordres d'enseignement confondus, sont plus nombreux que les filles à reprendre une année scolaire (tableau 1). En 2019-2020, 40 507 jeunes avaient redoublé au moins une fois sur un total de 998 088 élèves. Parmi eux, 63,0 %, soit plus de 25 520 élèves, étaient des garçons. Ainsi, plus de 5,0 % des garçons avaient repris au moins une année scolaire en 2019-2020, alors que le redoublement ne touchait que 3,1 % des filles.

Tableau 1 — *Fréquences des redoublements scolaires, de la première année du primaire à la cinquième année du secondaire, filles et garçons, 2019-2020*

		Élèves sans reprise	Élèves avec reprise	Total
Garçons	Nombre	486 153	25 520	511 673
	(% ligne)	95,0	5,0	100,0
	(% colonne)	50,8	63,0	51,3
Filles	Nombre	471 428	14 987	486 415
	(% ligne)	96,9	3,1	100,0
	(% colonne)	49,2	37,0	48,7
Total nombre		957 581	40 507	998 088
Total (% colonne)		100,0	100,0	100,0

Source : MEQ.

Toujours en 2019-2020, le programme de la formation générale des jeunes au Québec comptait 110 961 élèves nés ailleurs qu'au Canada et aux États-Unis (tableau 2). Parmi eux, 6,2 % avaient repris une année scolaire, alors que le taux de redoublement pour tous les autres élèves se chiffrait à 3,8 %. De plus, le tableau 3 révèle que près d'un jeune sur cinq en 2019-2020 a éprouvé des difficultés d'apprentissage exigeant un plan d'intervention pour un total de 190 242 élèves. De ce nombre, 18 840 jeunes, soit près de 10,0 %, étaient des doubleurs. De tous les doubleurs en 2019-2020, 46,5 % éprouvaient des difficultés d'apprentissage accompagnées d'un plan d'intervention. Enfin, parmi les 45 893 jeunes vivant avec un handicap physique ou intellectuel, 17,0 %, soit 7 791 élèves, ont dû reprendre une année scolaire (tableau 4).

Tableau 2 — *Fréquences des redoublements scolaires, de la première année du primaire à la cinquième année du secondaire, élèves nés à l'étranger et au Canada ou aux États-Unis, 2019-2020*

		Élèves sans reprise	Élèves avec reprise	Total
Né au Canada ou É.U.	Nombre	853 475	33 652	887 127
	(% ligne)	96,2	3,8	100,0
	(% colonne)	89,1	83,1	88,9
Né à l'étranger	Nombre	104 106	6 855	110 961
	(% ligne)	93,8	6,2	100,0
	(% colonne)	10,9	16,9	11,1
Total nombre		957 581	40 507	998 088
Total (% colonne)		100,0	100,0	100,0

Source : MEQ.

Tableau 3 — *Fréquences des redoublements scolaires, de la première année du primaire à la cinquième année du secondaire, élèves avec ou sans difficultés d'apprentissage, 2019-2020*

		Élèves sans reprise	Élèves avec reprise	Total
Sans difficulté d'apprentissage	Nombre	786 179	21 667	807 846
	(% ligne)	97,32	2,68	100,00
	(% colonne)	82,10	53,49	80,94
Avec difficulté d'apprentissage	Nombre	171 402	18 840	190 242
	(% ligne)	90,10	9,90	100,00
	(% colonne)	17,90	46,51	19,06
Total nombre		957 581	40 507	998 088
Total (% colonne)		100,00	100,00	100,00

Source : MEQ.

Tableau 4 — *Fréquences des redoublements scolaires, de la première année du primaire à la cinquième année du secondaire, élèves avec ou sans difficultés d'apprentissage, 2019-2020*

		Élèves sans reprise	Élèves avec reprise	Total
Élèves sans handicap	Nombre	919 479	32 716	952 195
	(% ligne)	96,6	3,4	100,0
	(% colonne)	96,0	80,8	95,4
Élèves avec handicap	Nombre	38 102	7 791	45 893
	(% ligne)	83,0	17,0	100,0
	(% colonne)	4,0	19,2	4,6
Total nombre		957 581	40 507	998 088
Total (% colonne)		100,0	100,0	100,0

Source : MEQ.

L'environnement socioéconomique de l'élève est également jugé important dans sa réussite éducative. Le tableau 5 montre la part des effectifs scolaires de 2019-2020 en situation de redoublement pour cinq paliers (quintiles) du revenu des familles par aire de diffusion. Les chiffres montrent sans ambiguïté la prévalence plus importante des redoublements scolaires chez les élèves vivants dans des secteurs où les familles sont moins nanties sur le plan économique. Ainsi, dans des secteurs où le revenu familial moyen était inférieur à 66 924 \$ par année, 5,3 % des élèves avaient redoublé une année scolaire. Le pourcentage ne s'élevait toutefois qu'à 2,2 % pour les élèves vivant dans des secteurs où le revenu moyen des familles était supérieur à 120 703 \$ par année.

Tableau 5 — *Taux de redoublement selon les quintiles du revenu familial par aire de diffusion, de la première année du primaire à la cinquième année du secondaire, 2019-2020*

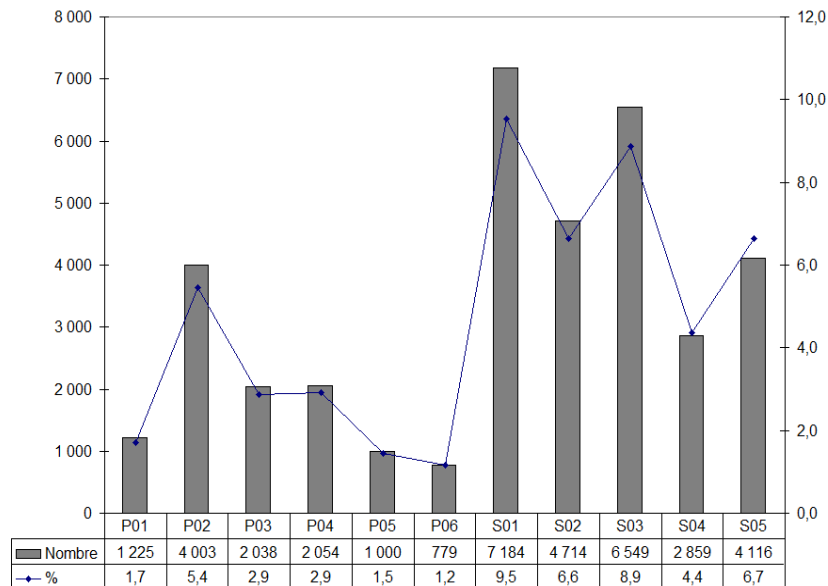
% Reprise scolaire	
Revenu familial (R)	
R ≤ 66 924 \$	5,3%
66 924 \$ < R ≤ 83 143 \$	5,0%
83 143 \$ < R ≤ 99 557 \$	4,0%
99 557 \$ < R ≤ 120 703 \$	3,2%
R > 120 703 \$	2,2%

Sources : MEQ, Statistique Canada.

Profil statistique de la reprise scolaire pour la cohorte de la première année du primaire en 2006-2007

Cette étude traite des cohortes des élèves inscrits à la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007, ce qui a permis d'évaluer l'impact du redoublement sur la réussite scolaire. Les statistiques ci-après concernent uniquement la cohorte de 2006-2007. Comme dans les statistiques présentées plus haut, les redoublements sont plus fréquents durant les premières années du primaire et du secondaire (graphique 4). Ainsi, 5,4 % des élèves de la cohorte de 2006-2007 ont dû reprendre la deuxième année du primaire, 2,9 % ont repris la troisième année et un pourcentage identique est observé pour la quatrième année. Le taux de redoublement est de 1,5 % pour la cinquième année du primaire et de 1,2 % pour la sixième année. Au secondaire, la fréquence des reprises grimpe jusqu'à 9,5 % et à 6,6 % pour la première et la deuxième année, et se chiffre à 8,9 % pour la troisième année. Les redoublements sont un peu moins importants durant les deux dernières années du secondaire.

Graphique 4 — *Fréquence de la reprise d'une année scolaire, cohorte de la première année du primaire en 2006-2007*



Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

Au total, 20 346 élèves, soit 29 % de ceux de la cohorte de 2006-2007, ont été dans l'obligation de reprendre une année ou plus (tableau 6). Parmi ces élèves, près de 11 000 n'ont redoublé qu'une fois, tandis que près d'un élève sur dix (8,1 %) l'a fait à deux reprises. Des cas de plus de deux redoublements sont également observés. Le tableau 7 indique que plus de la moitié (54,8 %) des redoublements se produisent uniquement au secondaire. Un peu plus de 3 000 élèves de la cohorte n'ont repris une année qu'au primaire et près de 30 %, soit 5 762 élèves, ont connu un épisode de redoublement tant au primaire qu'au secondaire.

Tableau 6 — *Nombres d'élèves ventilés selon le nombre de redoublements effectués entre la première année du primaire et la dernière année du secondaire, cohorte de la première année du primaire en 2006-2007*

Fréquence des reprises	Nombre d'élèves	%
0	49 895	71,0
1	10 815	15,4
2	5 705	8,1
3	2 232	3,2
4	895	1,3
5	389	0,55
6	186	0,26
7	70	0,10
8	34	0,05
9	8	0,01
10	7	0,01
11	5	0,01
<i>Total élèves avec reprise</i>	<i>20 346</i>	<i>29%</i>
Total	70 241	100,0

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

Tableau 7 — *Nombres d'élèves ventilés selon la période scolaire de redoublement, cohorte de la première année du primaire en 2006-2007*

Période de reprise	Nombre d'élèves	%
Uniquement au primaire	3 440	16,9
Uniquement au secondaire	11 144	54,8
Primaire et secondaire	5 762	28,3
Total	20 346	100,0

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

Les statistiques qui précèdent offrent une vue partielle de la dynamique du redoublement scolaire. Le tableau 8 présente les nombres d'élèves pour chacun des 45 premiers des 276 parcours de redoublement observés dans la BDLEQ. Les parcours sont affichés selon le nombre décroissant d'élèves. L'objectif de cette analyse est de déterminer les parcours les plus fréquents, ce qui sera utile ultérieurement pour la constitution de groupes de traitement. Les parcours les plus fréquents sont ceux où le redoublement se produit une seule fois durant le cheminement scolaire de l'élève. Ainsi, sur les 20 346 élèves ayant repris une année scolaire ou plus, 2 076 ont redoublé uniquement la première secondaire (10,2 %), 1 900, seulement la troisième secondaire (9,3 %) et 1 744, la deuxième secondaire (8,6 %). Enfin, un nombre comparable d'élèves (1 723, soit 8,5 %) n'ont redoublé que la dernière année du secondaire. Les autres parcours les plus fréquents caractérisés par un seul redoublement sont la quatrième année du secondaire (1 092 élèves), la deuxième année du primaire (1 010 élèves) et la quatrième année du primaire (679 élèves). De plus, les autres parcours indiqués au tableau 8 montrent que, pour plusieurs élèves, le redoublement se produit à différents moments du cheminement scolaire.

Tableau 8 — *Nombres d'élèves selon leur parcours de redoublement scolaire, cohorte de la première année du primaire en 2006-2007*

Élèves (N)	Reprise (N)	P1	P2	P3	P4	P5	P6	S1	S2	S3	S4	S5
2076	1	0	0	0	0	0	0	1	0	0	0	0
1900	1	0	0	0	0	0	0	0	0	1	0	0
1744	1	0	0	0	0	0	0	0	1	0	0	0
1723	1	0	0	0	0	0	0	0	0	0	0	1
1092	1	0	0	0	0	0	0	0	0	0	1	0
1010	1	0	1	0	0	0	0	0	0	0	0	0
679	1	0	0	0	1	0	0	0	0	0	0	0
591	2	0	0	0	0	0	0	1	0	1	0	0
560	2	0	1	0	0	0	0	1	0	0	0	0
531	1	0	0	1	0	0	0	0	0	0	0	0
518	2	0	0	0	0	0	0	0	1	1	0	0
383	1	0	0	0	0	1	0	0	0	0	0	0
331	2	0	1	0	0	0	0	0	0	1	0	0
279	2	0	1	0	0	0	0	0	1	0	0	0
271	1	0	0	0	0	0	1	0	0	0	0	0
268	2	0	0	0	1	0	0	1	0	0	0	0
263	1	1	0	0	0	0	0	0	0	0	0	0
218	2	0	0	0	0	0	0	1	1	0	0	0
215	2	0	0	1	0	0	0	1	0	0	0	0
191	2	0	0	0	0	0	0	0	0	1	1	0
182	2	0	0	0	0	0	0	1	0	0	1	0
181	2	0	0	1	0	0	0	0	0	1	0	0
178	2	0	0	0	1	0	0	0	0	1	0	0
165	3	0	1	0	0	0	0	1	0	1	0	0
161	2	0	0	0	0	0	0	0	1	0	1	0
156	2	0	0	0	0	0	0	0	1	0	0	1
136	2	0	0	0	0	1	0	1	0	0	0	0
134	2	0	0	0	0	0	0	0	0	1	0	1
134	2	0	0	0	1	0	0	0	1	0	0	0
133	2	0	0	0	0	0	0	1	0	0	0	1
121	2	0	0	1	0	0	0	0	1	0	0	0
103	2	1	0	0	0	0	0	1	0	0	0	0
102	2	0	0	0	0	1	0	0	0	1	0	0
98	2	0	1	0	0	0	0	0	0	0	1	0
97	2	0	1	0	0	0	0	0	0	0	0	1
95	2	0	0	0	0	0	0	0	0	0	1	1
90	2	0	0	0	0	0	1	1	0	0	0	0
82	3	0	0	0	1	0	0	1	0	1	0	0
79	2	0	0	0	1	0	0	0	0	0	1	0
77	2	0	0	0	0	0	1	0	0	1	0	0
74	2	1	0	0	0	0	0	0	0	1	0	0
73	2	0	0	0	0	1	0	0	1	0	0	0
69	3	0	0	1	0	0	0	1	0	1	0	0
67	3	0	1	0	0	0	0	0	1	1	0	0
63	2	0	0	0	1	0	0	0	0	0	0	1

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

Le tableau 9 dresse, pour sa part, un portrait comparatif des élèves avec ou sans redoublement de la cohorte de 2006-2007 sur le plan de leur réussite scolaire et de leurs environnements scolaire et socioéconomique. La colonne 1 du tableau indique que 6,9 % des élèves de la cohorte ont quitté le secondaire sans diplôme ni qualification. Le taux d'abandon se chiffre à 18,9 % pour les élèves ayant redoublé une année ou plus, alors qu'il se limite à 1,8 % pour les autres. Les notes aux épreuves uniques sont beaucoup plus faibles pour les élèves ayant redoublé une année scolaire. En outre, les résultats des élèves ayant repris une année sont à peine plus élevés que la note officielle de passage (60 %) établie pour les examens conventionnés du ministère de l'Éducation. En effet, pour les doubleurs, les notes moyennes aux épreuves uniques sont de 62,8 % pour le français écrit, de 61,9 % pour les mathématiques, de 62,5 % pour l'histoire, de 62,1 % pour la science et de 73,7 % pour la langue anglaise. La colonne 4 du tableau 9 montre des différences de résultats entre les deux groupes d'élèves, avec ou sans redoublement, statistiquement significatives.

De plus, les informations présentées dans le tableau 9 révèlent qu'un peu plus du quart seulement des élèves (27,8 %) qui reprennent une année scolaire obtiennent un diplôme d'études secondaires (DES), alors que le taux de diplomation se chiffre à 90,5 % chez les élèves au parcours standard. Il a été évoqué, dans la section 2, qu'il est possible pour les élèves du secondaire qui éprouvent des difficultés scolaires d'adhérer à des programmes écourtés, notamment aux programmes de formation menant à un métier semi-spécialisé. Ainsi, 17,3 % des élèves ayant repris une année scolaire ont obtenu un diplôme dans l'un ou l'autre de ces champs de formation. Enfin, 37,6 % des doubleurs ont poursuivi leurs études à la FGA sans toutefois avoir reçu un diplôme.

Par ailleurs, le tableau 9 montre d'importantes disparités familiales et socioéconomiques entre les élèves qui redoublent et les autres. D'abord, comme il a été mentionné plus haut, les filles sont moins nombreuses à reprendre une année que les garçons : seulement 37,8 % des doubleurs sont des filles. Ensuite, le pourcentage d'élèves nés à l'étranger dans le groupe des doubleurs (4,3 %) est moins important que celui rapporté pour les autres élèves (5,6 %). La compilation statistique montre également que les élèves les moins âgés de la cohorte, c'est-à-dire ceux nés durant le troisième trimestre de l'année, sont plus susceptibles d'avoir redoublé, alors même qu'une date d'anniversaire est somme toute aléatoire. Pour tous les autres trimestres de l'année, les pourcentages de doubleurs sont comparables sinon moins élevés. Ce lien entre la date de naissance et le redoublement est particulièrement utile, puisqu'il rejoint plusieurs études présentant des constats comparables (Pedraja-Chaparro, Santin et Simancas, 2015). La date de naissance servira notamment à instrumenter le redoublement dans les estimations économétriques.

En outre, les chiffres du tableau 9 montrent que le revenu familial des habitants du secteur de résidence des élèves qui reprennent une année (80 622 \$) est en moyenne plus faible que celui du secteur des élèves au parcours standard (98 019 \$). La sous-scolarisation des parents est également plus fréquente dans le groupe des doubleurs. Les calculs statistiques indiquent que 18,4 % des doubleurs résident dans des aires de diffusion où plus des deux tiers des mères sont sans diplôme ni qualification. La part est de 9,2 % en ce qui concerne les élèves de la même cohorte qui n'ont pas redoublé au primaire ou au secondaire. Quant aux doubleurs dont le père est sans diplôme, ils sont aussi plus nombreux (35,2 %) comparativement au groupe des non-doubleurs (23,8 %). De plus, les élèves qui redoublent sont, toutes proportions gardées, plus nombreux à être vraisemblablement issus d'une famille monoparentale¹².

Enfin, le parcours éducatif diffère pour les élèves ayant repris une année. Ceux-ci sont plus nombreux à avoir fréquenté la maternelle 4 ans à demi-temps (11,7 %)¹³ et leur participation à des programmes d'enrichissement scolaire (8,3 %) ou à des projets pédagogiques conventionnés (4,3 %) est moins importante par rapport aux autres élèves. Le taux de fréquentation des écoles privées est également plus faible pour ces élèves : 7,6 % seulement pour les doubleurs comparativement à une estimation de 30,1 % pour les non-doubleurs. Le tableau 9 indique aussi que le recours à un plan d'intervention, qui représente la première étape généralement établie avant la reprise d'une année scolaire, est beaucoup plus fréquent pour les élèves ayant redoublé (72,4 %). Enfin, un taux plus élevé d'élèves vivant avec un handicap intellectuel ou physique est observé dans ce groupe.

¹² En pratique, en 2016, 19,6 % des familles québécoises ayant au moins un enfant de moins de 24 ans avaient la mère comme seul parent et ce pourcentage a peu changé depuis le début des années 2000. Ainsi, dans notre mesure, le pourcentage d'enfants vivant avec une mère chef de famille monoparentale est sous-estimé. Il faut se rappeler que nos mesures caractérisent l'environnement de l'enfant au regard de l'aire de diffusion : <https://statistique.quebec.ca/fr/produit/tableau/familles-selon-la-structure-la-presence-denfants-et-lage-des-enfants-quebec-1986-2016>.

¹³ Le programme de la maternelle 4 ans à demi-temps est destiné aux enfants d'âge préscolaire de l'île de Montréal dont le statut socioéconomique est jugé défavorable, c'est-à-dire dont les valeurs de l'IMSE sont regroupées dans les déciles 8, 9 et 10 de la variable.

Tableau 9 — *Statistiques comparatives, élèves avec ou sans redoublement, cohorte de la première année du primaire en 2006-2007*

	Total (1)	Élèves avec reprise (2)	Élèves sans reprise (3)	Écart (4)	
Indicateurs cheminement et réussite scolaires					
Taux de décrochage scolaire FGJ	6,9%	18,9%	1,8%	17,1	*
Notes aux épreuves ministérielles					
<i>Français écrit</i>	73,3	62,8	75,1	-12,3	*
<i>Mathématique</i>	73,6	61,9	78,1	-16,2	*
<i>Histoires</i>	75,3	62,5	77,5	-15,0	*
<i>Sciences</i>	72,3	62,1	74,2	-12,1	*
<i>Anglais</i>	79,8	73,7	80,7	-7,0	*
Diplomation					
<i>Diplôme d'études sec. (FGJ,FGA)</i>	72,3%	27,8%	90,5%	-62,7	*
<i>Autres (CFMS, DEP,autres)</i>	5,7%	17,3%	1,0%	16,3	*
Sans diplomation observée	15,6%	37,6%	6,6%	30,9	*
Variables individuelles					
Filles	49,0%	37,8%	53,7%	-15,9	*
Né à l'étranger	5,2%	4,3%	5,6%	-1,3	*
Né au premier trimestre	25,2%	24,3%	25,5%	-1,2	*
Né au second trimestre	26,6%	26,6%	26,6%	0,0	
Né au troisième trimestre	24,8%	28,2%	23,4%	4,9	*
Né au quatrième trimestre	23,5%	20,9%	24,5%	-3,7	*
Variables socioéconomiques					
Revenu moyen familial	93 060	80 623	98 271	-17 648	*
Mère monoparentale	3,4%	5,0%	2,8%	2,2	*
Mère sans diplôme	11,9%	18,4%	9,2%	9,2	*
Père sans diplôme	27,2%	35,2%	23,8%	11,4	*
Variables scolaires					
Maternelle 4 ans demi-temps	6,9%	11,7%	4,9%	6,7	*
Passe-Partout	12,4%	13,2%	12,1%	1,1	*
Projets pédagogiques particuliers	13,0%	2,8%	17,3%	-14,5	*
Programme d'enrichissement	10,3%	5,0%	12,6%	-7,5	*
Plan d'intervention élèves en difficulté	35,8%	72,4%	20,5%	51,9	*
Élèves avec handicap et plan d'inter.	4,7%	12,7%	1,3%	11,4	*
Écoles privées	23,4%	7,6%	30,1%	-22,5	*
Nombre d'élèves	68 433	20 218	48 215		

* Statistiquement différent au seuil de confiance de 5%

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

L'analyse statistique a montré qu'un nombre important d'élèves de la cohorte de 2006-2007 n'ont repris qu'une fois une année scolaire. Le tableau 10 présente une comparaison des élèves ayant redoublé une seule fois durant leur parcours et de ceux l'ayant fait à plus d'une reprise, ce qui permet de déterminer si les deux groupes se distinguent sur les plans scolaire et socioéconomique. Les chiffres indiquent des taux de décrochage scolaire importants pour les deux groupes. Cependant, le taux d'abandon pour les élèves n'ayant redoublé qu'une fois (15,5 %) s'avère moins important que celui des jeunes l'ayant fait plus d'une fois (24,6 %). Un fait intéressant est que les élèves ayant redoublé une seule année obtiennent en moyenne la note de passage, tandis que ceux ayant repris deux années ou plus échouent à toutes les épreuves à l'exception de l'examen d'anglais, langue seconde. Le taux d'obtention d'un DES est également beaucoup plus important pour le premier groupe de doubleurs (41,9 %) que pour le second (12,1 %).

De plus, les chiffres indiquent un pourcentage de filles légèrement plus élevé dans le groupe d'un seul redoublement. Pour ce qui est des pourcentages d'élèves nés à l'étranger, ils sont identiques, peu importe le nombre de reprises. Cependant, les deux groupes diffèrent de façon particulière au regard de l'âge relatif des doubleurs. Comme l'indique le tableau 10, le pourcentage d'élèves dont la date de naissance se situe dans le troisième trimestre de l'année est nettement plus important dans le groupe des deux reprises ou plus (31,2 %) que dans celui d'un seul redoublement (26,5 %). D'ailleurs, pour celui-ci, le trimestre de naissance affecte peu le regroupement des doubleurs. Dans ce contexte, la maturité physique et émotionnelle de l'élève comparativement aux autres jeunes de sa cohorte d'origine occupe un rôle non négligeable dans sa propension à redoubler.

Par ailleurs, les élèves d'un seul redoublement sont légèrement avantagés sur le plan socioéconomique par rapport à ceux qui ont redoublé plus d'une fois. Les parents du premier groupe d'élèves ont en moyenne de meilleurs revenus, sont plus scolarisés et sont un peu plus nombreux, toutes proportions gardées, à s'occuper conjointement de l'éducation des enfants. Sur le plan scolaire, un plus grand nombre d'élèves du groupe de ceux ayant repris deux années ou plus ont fréquenté la maternelle 4 ans à demi-temps. Dans leur cas, le taux de participation se chiffre à 15,0 %, alors qu'il est évalué à 8,3 % pour les élèves n'ayant redoublé qu'une fois¹⁴. Ceux-ci participent également de façon plus importante à des projets éducatifs particuliers. Enfin, le pourcentage d'élèves qui bénéficient d'un plan d'intervention pour des difficultés d'apprentissage est élevé dans les

¹⁴ Il est ici important de rappeler que la maternelle 4 ans était offerte en priorité dans les milieux ayant un IMSE élevé afin de favoriser le développement des enfants. Ces statistiques ne permettent pas de quantifier l'efficacité de cette mesure, mais permettent uniquement de faire un lien entre le milieu de vie et la probabilité de redoubler.

Tableau 10 — *Statistiques comparatives, élèves ayant redoublé une seule fois et élèves ayant redoublé à deux reprises ou plus, cohorte de la première année du primaire en 2006-2007*

	Total (1)	1 seule reprise (2)	2 reprises ou plus (3)	Écart (4)	
Indicateurs cheminement et réussite scolaires					
Taux de décrochage scolaire FGJ	19,7%	15,5%	24,6%	-9,1	*
Notes aux épreuves ministérielles					
<i>Français écrit</i>	63,7	65,3	58,9	6,4	*
<i>Mathématique</i>	60,9	62,9	55,6	7,3	*
<i>Histoires</i>	61,7	63,0	55,0	8,0	*
<i>Sciences</i>	62,4	64,1	56,6	7,5	*
<i>Anglais</i>	73,5	73,7	72,6	1,1	*
Taux de diplomation					
<i>Tout diplôme</i>	51,7%	58,7%	43,5%	15,2	*
<i>DES 7 ans</i>	28,1%	41,9%	12,1%	29,8	*
Variables individuelles					
Filles	37,3%	39,8%	34,5%	5,2	*
Né à l'étranger	3,8%	3,8%	3,8%	0,0	
Né au premier trimestre	24,2%	25,5%	22,8%	2,7	*
Né au second trimestre	26,5%	26,4%	26,7%	-0,4	
Né au troisième trimestre	28,7%	26,5%	31,2%	-4,7%	*
Né au quatrième trimestre	20,6%	21,6%	19,3%	2,3	*
Variables socioéconomiques					
Revenu moyen familial	80 415	82 243	78 287	3 955,7	*
Mère monoparentale	5,1%	4,3%	6,0%	-1,6	*
Mère sans diplôme	18,7%	16,7%	21,1%	-4,5	*
Père sans diplôme	35,4%	33,2%	37,9%	-4,6	*
Variables scolaires					
Maternelle 4 ans demi-temps	11,4%	8,3%	15,0%	-6,7	*
Passe-Partout	13,7%	15,2%	12,0%	3,1	*
Projets éducatifs conventionnés	3,6%	5,2%	1,7%	3,5	*
Programmes d'enrichissement	8,0%	10,9%	4,8%	6,1	*
Plan d'intervention élèves en difficulté	75,3%	71,1%	80,2%	-9,1	*
Élèves avec handicaps	13,1%	6,7%	20,6%	-13,9	*
Nombre d'élèves	18 495	9 949	8 546		

* Statistiquement différent au seuil de confiance de 5%

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

deux groupes, mais s'avère un peu plus important pour les élèves qui ont redoublé à plus d'une reprise. La proportion d'élèves avec un handicap est aussi plus importante dans ce groupe.

Il est également utile de s'intéresser au parcours des doubleurs pour mieux comprendre les disparités dans leur cheminement scolaire et nuancer, au besoin, la portée des résultats économétriques de cette étude. Le tableau 11 dénombre les élèves ayant repris une année ou plus selon trois regroupements : les élèves sans diplôme ni qualification (décrochage de la FGJ), ceux ayant obtenu un certificat de formation ou un diplôme et les élèves sans diplôme recensé, c'est-à-dire ceux qui ont quitté la FGJ pour la FGA, par exemple, ou ceux dont l'obtention d'un diplôme ne peut être confirmée en raison de la non-disponibilité des plus récentes données.

Comme il a été indiqué plus haut, près d'un élève sur cinq (18,9 %) ayant repris une année scolaire quittera le secondaire sans diplôme ni qualification. Cependant, le risque de décrochage est plus faible (14,0 %) pour les élèves dont le redoublement ne s'est produit qu'au primaire. Un peu plus d'un jeune sur deux réussit à obtenir un diplôme ou une qualification, ce qui correspond à 10 674 élèves. Une majorité d'entre eux, soit 6 493 élèves, reçoivent un DES. Le pourcentage de diplômés aux études secondaires est plus important parmi les élèves qui ne redoublent qu'au primaire (50,0 %).

Enfin, près de 30,0 % des élèves dénombrés au tableau 11 sont passés de la FGJ à la FGA. Pour ces élèves, aucun diplôme n'est recensé soit parce que le cours suivi à la formation des adultes n'en offre aucun, qu'ils n'ont pas obtenu de diplôme ou qu'il est impossible pour le moment de confirmer leur statut à cet égard. Ce tableau indique également que la participation aux épreuves uniques est limitée chez les élèves qui redoublent. En effet, un peu plus du tiers seulement ont participé aux examens de français écrit, de science et de mathématiques (séquence « Culture, société et technique »).

Tableau 11 — *Parcours scolaires et participation aux épreuves uniques des élèves ayant repris une année scolaire, cohorte de la première année du primaire en 2006-2007*

Parcours	Reprise au primaire uniquement		Reprise au secondaire		Total	
	N	%	N	%	N	%
Décrochage scolaire	483	14,0%	3 345	19,8%	3 828	18,9%
Diplomation						
DES (FGJ, AD)	1 720	50,0%	4 773	28,2%	6 493	32,1%
CFMS, CFPT (FGJ,AD)	282	8,2%	2 346	13,9%	2 628	13,0%
DEP (FGJ,AD)	79	2,3%	420	2,5%	499	2,5%
TDG (AD)	78	2,3%	662	3,9%	740	3,7%
Autres	66	1,9%	248	1,5%	314	1,6%
<i>Total</i>	<i>2 225</i>	<i>64,7%</i>	<i>8 449</i>	<i>50,0%</i>	<i>10 674</i>	<i>52,8%</i>
Sans diplomation observée	604	17,6%	5 112	30,2%	5 716	28,3%
Total	3 312		16 906		20 218	100,0%
Participation aux épreuves FGJ						
Français écrit	1 987	57,8%	5 668	33,5%	7 655	37,9%
Sciences	1 756	51,0%	5 221	30,9%	6 977	34,5%
Mathématiques	1 471	42,8%	5 176	30,6%	6 647	32,9%

Source : Calculs des auteurs à partir des données de la cohorte entrée en première année du primaire en 2006-2007 et suivie jusqu'en 2019-2020, soit trois années de plus que le parcours normal. Données du ministère de l'Éducation.

4.2.2 Profil statistique selon les données des enquêtes du PISA

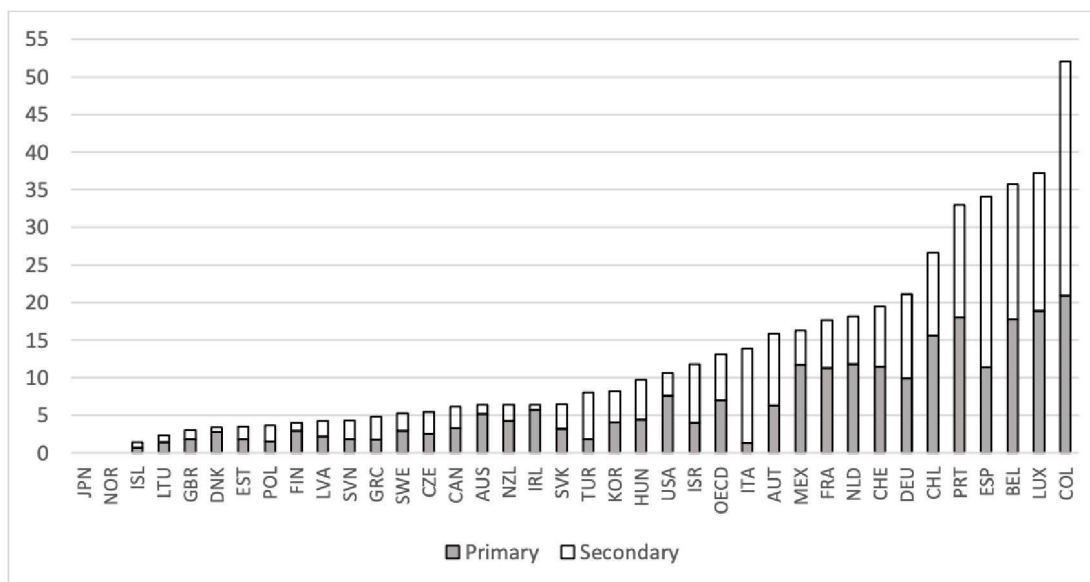
La question du redoublement est abordée dans les enquêtes du PISA menées en 2003, en 2009, en 2012, en 2015 et en 2018. Elle est posée dans le questionnaire destiné aux élèves. Cette question concerne le redoublement au primaire et au secondaire. À partir des réponses à cette question, il est possible de calculer le taux de redoublement au sein d'une province pour un cycle d'enquête donné. Dans l'analyse qui suit, le taux de redoublement indique le rapport entre le nombre d'élèves ayant repris au moins une année scolaire, dans la province et le cycle d'enquête en question, et le nombre d'élèves de cette province et de ce cycle.

Le graphique 5 montre avant tout que le redoublement scolaire est une pratique éducative qui varie énormément d'un pays à l'autre. Ainsi, un peu plus de 5 % des élèves du Canada ont repris une année scolaire, alors que la moyenne de l'OCDE se situe plutôt à près de 15 %. Le redoublement scolaire est plus important aux États-Unis, en France et aux Pays-Bas qu'il ne l'est au Canada. Ce graphique montre également que la reprise scolaire est utilisée principalement pour le soutien aux élèves du primaire dans plusieurs pays, dont

le Danemark, la Finlande, l’Australie, l’Irlande et les États-Unis. Au Canada, le taux de redoublement au secondaire avoisine 6 % comparativement à 3 % au primaire.

Graphique 5 — *Pourcentage d’étudiants ayant repris une année scolaire à l’âge de 15 ans, pays de l’OCDE, PISA 2018*

Figure 1. Students who have repeated at least one grade at age 15 (%). OECD countries, Programme for International Student Assessment (PISA) 2018



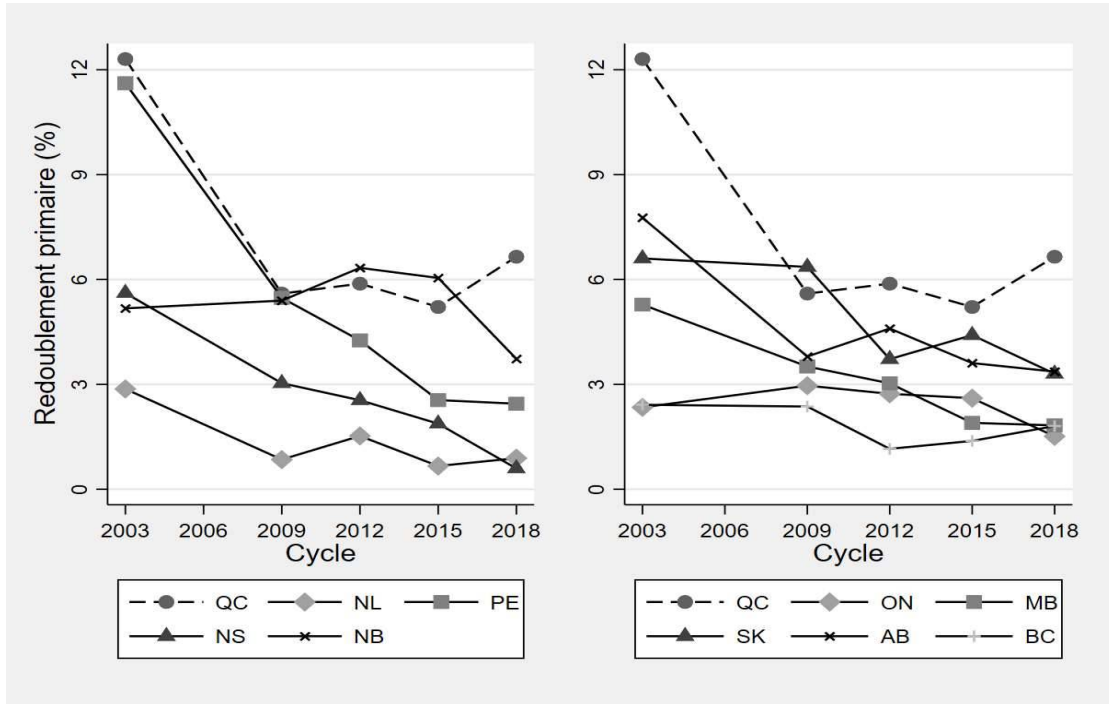
Le graphique 6 permet de porter un regard particulier sur l’évolution du taux de redoublement scolaire de chaque province canadienne à travers les cycles d’enquête du PISA. La section de gauche représente l’évolution de ce taux pour les provinces de l’Est et celle de droite, son évolution pour les provinces de l’Ouest¹⁵. Pour compléter ce graphique, le tableau 12 indique le taux de redoublement à travers les provinces et les cycles d’enquête du PISA.

Le Québec présente un taux de redoublement au primaire parmi les plus élevés au Canada. Ce taux est semblable à ceux du Nouveau-Brunswick. La différence entre le Québec et les autres provinces était particulièrement importante en 2003. Dès le cycle d’enquête suivant, on peut observer une diminution majeure du taux de redoublement au primaire au Québec. Cela correspond à un changement de régime par rapport au redoublement scolaire dans la province. Par la suite, de 2009 à 2018, alors que le Québec a maintenu son taux de redoublement pour la scolarité au primaire, les autres provinces ont constaté une baisse pour la même étape scolaire à chaque cycle d’enquête, si bien que le

¹⁵ Pour les graphiques suivants de cette section, les provinces ont été regroupées de la même façon en maintenant celles de l’Est à gauche et celles de l’Ouest à droite. Dans tous les graphiques qui suivent, le Québec est distingué par rapport aux autres provinces par le tracé d’une droite pointillée.

Québec est demeuré, jusqu'en 2018, parmi les provinces ayant le taux de redoublement au primaire le plus élevé.

Graphique 6 — Taux de redoublement au primaire par province à travers les cycles d'enquête du PISA



Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

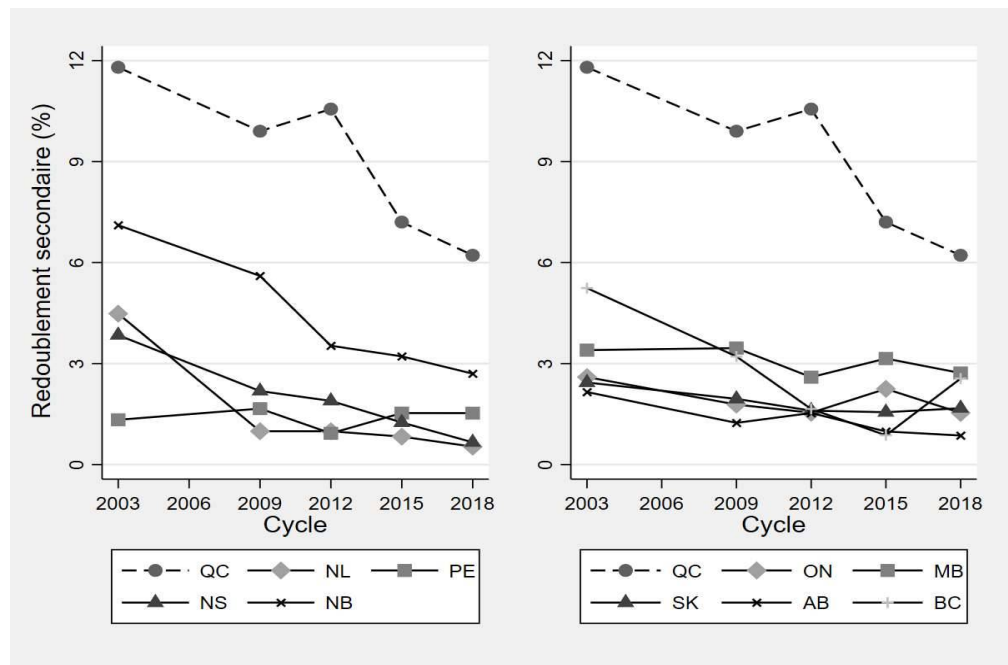
Tableau 12 — Taux de redoublement au primaire par province et par cycle d'enquête du PISA

	2003	2009	2012	2015	2018
NL	2,9	0,8	1,5	0,7	0,9
PE	11,6	5,5	4,3	2,6	2,4
NS	5,6	3,0	2,5	1,9	0,6
NB	5,2	5,4	6,3	6,0	3,7
QC	12,3	5,6	5,9	5,2	6,7
ON	2,3	3,0	2,7	2,6	1,5
MB	5,3	3,5	3,0	1,9	1,8
SK	6,6	6,4	3,7	4,4	3,3
AB	7,8	3,8	4,6	3,6	3,4
BC	2,4	2,4	1,2	1,4	1,8
CAN	6,0	4,0	3,8	3,2	3,1

Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Le graphique 7 présente l'évolution du taux de redoublement au secondaire, et ce, pour chaque province à travers les cycles d'enquête du PISA. Selon ce graphique, le Québec a le taux de redoublement au secondaire le plus élevé parmi tous les cycles d'enquête du

Graphique 7 — Taux de redoublement au secondaire par province à travers les cycles d'enquête du PISA



Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Tableau 13 — Taux de redoublement au secondaire par province et par cycle d'enquête du PISA

	2003	2009	2012	2015	2018
NL	4,5	1,0	1,0	0,8	0,5
PE	1,3	1,7	0,9	1,5	1,5
NS	3,8	2,2	1,9	1,3	0,7
NB	7,1	5,6	3,5	3,2	2,7
QC	11,8	9,9	10,6	7,2	6,2
ON	2,6	1,8	1,5	2,2	1,5
MB	3,4	3,5	2,6	3,2	2,7
SK	2,4	2,0	1,6	1,6	1,7
AB	2,2	1,2	1,5	1,0	0,9
BC	5,2	3,2	1,7	0,9	2,6
CAN	4,8	3,6	3,5	2,6	2,6

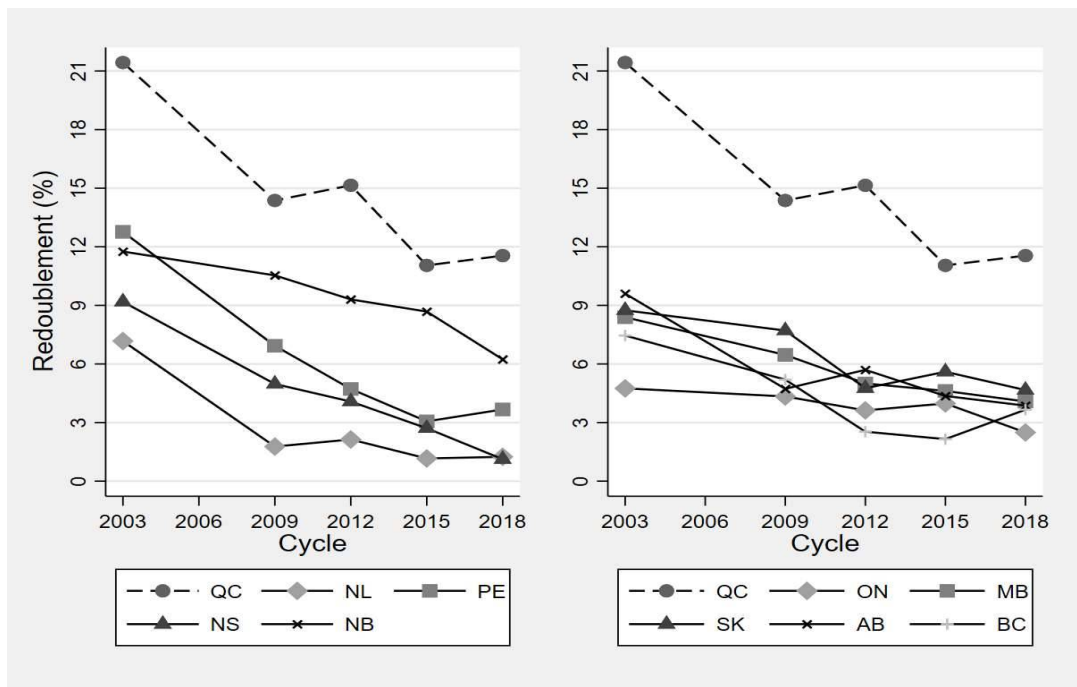
Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

PISA. Bien que le Québec observe une baisse de ce taux depuis 2003, celui-ci demeure au moins deux fois plus élevé ici que dans les autres provinces, qui constatent également une diminution de leur taux de redoublement au secondaire depuis 2003.

En regroupant les taux de redoublement du primaire et du secondaire, il est possible de calculer un taux combiné indiquant le nombre d'élèves ayant repris au moins une année au primaire ou au secondaire par rapport au nombre d'élèves de la province et pour un cycle d'enquête donné. Un élève ayant redoublé au primaire et au secondaire compte donc une seule fois dans le total des doubleurs. Cela explique pourquoi le taux affiché dans le tableau 14 ne correspond pas au total des taux des tableaux 12 et 13.

Le graphique 8 présente l'évolution du taux de redoublement au primaire et au secondaire pour chaque province à travers les cycles d'enquête du PISA. Il montre que le redoublement au Québec et dans les autres provinces a diminué. Toutefois, celui du Québec est de nouveau le plus élevé de toutes les provinces canadiennes à travers tous les cycles d'enquête.

Graphique 8 — Taux de redoublement au primaire et au secondaire par province à travers les cycles d'enquête du PISA



Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Tableau 14 — *Taux de redoublement au primaire et au secondaire par province et par cycle d'enquête du PISA*

	2003	2009	2012	2015	2018
NL	7,2	1,8	2,1	1,2	1,2
PE	12,8	6,9	4,7	3,1	3,7
NS	9,2	5,0	4,1	2,7	1,1
NB	11,8	10,5	9,3	8,7	6,2
QC	21,4	14,4	15,1	11,0	11,5
ON	4,8	4,3	3,6	4,0	2,5
MB	8,4	6,5	5,0	4,6	4,1
SK	8,7	7,7	4,8	5,6	4,7
AB	9,6	4,7	5,7	4,4	3,9
BC	7,5	5,2	2,5	2,2	3,7
CAN	10,2	7,1	6,7	5,2	5,1

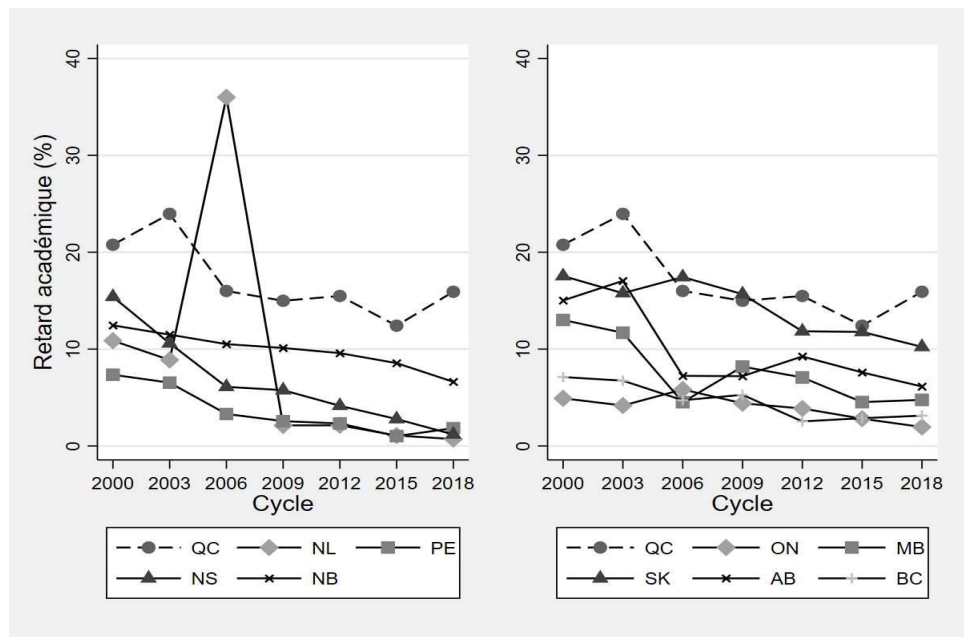
Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

À partir du niveau scolaire et du mois de naissance des élèves ayant participé aux enquêtes du PISA, il est possible de déterminer ceux qui sont en retard par rapport au niveau scolaire modal des jeunes de 15 ans. De cette façon, nous avons calculé le taux de retard scolaire pour chaque province à travers les sept cycles d'enquête du PISA¹⁶.

Ainsi, le graphique 9 présente l'évolution du taux de retard scolaire de chaque province à travers les cycles d'enquête du PISA. Tout comme le taux de redoublement, le taux de retard est généralement plus élevé au Québec que dans les autres provinces. Toutefois, cet indicateur semble être une mesure imparfaite du taux de redoublement, particulièrement pour la province de Terre-Neuve-et-Labrador (NL), qui affichait une valeur anormalement élevée en 2006.

¹⁶ Au Canada, le niveau scolaire modal des élèves de 15 ans est de 10, ce qui correspond à la quatrième secondaire. Au Québec, les élèves de 15 ans nés après le 30 septembre sont généralement en troisième secondaire, étant donné l'âge d'admission à l'école. Il faut en tenir compte dans le calcul du taux de retard scolaire. Cette particularité se trouvait aussi en Nouvelle-Écosse jusqu'en 2019, étant donné que cette province a modifié en 2008 le critère d'admission à l'école, le faisant passer du 1^{er} octobre au 31 décembre. Cette réforme n'affecte pas les retards scolaires de la cohorte de 2018, dont l'année de naissance est 2002. Les élèves de cette cohorte avaient donc tous déjà 5 ans (l'âge d'admission) au moment de l'entrée en vigueur de la réforme. L'Île-du-Prince-Édouard a, pour sa part, modifié l'âge d'admission à partir de 2003 en changeant le critère d'admission du 31 janvier et en reculant d'un mois chaque année jusqu'à ce que ce critère soit établi au 31 août. Les cohortes antérieures à l'année 2015 ont donc le 31 janvier comme critère, la cohorte de 2015 a le 31 novembre et celle de 2018, le 31 août.

Graphique 9 — Taux de retard scolaire par province à travers les cycles d'enquête du PISA



Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Tableau 15 — Taux de retard scolaire par province et par cycle d'enquête du PISA

	2000	2003	2006	2009	2012	2015	2018
NL	10,9	8,9	36,0	2,1	2,1	1,1	0,7
PE	7,4	6,5	3,3	2,6	2,3	1,0	1,8
NS	15,4	10,6	6,1	5,8	4,1	2,8	1,2
NB	12,5	11,5	10,5	10,1	9,6	8,6	6,6
QC	20,8	24,0	16,0	15,0	15,5	12,4	15,9
ON	4,9	4,2	5,8	4,4	3,9	2,8	2,0
MB	13,0	11,7	4,5	8,2	7,1	4,5	4,8
SK	17,5	15,8	17,4	15,7	11,8	11,8	10,2
AB	15,0	17,0	7,2	7,2	9,2	7,6	6,1
BC	7,1	6,7	4,7	5,3	2,5	2,9	3,1
CAN	12,7	11,9	11,2	8,1	7,8	6,1	6,6

Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Par ailleurs, la question du redoublement n'a jamais été posée aux élèves d'après les données canadiennes du programme TEIMS. Toutefois, il est possible d'utiliser la base de données de ce programme pour corroborer les taux de retard scolaire observés dans les données du PISA. Le tableau 16 présente le taux de retard scolaire calculé avec les données

du programme TEIMS pour le Québec et l'Ontario, qui sont les seules provinces à avoir participé à tous les cycles d'enquête. Les calculs indiquent un taux de redoublement au moins deux fois plus élevé au Québec qu'en Ontario. Bien que des différences puissent être notées entre les taux de redoublement du PISA et du programme TEIMS, il demeure qu'une tendance à la baisse est observable dans les deux provinces. De plus, le Québec montre un taux de redoublement beaucoup plus élevé que celui de l'Ontario, autant dans les données du PISA que dans celles du programme TEIMS.

Tableau 16 — *Taux de retard scolaire par province et par cycle d'enquête du programme TEIMS — Grade 8*

	1995	2003	2007	2011	2015	2019
QC	23,5	18,4	17,3	14,4	6,8	7,3
ON	9,7	4,6	3,5	2,9	0,4	0,3

Source : Calculs des auteurs à partir des microdonnées du programme TEIMS pour le Canada.

Caractéristiques des élèves qui ont redoublé dans les données du PISA

Les données du PISA contiennent plusieurs informations sur les élèves et les écoles qu'ils fréquentent, permettant ainsi de caractériser l'échantillon de ceux qui redoublent une année scolaire. Ces caractéristiques sont l'âge, le sexe, le niveau scolaire, la langue parlée à la maison, le statut d'immigration, le type de structure familiale, le nombre de livres à la maison, les niveaux d'éducation du père et de la mère, la nature privée ou publique de l'école et la taille démographique de la ville dans laquelle elle est située.

Les tableaux 17 à 19 présentent les caractéristiques sociodémographiques des élèves avec ou sans redoublement ¹⁷. Plus particulièrement, ils montrent, pour chaque caractéristique, le nombre d'élèves correspondant suivi, entre parenthèses, par la proportion de ceux possédant cette caractéristique au sein de leur groupe. La première colonne indique ces statistiques pour le groupe de doubleurs, alors que la seconde colonne fait référence au groupe de non-doubleurs. Ces statistiques révèlent, entre autres, qu'une plus grande proportion de doubleurs sont des garçons et que le pourcentage de garçons parmi les doubleurs est vraiment très près de ce qui a été observé plus haut pour la BDLEQ. On remarque de plus que les doubleurs vivent plus souvent dans une structure familiale

¹⁷ Les données du PISA pour 2000 et 2006 n'ont pas été utilisées pour calculer ces statistiques étant donné que la question du redoublement n'a pas été posée lors de ces deux cycles d'enquête.

monoparentale (27 %) ou mixte (9 %), possèdent moins de livres à la maison, étudient davantage dans des écoles publiques et ont des parents moins scolarisés. Ces statistiques sont conformes à celles des données administratives.

Tableau 17 — Répartition des doubleurs et des non-doubleurs selon leurs caractéristiques socioéconomiques

Caractéristiques	Doubleurs	Non-doubleurs
Secondaire 3	5 071 (62.40%)	9 462 (8.88%)
Secondaire 4	2 116 (26.04%)	95 531 (89.67%)
Garçon	4 855 (59.65%)	52 026 (48.83%)
Langue à la maison correspond à celle du test	6 326 (81.05%)	85 469 (83.66%)
École privée	325 (4.10%)	8 110 (7.79%)
Natif	6 466 (82.88%)	83 188 (81.84%)
Immigrant de première génération	492 (6.31%)	9 302 (9.15%)
Non-natif	844 (10.82%)	9 163 (9.01%)
Village (moins de 3000 hab.)	1 309 (16.60%)	17 044 (16.44%)
Petite ville (3 000 à 15 000 hab.)	1 937 (24.57%)	22 633 (21.82%)
Ville de taille moyenne (15 000 à 100 000 hab.)	1 968 (24.96%)	21 825 (21.05%)
Grande ville (100 000 à 1 000 000 hab.)	1 868 (23.69%)	30 323 (29.24%)
Grands centres urbains (plus de 1 000 000 hab.)	803 (10.18%)	11 880 (11.46%)
Nombre d'élèves	8 139	107 276

Note : La répartition en pourcentage au sein de chaque groupe est affichée entre parenthèses.

Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Tableau 18 — Répartition des doubleurs et des non-doubleurs selon leurs caractéristiques sociodémographiques

Caractéristiques	Doubleurs	Non-doubleurs
<u>Types de structure familiale</u>		
Famille monoparentale	1 528 (26.91%)	9 410 (14.93%)
Famille nucléaire	3 499 (61.61%)	50 658 (80.40%)
Famille mixte	501 (8.82%)	2 308 (3.66%)
<u>Nombre de livres à la maison</u>		
Aucun livre	1 636 (20.71%)	9 299 (9.12%)
1 à 10 livres	1 527 (19.33%)	12 905 (12.66%)
11 à 50 livres	2 321 (29.38%)	31 048 (30.46%)
51 à 100 livres	1 200 (15.19%)	21 911 (21.49%)
101 à 250 livres	730 (9.24%)	17 374 (17.04%)
251 à 500 livres	486 (6.15%)	9 404 (9.22%)
Nombre d'élèves	8 139	107 276

Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

Tableau 19 — Répartition des doubleurs et des non-doubleurs selon leurs caractéristiques sociodémographiques

Caractéristiques	Doubleurs	Non-doubleurs
<u>Niveau d'éducation du père</u>		
Père - ISCED 0	116 (1.57%)	465 (0.47%)
Père - ISCED 1	295 -4%	1 379 (1.38%)
Père - ISCED 2	982 (13.31%)	6 715 (6.74%)
Père - ISCED 3	137 (1.86%)	1 098 (1.10%)
Père - ISCED 4	2 860 (38.76%)	36 261 (36.37%)
Père - ISCED 5	1 223 (16.57%)	18 212 (18.27%)
Père - ISCED 6	1 766 (23.93%)	35 559 (35.67%)
<u>Niveau d'éducation de la mère</u>		
Mère - ISCED 0	82 (1.06%)	320 (0.31%)
Mère - ISCED 1	160 (2.07%)	785 (0.77%)
Mère - ISCED 2	830 (10.73%)	4 074 -4%
Mère - ISCED 3	134 (1.73%)	855 (0.84%)
Mère - ISCED 4	2 991 (38.67%)	32 694 (32.07%)
Mère - ISCED 5	1 537 (19.87%)	23 447 -23%
Mère - ISCED 6	2 001 (25.87%)	39 773 (39.01%)
Nombre d'élèves	8 139	107 276

Source : Calculs des auteurs à partir des microdonnées du PISA pour le Canada.

5 Approches méthodologiques

À l'instar de toutes les évaluations de programmes, cette étude s'est heurtée à plusieurs difficultés statistiques propres à la mesure des effets du redoublement sur la réussite scolaire. Premièrement, la section 4 a montré que les élèves qui reprennent une année ont des caractéristiques scolaires et socioéconomiques statistiquement différentes. Il est donc probable que les écarts de réussite observés entre ceux ayant vécu un redoublement et les autres ne soient pas entièrement attribuables à celui-ci, mais également à des caractéristiques et aptitudes différentes. Malheureusement, nos données ne nous permettent pas d'observer tous les déterminants de la réussite. Par exemple, les aptitudes d'apprentissage de l'élève font partie des déterminants inobservables de la réussite scolaire. Exclues des estimations économétriques, les effets estimés du redoublement pourraient s'avérer statistiquement erronés.

Deuxièmement, comme il a été mentionné auparavant, les sources de données présentent chacune des limites dans leur capacité à réduire l'incidence des facteurs confondants inobservables sur l'estimation des effets du redoublement. Ainsi, il est nécessaire d'utiliser plusieurs méthodes économétriques pour le traitement statistique des éléments inobservables et de l'hétérogénéité des effets de la reprise scolaire afin de valider la robustesse des résultats.

Les modèles de sélection de l'échantillon à un et à deux régimes sont bien adaptés à l'évaluation des effets du redoublement sur les notes scolaires pour deux raisons : 1) l'hypothèse de normalité des erreurs est particulièrement crédible dans le cas des notes scolaires; 2) le recours au troisième trimestre de naissance à titre de condition d'exclusion facilite la détermination de l'effet de traitement dans ce type de modèle. Ce renseignement sur la date de naissance des élèves facilite également l'emploi de la méthode des variables instrumentales, qui se révélera particulièrement utile dans l'exploitation des données du PISA.

Sur le plan psychologique, il est raisonnable de penser que les élèves anticipent des gains quant à l'obtention du diplôme d'études secondaires. Bien entendu, les attentes peuvent être imparfaites et c'est en partie cette évaluation différenciée des gains de l'éducation qui génère des choix différents. De plus, les différences dans les aptitudes et la facilité à apprendre modifient les gains nets de la scolarisation. En effet, pour certains, les coûts (sociaux, psychologiques ou imprévus) relatifs à la scolarité sont si élevés que le décrochage scolaire s'avère la meilleure option selon leurs critères d'optimisation. Par exemple, certains élèves quittent hâtivement leur parcours scolaire sans diplôme et avant même la reprise. Pour ces élèves, il est probable que l'effort (coût) à consentir dans le cas

d'un redoublement soit jugé excessif. Ainsi, alors que, pour plusieurs, les gains associés à l'obtention d'un diplôme d'études secondaires, escomptés à long terme, sont avantageux, pour d'autres, ces avantages sont considérés comme étant à très court terme. Dans ce cas, certains opteront pour la formation professionnelle.

Au Québec, les élèves qui n'ont pas encore obtenu un diplôme au secondaire bénéficient d'un panier de programmes éducatifs adaptés à leurs besoins. Il a été mentionné, dans les sections 2 et 4, qu'ils peuvent opter pour un programme conduisant à un diplôme autre que le DES, dont le certificat de formation à un métier semi-spécialisé (CFMS). Dans ce contexte, plusieurs choix s'offrent à eux et le modèle de probit ordonné se prête bien à l'analyse de ces choix. Il est à noter que le décrochage scolaire s'ajoute aux options disponibles et que la reprise scolaire est l'un des facteurs susceptibles de l'expliquer. Les effets du redoublement sur la diplomation peuvent donc être considérés selon des modèles à choix multiples.

Cette étude aborde la modélisation des indicateurs de la réussite scolaire et l'estimation des effets du redoublement d'un point de vue paramétrique, comme il est décrit plus haut, et selon des méthodes non paramétriques, lesquelles ont l'avantage d'éviter les biais de spécification. La méthode de l'appariement par les scores de propension (ASP) a notamment été employée. Elle a l'avantage d'assurer une comparaison efficace du groupe de traitement et du groupe témoin selon un recoupement des scores de propension de ces deux groupes restreint à un support commun. En revanche, cette méthode est tributaire de l'hypothèse d'indépendance conditionnelle et, du coup, ne tient pas compte de l'endogénéité inobservable de la variable de traitement. Le test de Rosenbaum permet toutefois de vérifier si l'endogénéité inobservable est un enjeu pour les estimations¹⁸.

Les méthodes d'apprentissage machine (AM) et, en particulier, les forêts aléatoires causales appliquées à l'évaluation des effets de traitement connaissent un essor important depuis quelques années (Athey, 2015; Linden et Yarnold, 2016; Athey et Imbens, 2019; Athey et Wager, 2019). L'AM permet d'estimer efficacement les effets du redoublement sur la réussite scolaire en raison, d'une part, de la quantité importante de données administratives disponibles pour cette étude et, d'autre part, de la flexibilité de son approche pour la comparaison de groupes (de traitement et témoin) d'élèves comparables sur les plans individuel, familial et socioéconomique. À l'instar de l'ASP, l'efficacité de l'AM dépend de l'hypothèse d'indépendance conditionnelle. Néanmoins, les plus récents développements, en particulier les forêts aléatoires causales avec variables instrumentales

¹⁸ Voir P. R. Rosenbaum (2002). *Observational Studies*. Springer Series in Statistics, 375 p.

(Chen et Hsiang, 2019), contribuent au relâchement de l'HIC et à la détermination de l'effet de traitement.

5.1 Bref rappel du problème de l'évaluation

Cette étude visait à évaluer les effets du redoublement sur quatre indicateurs de la réussite scolaire, soit les notes aux épreuves uniques, la diplomation, le taux de sortie sans diplôme ni qualification à la FGJ et les scores aux tests de PISA. Elle visait aussi à connaître les effets à court et à long terme de cette mesure en permettant des effets différenciés selon que la reprise a eu lieu au primaire ou au secondaire. La littérature citée plus haut suggère en effet que les effets de la reprise varient selon le niveau scolaire de l'élève au moment de celle-ci.

La modélisation économétrique de la réussite scolaire représente un enjeu fondamental et courant dans le secteur de l'éducation en raison des renseignements restreints sur les indicateurs de réussite et leurs déterminants (Todd et Wolpin, 2003). Comme il a été mentionné plus haut, il serait notamment utile d'avoir accès à des données précises sur l'environnement familial de l'élève ainsi que sur ses aptitudes innées ou acquises. Au regard des indicateurs de réussite, dans les données administratives du Ministère, le suivi longitudinal ne permet de mesurer que la propension d'un élève à quitter l'école sans diplôme ni qualification, puisque les données sur les résultats des élèves ne se trouvent pas dans les bases de données.

Nous décrivons ici de manière succincte le problème auquel nous faisons face (Borland, Wilkins et Tseng, 2005; Caliendo et Hujer, 2005; Heckman, Lalonde et Smith, 1999; Winship et Morgan, 1999). D'un point de vue formel, le résultat observé Y de l'élève i peut s'écrire de la façon suivante :

$$Y_i = D_i Y_i^1 + (1-D_i) Y_i^0, \quad (1)$$

où

$$D_i = \begin{cases} 1 & \text{si l'élève reprend une année scolaire} \\ 0 & \text{autrement} \end{cases}$$

et Y_i dénote la valeur de l'un des indicateurs de la réussite scolaire de l'élève i . Il peut s'agir de l'un des trois indicateurs décrits plus haut. Le résultat observé, Y_i , est soit Y_i^1 si l'élève a repris une année scolaire, soit Y_i^0 autrement.

L'équation (1) peut aussi s'écrire de la façon suivante :

$$Y_i = Y_i^0 + \delta_i D_i. \quad (2)$$

Dans cette équation, le paramètre δ_i indique la différence entre Y_i^1 et Y_i^0 , soit l'effet du redoublement sur la réussite scolaire d'un élève. Ce paramètre est précisément ce que nous cherchons à déterminer. Il est toutefois indéterminé pour un élève en particulier, puisque Y_i^1 et Y_i^0 ne peuvent être observés simultanément. Là se trouve le cœur du problème de l'évaluation. Pour pallier ce problème, il suffit de s'intéresser à l'incidence la plus probable du traitement ou à son espérance mathématique. Cette valeur attendue de l'effet de traitement n'est liée à aucun individu en particulier. Elle fait plutôt référence à l'impact de la mesure si un *très* grand nombre d'élèves redoublent une année scolaire. En pratique, la valeur attendue de l'effet du redoublement peut être estimée de manière statistique ou économétrique avec les contraintes de restriction des données décrites plus haut. Par ailleurs, il est particulièrement avisé de s'intéresser à l'effet attendu de la mesure, non pas sur un élève tiré au hasard parmi tous les effectifs scolaires, mais précisément sur les doubleurs ($D_i = 1$). En effet, ce ne sont pas tous les élèves qui vont redoubler. Déterminer l'effet du redoublement sur des élèves qui ne reprendront jamais une année n'a aucun intérêt. Ce que l'on cherche à mesurer est essentiellement l'effet du redoublement sur les élèves dont les caractéristiques les prédisposent fortement à reprendre une année étant donné leurs difficultés d'apprentissage ponctuelles ou récurrentes. Donc, l'effet attendu du redoublement scolaire sur les doubleurs peut être formalisé par l'équation suivante :

$$E(\delta_i | D_i = 1) = E(Y_i^1 | D_i = 1) - E(Y_i^0 | D_i = 1). \quad (3)$$

Dans cette équation, le terme $E(Y_i^1 | D_i = 1)$ indique la valeur attendue de Y_i^1 pour les doubleurs et $E(Y_i^0 | D_i = 1)$, leur valeur attendue si les mêmes élèves n'avaient pas repris une année scolaire. De toute évidence, cette valeur ne peut être observée. C'est ce qu'on appelle le facteur contrefactuel du redoublement scolaire, soit la réussite scolaire des doubleurs s'ils n'avaient pas repris une année. Dès lors, l'expression (3) ne peut servir, comme elle est présentée, au calcul de l'impact du redoublement. Il est toutefois possible de la remanier pour qu'elle s'avère utile à ce calcul. Il suffit de poser la condition que les termes $E(Y_i^0 | D_i = 1)$ et $E(Y_i^0 | D_i = 0)$ sont égaux, c'est-à-dire que la réussite scolaire des doubleurs et celle des non-doubleurs sont en moyenne identiques en l'absence de redoublement. Dans ce contexte, l'effet moyen de la reprise sur les doubleurs peut s'écrire comme suit :

$$\delta = E(\delta_i | D_i = 1) = E(Y_i^1 | D_i = 1) - E(Y_i^0 | D_i = 0). \quad (4)$$

Cette expression correspond tout simplement au calcul de la différence entre la réussite scolaire des doubleurs et celle des non-doubleurs. Elle offre malheureusement un estimateur vraisemblablement biaisé de l'effet du redoublement, puisque la propension à reprendre une année scolaire n'est pas un événement aléatoire et qu'il est donc peu probable que $E(Y_i^0|D_i = 1) = E(Y_i^0|D_i = 0)$. En effet, comme il a été rapporté à la section 3 et montré d'un point de vue statistique à la section 4, plusieurs facteurs socioéconomiques distinguent les doubleurs des autres élèves. Le paramètre δ de l'équation (4) « souffre » d'un biais de sélection classique.

L'hypothèse d'indépendance conditionnelle permet de pallier en partie ce problème de sélection. Elle requiert que la totalité des facteurs qui déterminent aussi bien le redoublement que la réussite scolaire soit prise en considération. Formellement, cet énoncé s'écrit comme suit :

$$Y^0 \perp D|X, \tag{5}$$

où \perp indique l'indépendance entre Y_0 et D étant donné X .

Par cette hypothèse, il est alors possible de déterminer l'effet moyen du redoublement sur les doubleurs par le calcul suivant :

$$\begin{aligned} E(\delta_i|D_i = 1, X) &= E(Y_i^1|D_i = 1, X) - E(Y_i^0|D_i = 0, X) \\ &\text{à la condition que :} \\ E(Y_i^0|D_i = 1, X) &= E(Y_i^0|D_i = 0, X) = E(Y_i^0|X), \end{aligned} \tag{6}$$

où X symbolise le vecteur de tous les déterminants de la réussite scolaire et du redoublement. Le terme $E(Y_i^0|X)$ indique que le résultat attendu en l'absence de la mesure est invariant à l'égard du statut de participation. Les régressions statistiques et économétriques par moindres carrés ordinaires, par exemple, s'appuient sur l'HIC¹⁹.

Cette hypothèse est cependant loin d'être triviale. En pratique, elle requiert l'accès à des données très riches. Elle sollicite également une réflexion minutieuse, guidée par les divers champs scientifiques concernés, à propos des facteurs qui ont un effet ou non sur la participation au redoublement *et* à la réussite scolaire²⁰. En dépit de ces considérations et étant donné que les données observationnelles s'avèrent incomplètes, un nombre

¹⁹ Il faut cependant spécifier que les moindres carrés ordinaires ne permettent pas nécessairement d'estimer l'effet moyen d'un programme ou d'une mesure sur les participants.

²⁰ La section 3 a fait état, notamment, des facteurs psychologiques qui peuvent amener un élève à reprendre une année scolaire.

indéterminé de variables dites confondantes échappent en général aux analyses statistiques et économétriques. Elles sont communément nommées « facteurs confondants inobservables ».

En raison des limites concernant les données d'enquête ou de source administrative, la reprise scolaire est, selon toute vraisemblance, une variable endogène dans l'explication de la réussite scolaire. Cette étude vise donc à neutraliser l'incidence de ce biais dans l'estimation des effets de la reprise scolaire en exploitant différentes méthodes tenant compte de l'endogénéité du redoublement. Ces méthodes sont décrites plus bas. De plus, pour la vérification de la robustesse des estimations du redoublement scolaire, cette étude a recours à une stratégie d'évaluation dans laquelle sont employées aussi bien les données administratives du MEQ (BDLEQ) que les données d'enquête du PISA. L'utilisation de la BDLEQ permet d'exploiter les avantages d'un très grand nombre d'observations sur plusieurs cohortes d'élèves. Les données d'enquête du PISA offrent, quant à elles, la possibilité de tenir compte d'un nombre plus conséquent de déterminants socioéconomiques de la réussite scolaire ou, à tout le moins, d'une mesure plus précise de ces déterminants.

5.2 Méthodes paramétriques de modélisation des indicateurs de la réussite scolaire : description sommaire

5.2.1 Modélisation des notes scolaires

Les notes scolaires colligées par le MEQ sont disponibles pour la quatrième et la cinquième secondaire uniquement en ce qui concerne les épreuves menant à un diplôme à la FGJ. Ces données ne permettent donc pas un suivi longitudinal des résultats d'un élève. Elles ne sont pas disponibles pour l'ensemble des élèves, mais uniquement pour ceux qui se rendent jusqu'à la quatrième ou à la cinquième secondaire en FGJ. Tous les jeunes qui quittent la FGJ avant la tenue de ces examens sont donc exclus des analyses. Étant donné la nature transversale de ces notes, leur modélisation aux épreuves uniques doit reposer sur les approches économétriques, également en coupes transversales.

Nous commençons donc l'analyse par l'utilisation d'un modèle classique de sélection de l'échantillon avec changement de régime pour évaluer les effets du redoublement sur les notes scolaires. Cette approche instaure un cadre d'analyse bien adapté à la modélisation combinée du redoublement et des notes. Dans ce modèle, la propension d'un élève à reprendre une année scolaire est symbolisée par une variable latente, $D_i^* = Y_i^1 - Y_i^0 - C_i$, où Y_i^1 est la note d'un élève i qui redouble une fois ou plus, Y_i^0 est la note à l'épreuve si l'élève ne redouble pas et C_i , le coût anticipé du redoublement. L'élève reprend une année

si la différence entre les gains escomptés du redoublement ($Y_i^1 - Y_i^0$) et son coût est nulle ou positive²¹. Seul le choix de l'élève ou celui des intervenants concernés (enseignant, parent, direction d'école) est observé et se décrit par la fonction indicatrice,

$$D(Z)_i = 1[D_i^*(Z) \geq 0] = 1[Z_i\theta + U_i^D \geq 0], \quad (7)$$

où $Z_i\theta + U_i^D$ modélise de façon linéaire la décision dichotomique de l'élève (D). Le terme Z_i est un vecteur de variables observées (ex. : revenu familial du secteur de résidence, scolarité des parents du secteur de résidence, trimestre de naissance du jeune) et le terme U_i^D , un vecteur de facteurs inobservés (ex. : aptitude cognitive de l'élève, troubles familiaux ponctuels ou récurrents, relations avec les camarades de classe). Ces deux termes déterminent la propension de l'élève au redoublement. Les espérances conditionnelles de Y_i^1 et de Y_i^0 sont décrites respectivement par $X'\beta^1$ et $X'\beta^0$, où X est un vecteur de facteurs explicatifs de la réussite scolaire. Le modèle classique avec changement de régime s'écrit donc de la manière suivante :

$$Y_i^1 = X_i'\beta^1 + U_i^1 \quad (8)$$

$$Y_i^0 = X_i'\beta^0 + U_i^0, \quad (9)$$

où U_i^1 et U_i^0 sont des termes d'erreur à moyenne nulle et indépendants de X . L'équation 8 décrit les résultats scolaires des doubleurs, tandis que l'équation 9 montre les résultats des non-doubleurs.

Les erreurs de chaque équation sont censées suivre une loi normale trivariée, ce qui est relativement raisonnable compte tenu de la normalité observée dans la distribution des notes scolaires :

$$\begin{bmatrix} U^D \\ U^1 \\ U^0 \end{bmatrix} \sim N \left(0, \begin{bmatrix} \sigma_{U^D}^2 & \sigma_{U^1 U^D} & \sigma_{U^0 U^D} \\ \sigma_{U^1 U^D} & \sigma_{U^1}^2 & \sigma_{U^1 U^0} \\ \sigma_{U^0 U^D} & \sigma_{U^1 U^0} & \sigma_{U^0}^2 \end{bmatrix} \right), \quad (10)$$

²¹ Certains peuvent s'opposer à ce modèle, notamment pour les élèves du primaire, puisque l'élève qui échoue aux tests scolaires doit nécessairement prolonger son cycle d'apprentissage ou reprendre l'année où il a échoué. La reprise scolaire n'est pas issue d'un processus décisionnel idiosyncratique, c'est-à-dire qui dépend de la situation précise de l'élève. La mesure est plutôt jugée profitable pour tous les élèves en situation d'échec, peu importe son coût. Au secondaire, bien que l'obligation de reprendre une année soit aussi dictée par une mesure universelle jugée bénéfique, l'élève, compte tenu de son âge et de sa maturité, peut décider lui-même de reprendre ou non l'année qu'il n'a pas réussie. Dans le cas d'un refus de redoubler, il peut, suivant sa décision, quitter l'école sans diplôme ni qualification.

où σ_{UD}^2 est la variance du terme d'erreur dans l'équation du redoublement scolaire (7) et $\sigma_{U^1}^2$ et $\sigma_{U^0}^2$ sont respectivement les variances des erreurs des équations 8 et 9. Le paramètre $\sigma_{U^1U^0}$ indique la covariance entre U^1 et U^0 et σ_{U^0UD} , la covariance entre U^0 et U^D .

L'estimation des équations 7 à 9 peut être abordée selon deux approches. La première approche consiste à estimer le modèle en deux étapes, soit en estimant de manière asymptotiquement convergente les paramètres du modèle (Maddala, 1983; Heckman, Tobias et Vytlačil, 2001). La deuxième approche, celle retenue dans cette étude, vise à maximiser la fonction de vraisemblance du système des équations 7 à 9 sous les conditions de normalité jointe des termes d'erreur (Maddala, 1983; Lokshin et Sajaia, 2004). Le modèle est identifiable par non-linéarité. La détermination du modèle sera facilitée toutefois par la disponibilité d'une ou de plusieurs variables instrumentales. À l'instar d'autres recherches, dans cette étude, le trimestre de naissance de l'élève est utilisé à titre d'instrument du redoublement scolaire (Gonzalez-Betancour et Lopez-Puig, 2016). La procédure *MoveStay* du logiciel Stata, développée par Lokshin et Sajaia, est employée afin de maximiser la fonction de vraisemblance du modèle à deux régimes.

5.2.2 Modélisation combinée de la diplomation et de la non-diplomation (formation générale des adultes, décrochage scolaire) : un modèle de probit ordonné avec endogénéité de la variable de traitement

Au cours de son cheminement scolaire, l'élève doit faire plusieurs choix en lien avec son parcours éducatif. Ainsi, il doit choisir le ou les diplômes qu'il souhaite obtenir. L'élève pourra, par exemple, choisir parmi les trois options suivantes : obtenir un DES en sept ans, obtenir un diplôme autre qu'un DES en sept ans ou terminer son parcours scolaire sans diplôme ni qualification. Pour faire ce choix, l'élève doit tenir compte de ses préférences quant à l'emploi qu'il souhaite occuper une fois adulte. Par exemple, envisage-t-il une profession ou un métier, ou préfère-t-il un emploi demandant un faible niveau de qualification? La réponse à cette question influence ses choix éducatifs. De plus, la décision de l'élève vis-à-vis de sa formation immédiate et future sera en partie guidée par les gains anticipés concernant les différents emplois considérés. Ces gains sont variés. Ils peuvent être économiques, c'est-à-dire axés sur le revenu et les avantages sociaux. Ils peuvent aussi être environnementaux et être liés à la qualité de milieu de vie offerte par un emploi. Enfin, ils peuvent être psychologiques et avoir trait, par exemple, au bien-être envisagé par l'occupation d'un emploi particulier.

Dans cette étude, compte tenu des renseignements statistiques rapportés à la section 4, l'élève a donc trois options au regard de ses choix scolaires : obtenir un DES en sept ans,

obtenir un diplôme autre qu'un DES en sept ans ou terminer son parcours scolaire sans diplôme ni qualification. Le modèle de probit ordonné sert à modéliser le choix de l'élève par rapport à ces trois options. Ce modèle a été largement utilisé dans le secteur de l'éducation pour expliquer, entre autres choses, le choix des élèves au regard des études postsecondaires (Nguyen et Taylor, 2003) ou le rôle des préférences parentales dans leurs décisions scolaires (Bellani et Ortiz-Gervasi, 2021).

Le probit ordonné comme modèle explicatif du choix de scolarisation peut paraître a priori injustifié, puisqu'aucun ordre hiérarchique ne guide en apparence la décision de l'élève. Dans les faits, les options qui s'offrent à lui peuvent être organisées en fonction d'un ordre implicite des rendements *anticipés* des diplômes et des préférences par rapport aux gains socioéconomiques à court terme et à long terme. La probabilité de redoubler une année est l'une des variables susceptibles de déterminer le choix de scolarisation de l'élève. S'il anticipe le fait d'échouer et d'avoir à reprendre une année, il est possible qu'il oriente son choix selon ses préférences. Le redoublement est présumé endogène dans les décisions latentes de l'élève au regard de ses options scolaires.

Dans un modèle ordonné, la variable de résultat Y_i est ordinale et indique le choix de l'élève i au regard des trois options citées plus haut. Dans ce contexte de choix à trois options, la variable prendra par convention la valeur 1, 2 ou 3. Dans ce modèle, on inclut également la variable inobservable Y_i^* . Cette variable représente les gains socioéconomiques anticipés par l'élève au regard de son choix de formation et de son occupation sur le marché du travail. Enfin, ce modèle repose aussi sur X_i , un vecteur regroupant les facteurs explicatifs du choix de formation de l'élève. Le modèle de probit ordonné s'écrit de la manière suivante :

$$\begin{aligned} Y_i^* &= X_i\beta + \delta_i D_i + u_i \\ u_i &\sim N(0, 1). \end{aligned} \tag{11}$$

Ainsi, ce modèle s'appuie sur l'hypothèse selon laquelle Y_i^* dépend de X_i de façon linéaire. Dans ce modèle, β est un vecteur de paramètres sans constante et δ_i capte l'effet du redoublement scolaire. Puisque la variable Y_i^* est inobservable, sa relation avec la variable aléatoire observée Y_i est la suivante :

$$\begin{aligned} Y_i &= 1 \text{ si } \gamma_2 < Y_i^* \leq +\infty \text{ (gains anticipés élevés } \rightarrow \text{ DES)} \\ Y_i &= 2 \text{ si } \gamma_1 < Y_i^* \leq \gamma_2 \text{ (gains anticipés modérés } \rightarrow \text{ FP, CFMS)} \\ Y_i &= 3 \text{ si } -\infty < Y_i^* \leq \gamma_1 \text{ (gains anticipés nuls ou négatifs } \rightarrow \text{ abandon scolaire).} \end{aligned} \tag{12}$$

Les paramètres γ_1 et γ_2 fixent le domaine de Y^* pour trois options associées aux gains anticipés. Le modèle est estimé par un maximum de vraisemblance. La corrélation présumée entre D_i et u_i nécessite de tenir compte d'un biais d'endogénéité dans les effets estimés du redoublement. Tous les détails de l'élaboration et de l'estimation de cette approche sont présentés dans la documentation relative à la commande « eoprobit » du logiciel Stata 15.

5.3 Méthodes non paramétriques de modélisation des indicateurs de la réussite scolaire

5.3.1 Appariement par les scores de propension

La méthode de l'appariement par les scores de propension est utilisée dans l'évaluation des effets des programmes publics depuis de nombreuses années (Dehejia et Wahba, 2002; Heckman, Ichimura et Todd, 1998; Hirano, Imbens et Ridder, 2003). Sous certaines conditions, il est possible de considérer que le redoublement scolaire présente les caractéristiques d'une expérience aléatoire classique. Ainsi, cette méthode permet de corriger les données de sorte que les effets estimés du programme soient exempts de biais statistiques. En quelques mots, la correction consiste essentiellement à déterminer un groupe témoin à partir duquel il est possible de comparer les doubleurs. On cherche ici à trouver des élèves n'ayant pas redoublé, mais possédant les mêmes caractéristiques que les élèves ayant repris une année. Ces élèves du groupe témoin servent à définir la situation contrefactuelle qui est autrement inobservable.

Cette méthode présente comme principal avantage de ne pas recourir à des formes entièrement paramétriques pour les facteurs explicatifs sur les facteurs explicatifs et les termes d'erreur par rapport aux méthodes de régression standards. Elle constitue donc une méthode dite semi-paramétrique. Cette approche repose toutefois sur deux hypothèses fondamentales assez restrictives et peu triviales. La première est l'HIC, décrite à la section 5.1. La seconde stipule qu'il doit exister un groupe de comparaison. Cette condition s'écrit comme suit :

$$0 < P(D = 1|X) < 1. \tag{13}$$

Cette hypothèse garantit que, pour chaque doubleur aux caractéristiques X , il existe au moins un élève qui lui est comparable. Autrement dit, elle confirme que chaque élève a une

probabilité non nulle ou non unitaire de reprendre une année. L'autre hypothèse est la suivante :

$$Y^0 \perp D|X \Rightarrow Y^0 \perp P(X). \quad (14)$$

Cette propriété permet d'associer deux élèves, un doubleur et un non-doubleur, sur la base d'un score de propension au redoublement, lequel forme un indicateur unidimensionnel du vecteur X de facteurs explicatifs. Comme le score est une variable continue, il existe plusieurs concepts d'appariement des élèves avec ou sans redoublement. Dans cette étude, la méthode de l'appariement des plus proches voisins est privilégiée. Cette méthode consiste à sélectionner un nombre fixe de non-doubleurs dont les écarts de score par rapport au doubleur sont inférieurs à une valeur prédéterminée.

5.3.2 Apprentissage machine : forêt aléatoire causale

L'apprentissage machine est un sous-champ de l'intelligence artificielle et est constitué d'algorithmes (règles de décision) qui apprennent à détecter les motifs (*patterns*) qui mènent à des événements tels que le redoublement scolaire. De façon générale, les données dont dispose l'analyste sont divisées en deux sous-ensembles. Le premier, l'*ensemble d'apprentissages*, regroupe de 60 % à 70 % des observations et est utilisé pour « enseigner » à la machine à reconnaître les motifs (apprentissage supervisé). Le second, l'*ensemble de tests*, est utilisé pour vérifier la validité des prédictions effectuées à partir des motifs déterminés à l'aide de l'ensemble d'apprentissages. Les algorithmes reposent tous sur la construction d'« arbres de décision » (ou « arbres de prédiction »). L'algorithme nommé « forêt aléatoire » a la particularité de générer un nombre arbitrairement grand d'arbres de décision. Pour y parvenir, il extrait des sous-échantillons de l'ensemble de tests et sélectionne aléatoirement un sous-ensemble des variables utilisées aux fins de modélisation de chaque arbre. Cela permet de prendre en compte de façon très flexible les nombreuses interactions pouvant exister entre les variables. De plus, cette façon de procéder permet de déterminer les variables ayant le plus grand pouvoir prédictif de la variable d'intérêt (redoublement scolaire).

L'apprentissage machine a pour but de créer un modèle qui prédit la valeur d'une variable endogène (redoublement) à partir de plusieurs variables exogènes (caractéristiques de l'élève, environnement social, etc.). Une de ces variables est sélectionnée à chaque nœud intérieur de l'arbre selon une méthode qui dépend de l'algorithme utilisé. Chaque feuille (ou nœud terminal de l'arbre) représente soit une valeur de la variable

endogène (résultat à une épreuve unique), soit une distribution de probabilité des diverses valeurs possibles de celle-ci (décrochage, type de diplôme obtenu, etc.). La combinaison des valeurs des variables exogènes est représentée par le chemin de la racine jusqu'aux feuilles de l'arbre. Celui-ci est en général construit en divisant l'ensemble de données en sous-ensembles (nœuds) en fonction de la valeur d'une variable exogène (nœud). De manière récursive, ce processus est répété à chaque nœud. La récursion se termine lorsque la séparation n'améliore plus la prédiction. Pour choisir la variable de séparation (nœud), les algorithmes testent toutes les variables une à une et sélectionnent celle qui optimise l'indice d'impureté de Gini ou l'indice d'entropie.

L'indice d'impureté de Gini permet de mesurer la probabilité qu'une observation choisie au hasard soit mal classée dans un nœud. Cet indice est calculé comme suit :

$$\text{Gini} = 1 - \sum_{i=1}^k p_i^2, \quad (15)$$

où p_i est la probabilité de redoublement. L'indice de Gini pour un nœud quelconque est limité à une valeur de 0 à 0,5. Par exemple, si l'on s'intéresse au redoublement, la variable endogène peut prendre deux valeurs « k » : redoubler ou non. Dans ce cas, l'indice de Gini est donné par :

$$\text{Gini} = 1 - p_1^2 - p_2^2, \quad (16)$$

où p_1 est la probabilité de redoubler et p_2 , celle de ne pas redoubler. Si la probabilité est de 50 % dans les deux cas, alors l'indice de Gini est égal à 0,5. Si la proportion de redoubleurs est de 100 %, alors l'indice vaut 0. À chaque nœud, l'algorithme effectue un test pour vérifier, sur la base de cet indice, si un découpage supplémentaire à partir de chacune des variables améliore suffisamment l'indice pour justifier la création de nœuds supplémentaires.

L'indice d'entropie est une autre mesure du « désordre » dans un nœud et est fondé sur la théorie de l'information. Il est semblable à l'indice de Gini, mais limité à une valeur de 0 à 1 et demande légèrement plus de temps pour le calcul. Il est défini comme suit :

$$\text{Entropie} = - \sum_{i=1}^k p_i \log_2 p_i. \quad (17)$$

Les forêts aléatoires sont liées aux méthodes d'appariement par noyaux ainsi qu'aux méthodes du plus proche voisin en ce qu'elles font des prédictions à partir d'une moyenne

pondérée des observations «à proximité». En effet, comme les comparaisons sont effectuées à l'intérieur des feuilles de chaque arbre et que ces feuilles regroupent des observations identiques du point de vue des variables exogènes, cela revient à apparier les individus de façon totalement non paramétrique. Les forêts diffèrent toutefois en ce qu'elles sont entièrement basées sur les données pour déterminer quelles observations reçoivent plus de poids, ce qui est particulièrement important en présence d'interactions complexes entre les variables.

L'algorithme « forêt aléatoire causale », développé par Athey et Imbens (2019), consiste à estimer l'effet causal du redoublement sur les indicateurs de performance à l'intérieur de chaque feuille de chaque arbre de la forêt. Il s'agit simplement de faire régresser les indicateurs de performance concernant le redoublement scolaire sans aucune autre variable explicative à l'intérieur de chaque feuille de chaque arbre. L'inclusion d'autres variables explicatives est en effet inutile, puisque les individus appartenant à une même feuille sont identiques. On dispose alors d'autant d'estimateurs de l'effet causal qu'on trouve d'arbres dans la forêt. Cela permet d'estimer la distribution des effets causaux et non seulement leur moyenne. Par ailleurs, il est possible de relier les effets causaux estimés à l'aide de la forêt aléatoire aux variables considérées par l'algorithme pour déterminer lesquelles sont responsables de l'hétérogénéité des effets de traitement. L'estimation de cette hétérogénéité est certainement un des principaux avantages de cette approche. Enfin, il est également possible d'exploiter des variables instrumentales dans le contexte des forêts aléatoires causales (voir Athey and Wager, 2019).

6 Résultats

Jusqu'à présent, nous avons simplement décrit la relation empirique entre le redoublement scolaire et les caractéristiques des élèves (section 4). Dans les pages qui suivent, les méthodes d'estimation présentées dans la section 5 permettront d'établir une relation de causalité entre le redoublement et la réussite des jeunes. Nous débutons par une analyse des déterminants du redoublement scolaire (section 6.1). Ensuite, nous abordons les effets du redoublement sur les indicateurs de la réussite scolaire (section 6.2). Nous concluons avec les effets du redoublement sur les résultats du PISA (section 6.3).

6.1 Déterminants de la propension au redoublement scolaire

Cette sous-section contient des estimations permettant de mieux comprendre les déterminants du redoublement scolaire. Le tableau 20 présente les facteurs présumés

explicatifs²² de la propension au redoublement scolaire. Le lien entre chacun de ces facteurs et le redoublement est estimé à partir d'un modèle de probit. L'ensemble des élèves inscrits à la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007 sont inclus dans nos estimations. La deuxième colonne de ce tableau affiche les coefficients estimés du modèle, alors que la troisième indique la moyenne des effets marginaux des facteurs explicatifs. La dernière colonne montre la statistique *t-student*.

Sur le plan individuel, les résultats présentés dans le tableau 20 signifient d'abord que les filles ont une probabilité de 12,0 points de pourcentage (pp) inférieure à celle des garçons de redoubler une année, toutes choses égales d'ailleurs. Un élève né à l'étranger voit son risque de redoublement réduit de 3,5 pp, alors qu'un enfant né au troisième trimestre de l'année a une probabilité de redoublement plus élevée de 4,1 pp comparativement aux autres élèves. Les sections 3 et 4 ont souligné l'importance de l'âge relatif de l'élève dans la reprise d'une année scolaire et le fait que les plus jeunes d'une cohorte avaient une propension plus importante que les autres à développer des difficultés d'apprentissage et à reprendre une année scolaire. Les résultats du tableau 20 montrent sans équivoque l'impact de l'âge sur le redoublement scolaire. L'effet consigné suggère que, parmi une cohorte de 70 000 élèves qui commencent leur première année du primaire à une année donnée, 700 d'entre eux seront des doubleurs uniquement parce qu'ils sont les plus jeunes parmi leurs camarades de classe ou l'ensemble des élèves de l'établissement scolaire fréquenté²³.

De plus, un environnement familial sain et harmonieux est nécessaire au bon cheminement éducatif des enfants tout comme les ressources éducatives mises à la disposition de la famille. Une hausse de 10 000 \$ du revenu familial du secteur de résidence est associée à une baisse de 1,0 pp de la propension de l'élève à reprendre une année scolaire. Cette propension augmente en revanche si la mère est seule responsable de la garde de ses enfants (+4,0 pp) ou si le père et/ou la mère sont sans diplôme ni qualification (+2,7 pp).

²² Plusieurs variables de cette étude n'ont pas nécessairement un lien de cause à effet fondamental pour la propension au redoublement ou la réussite scolaire des élèves. Autrement dit, l'effet estimé d'un facteur peut découler davantage de déterminants corrélés à celui-ci. Par exemple, la participation à un projet éducatif agréé peut indiquer davantage l'aptitude scolaire d'un élève que l'effet du projet lui-même. Par ailleurs, le changement d'établissement durant le parcours scolaire peut être associé à l'anxiété qu'éprouve l'élève et qui est causée directement par ce changement, mais pourrait aussi indiquer un effet anxiogène dû à une séparation familiale.

²³ Il a été supposé que le nombre d'élèves par trimestre de naissance est identique et que l'effet estimé du troisième trimestre de naissance sur la propension au redoublement est aussi le même pour tous les élèves de ce trimestre. Par conséquent, la division par 4 d'un nombre de 70 000 élèves suivie de la multiplication du résultat (17 500) par 0,041 est égale à 700.

Cette étude visait à tenir compte des traits cognitifs et non cognitifs des élèves par l'entremise de leur fréquentation de programmes préscolaires et scolaires²⁴. Ainsi, les données du tableau 20 montrent qu'une participation au programme Passe-Partout réduit de 2,3 pp la probabilité de redoubler une année, tandis que la fréquentation de la maternelle 4 ans à demi-temps ou les facteurs corrélés avec cette fréquentation haussent cette propension de 5,4 pp. De plus, la fréquentation d'une école privée au secondaire est associée à une baisse importante de la probabilité de reprise scolaire (-18,3 pp) tout comme la participation à un projet pédagogique particulier (-21,4 pp)²⁵. Par ailleurs, le changement d'établissement scolaire, à l'exception de celui qui découle du passage de la dernière année du primaire au premier cycle du secondaire, ne favorise pas le cheminement éducatif régulier de l'élève. En outre, ce changement est ressenti de façon importante, puisqu'il est lié à une hausse du risque de redoublement de 9,5 pp au primaire et de 19,9 pp au secondaire.

L'environnement scolaire joue également un rôle important au regard du cheminement scolaire et de la réussite éducative des élèves. Le tableau 20 présente plusieurs indicateurs du réseau scolaire qui peuvent contribuer à cette réussite. Le statut d'emploi des enseignants titulaires apparaît favorable au cheminement éducatif, puisque chaque hausse de 1,0 pp pour les enseignants titulaires à temps plein dans les établissements scolaires au primaire réduit de 17,1 pp le risque de redoublement. Les enseignants comptant le plus d'années d'ancienneté contribuent aussi à une réduction du risque. Trois indicateurs du statut socioéconomique de l'école ont été insérés dans les analyses. Ainsi, il est estimé qu'une hausse de l'IMSE, des revenus familiaux moins élevés ou un pourcentage plus important d'élèves nés à l'étranger parmi les jeunes fréquentant l'établissement sont associés à une augmentation de la probabilité de redoublement scolaire chez un élève. Celle-ci diminue pour les écoles avec des effectifs plus importants. Le tableau 20 indique également que l'année de construction d'un immeuble scolaire, bien que son effet soit négligeable, est associée à la propension d'un élève à reprendre une année scolaire. Ainsi, le jeune qui fréquente un établissement situé dans un immeuble de construction relativement récente par rapport aux autres établissements d'enseignement au Québec risque moins de redoubler. Bien entendu, ce coefficient ne permet pas de conclure que l'année de construction de l'immeuble cause un retard scolaire. Toutefois, cette variable

²⁴ Il est présumé, par exemple, qu'un élève qui participe à un projet pédagogique particulier qui est agréé par le ministère de l'Éducation possède les aptitudes nécessaires pour répondre aussi bien aux exigences scolaires qu'aux exigences sportives dans le contexte d'une concentration des heures allouées à l'apprentissage scolaire.

²⁵ Ces facteurs sont vraisemblablement corrélés avec les aptitudes de l'élève ou les autres facteurs inobservables dans cette étude et, du coup, entachés de biais d'endogénéité.

capte possiblement, au regard des écoles, des effets inobservables qui affectent le rendement scolaire (désuétude des installations sportives, bibliothèque, etc.).

Finalement, précisons que ce modèle permet d'estimer la probabilité de redoubler selon les caractéristiques de l'élève et de son milieu. L'inclusion du troisième trimestre de naissance permet de générer une variation exogène par rapport à la réussite scolaire dans la probabilité de redoublement. En effet, il n'y a aucune raison de penser que les élèves nés à un trimestre donné sont différents sur le plan socioéconomique, démographique ou génétique de ceux nés aux autres trimestres. Cette variable instrumentale sera mise à profit dans les estimations économétriques pour contribuer à la détermination des effets du redoublement scolaire.

Tableau 20 — Résultats de l'estimation de la propension à reprendre une année scolaire par le modèle de probit, cohorte de la première année du primaire en 2006-2007

Variables	Coefficient	Moyennes des effets marginaux individuels	t-stat
	(2)	(3)	(4)
Caractéristiques personnelles			
Fille	-0,422	-0,118	-73,916
Né à l'étranger	-0,130	-0,035	-9,137
Né au troisième trimestre de l'année de naissance	0,148	0,041	23,020
Environnement familial			
Mère et/ou père sans diplôme (P1 à S5)	0,098	0,027	15,382
Revenu familial moyen aire de diffusion (P1 à S5)	-0,037	-0,010	-25,722
Mère monoparentale (P1 à S5)	0,140	0,040	8,524
Programmes et parcours scolaire			
Programme passe-partout (4 ans)	-0,083	-0,023	-9,324
Maternelle 4 ans demi-temps	0,191	0,054	15,904
Projet éducatif particulier (S1 à S5)	-0,824	-0,214	-111,746
Fréquentation école privée (S1 à S5)	-0,734	-0,186	-76,095
Changement d'école (P1 à P6)	0,339	0,095	58,322
Changement d'école (S1 à S5)	0,646	0,190	110,081
Environnement scolaire			
Taux enseignant titulaire temps plein (P1 à P6)	-0,620	-0,171	-17,868
IMSE (P1 à S5)	0,006	0,002	8,379
% né à l'étranger (P1 à P6)	0,333	0,092	6,342
% né à l'étranger (S1 à S5)	0,403	0,111	9,420
Année de construction immeuble (P1 à S5)	-0,001	-0,000	-3,987
Année d'ancienneté enseignants titulaires (P1 à P6)	-0,011	-0,003	-11,405
Revenu familial moyen école (P1 à P6)	-0,003	-0,001	-1,196
Revenu familial moyen école (S1 à S5)	-0,071	-0,020	-32,286
705 <= Effectifs école < 1016	-0,108	-0,029	-13,151
1016 <= Effectifs école < 1360	-0,102	-0,028	-12,154
1360 > Effectifs école	-0,016	-0,004	-1,723
Constante	2,886		6,010
Nombre d'observations	264 380		
Pseudo-R2	0,207		

6.2 Effets du redoublement scolaire sur les indicateurs de la réussite scolaire : données administratives

6.2.1 Notes aux épreuves uniques

Nous abordons maintenant les effets du redoublement sur les notes relatives à cinq épreuves uniques auxquelles nous avons eu accès. Il s'agit en premier lieu des trois épreuves de quatrième secondaire, à savoir celles de mathématiques (séquence « Culture, société et technique » [CST]), d'histoire et de science, ainsi qu'en deuxième lieu des deux épreuves de cinquième secondaire, soit celles de français écrit et d'anglais, langue seconde. Les tableaux 21 à 24 présentent les effets estimés du redoublement scolaire sur les notes à ces épreuves pour la cohorte des élèves inscrits à la première année du primaire en 2006-2007. Pour ces estimations, nous nous concentrons uniquement sur une cohorte afin d'évaluer les effets de la reprise selon deux groupes de comparaison, le premier étant les camarades de la cohorte d'entrée au primaire du doubleur et le second, ses nouveaux compagnons depuis le redoublement. Les régressions, avec l'emploi du second groupe de comparaison, portent sur le même examen effectué en 2016-2017 pour les épreuves de quatrième secondaire et en 2017-2018 pour les épreuves de cinquième secondaire. Ainsi, seuls les élèves ayant redoublé une fois composent le groupe de traitement.

Dans les tableaux 21 à 24, les effets estimés du redoublement sur les notes aux épreuves uniques sont ventilés selon les méthodes d'estimation employées et le moment du ou des redoublements²⁶. Cette ventilation vise à vérifier si la mesure appliquée tôt ou tard dans le parcours scolaire de l'élève comporte des avantages ou des inconvénients particuliers. Ainsi, les tableaux se divisent en trois sections : la section A présente les effets du redoublement, qu'il ait eu lieu au primaire ou au secondaire, la section B montre ses effets au secondaire, alors que la section C rapporte les effets qu'il a produits au primaire. Ainsi, la variable de résultat prend la valeur 1 si l'étudiant a redoublé au moins une fois au primaire et/ou au secondaire (section A), au secondaire (section B) ou au primaire (section C), sinon elle prend la valeur 0. La première ligne de ces tableaux indique les écarts de résultats entre les élèves ayant redoublé et les autres, sans variable de contrôle ni modèle. Nous ne parlerons donc pas des effets du redoublement dans ce cas. La deuxième ligne montre les effets du redoublement estimés par la méthode des moindres carrés ordinaires (MCO). Ces estimations offrent une première indication des effets du redoublement en tenant compte de la sélection selon les facteurs observables. Les facteurs

²⁶ Les résultats détaillés des régressions économétriques sont présentés à l'annexe A. Par contre, par souci de concision, seulement un nombre restreint de résultats sont annexés à cette étude. Tous les résultats détaillés sont disponibles sur demande.

inobservables demeurent possiblement un problème à cette étape. La troisième ligne montre les effets estimés du redoublement obtenus par la méthode de l'appariement par les scores de propension selon les critères de sélection « plus proche voisin » des unités de comparaison. La quatrième ligne comporte les estimations produites par les forêts aléatoires causales. Les effets estimés par cette méthode touchent exceptionnellement les élèves de plusieurs cohortes, soit celles de la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007. Ces trois méthodes, soit les MCO, l'appariement par les scores de propension et les forêts aléatoires causales, s'appuient toutes sur l'HIC dans le calcul des effets du redoublement scolaire. Quant à la dernière ligne du tableau, elle concerne les estimations produites par les modèles classiques de sélection de l'échantillon à un régime par un maximum de vraisemblance²⁷. Dans ces modèles, la sélection selon les facteurs inobservables est prise en considération par le paramètre de corrélation des termes d'erreur entre les équations de participation et les notes scolaires.

Avant de poursuivre l'analyse des résultats du tableau 21, il est d'abord utile de s'intéresser aux coefficients des facteurs explicatifs autres que le redoublement pour les notes aux épreuves uniques afin de déterminer si les écarts entre les doubleurs et les non-doubleurs peuvent s'expliquer en partie par ces autres facteurs. L'analyse a été réalisée pour l'épreuve de français écrit de la cinquième année du secondaire et les résultats sont présentés au tableau A.1 de l'annexe A. Les chiffres du tableau indiquent que les filles obtiennent une note en moyenne supérieure à celle des garçons pour un écart estimé à près de 5 pp. Être né à l'étranger n'a aucune incidence sur la note à l'épreuve. En revanche, tout handicap physique ou intellectuel d'un élève hausse de 2 pp les résultats toutes choses égales d'ailleurs. Il est possible que l'effet de ce facteur soit fortement associé au soutien pédagogique prévu pour les élèves touchés par un handicap et que d'autres variables présentes dans le modèle captent l'effet d'un handicap sur les notes aux épreuves uniques. Il est aussi important de se rappeler que le sous-échantillon d'élèves ayant effectué cette épreuve n'est pas entièrement représentatif de l'ensemble des jeunes de la cohorte de 2006-2007 puisque, par exemple, 78,5 % de cette cohorte, soit 53 738 élèves, finiront par participer à l'épreuve de français écrit. Parmi eux, seuls 979 jeunes vivant avec un handicap physique ou intellectuel sur un total de 3 183 y prendront part. Ainsi, les élèves

²⁷ La procédure *MoveStay* offerte par le logiciel Stata et utilisée pour maximiser la fonction de vraisemblance du système à trois équations des modèles à deux régimes n'a pas produit d'estimations convergentes. Par conséquent, les modèles classiques de sélection de l'échantillon à deux régimes ont été délaissés au profit de ceux à un régime.

lourdement handicapés sont peu nombreux à participer à cette épreuve, tout comme les élèves accusant des retards importants de développement.

Au Québec, l'accès et le soutien à l'instruction pour une majorité d'élèves sont financés par les deniers publics. En principe, la réussite scolaire est théoriquement dissociée de la richesse parentale. Or, les estimations montrent que le revenu familial du secteur de résidence de l'élève a un lien positif avec la note à l'épreuve de français écrit. Ainsi, chaque augmentation de 10 000 \$ de ce revenu est associée à une hausse de 0,1 pp de la note obtenue. Dans ce contexte, il est présumé que le revenu des parents s'avère un indicateur d'aisance²⁸ du milieu de vie de l'élève facilitant son cheminement et sa réussite scolaires. La structure familiale probable de l'élève a également une incidence sur sa réussite. La monoparentalité de la mère est associée à une baisse de 1,2 pp de la note à l'épreuve²⁹. La sous-scolarisation des parents est également associée à des notes plus faibles. En effet, un élève ayant une mère et/ou un père sans diplôme ni qualification obtient une note plus faible de 0,4 pp en français écrit³⁰.

De plus, durant son parcours scolaire, l'élève, en accord avec ses parents, peut participer à des programmes de soutien à la réussite éducative ou à des projets parascolaires ou encore faire le choix de fréquenter une école privée. Les régressions économétriques tiennent compte de ces choix non pas pour une évaluation directe de leur incidence sur la réussite, mais plutôt pour la prise en considération des facteurs propres à l'élève ou à son environnement familial corrélés avec ces choix. Ainsi, un élève qui participe à un projet éducatif particulier conventionné pourrait posséder également de grandes aptitudes pour l'apprentissage scolaire qui demeurent largement inobservables dans les données, mais qui peuvent être captées indirectement par ces autres variables. De la même façon, pour les enfants de cette cohorte, seuls les tout-petits des milieux défavorisés de l'île de Montréal ont pu fréquenter le programme de la maternelle 4 ans à demi-temps. Enfin, la participation au programme Passe-Partout peut aussi indiquer l'intérêt et la motivation du parent pour un soutien au développement de l'enfant.

Les résultats rapportés dans le tableau A.1 montrent que la fréquentation d'une école privée est associée à une hausse de 2,9 pp de la note à l'épreuve de français écrit tout comme la participation à un projet éducatif particulier. En revanche, prendre part au programme de la maternelle 4 ans à demi-temps est associé à une baisse de cette note. Des

²⁸ Cet indicateur d'aisance peut être notamment lié au degré de préoccupations économiques de la famille et, par conséquent, au degré de tension familiale pouvant être associé à ces préoccupations.

²⁹ Ici, il ne s'agit pas d'un effet délibéré du parent qui fait en sorte que son enfant subisse les contrecoups de la garde monoparentale, mais plutôt des conséquences indirectes de cette garde sur la réussite de l'élève.

³⁰ Rappelons que les indicateurs dichotomiques bâtis à partir des données sur le recensement signalent si, oui ou non, une majorité (66 %) de parents sont sans diplôme ni qualification dans une aire de diffusion donnée.

milieux de vie scolaire et familial affermis et stables, les amis et la camaraderie peuvent aussi contribuer à soutenir la réussite scolaire. Les résultats indiquent que le changement d'établissement scolaire, à l'exception de celui causé par le passage du primaire au secondaire, nuit à la réussite de l'élève. Le coefficient est particulièrement négatif au secondaire.

Par ailleurs, l'environnement scolaire et la qualité de l'enseignement contribuent au succès des élèves. Ainsi, cette étude tient compte des effets du milieu scolaire dans la modélisation des indicateurs de la réussite éducative. Il est notamment présumé que la stabilité de l'enseignement primaire par la présence en classe d'un seul enseignant tout au long de l'année scolaire peut favoriser le bon cheminement éducatif de l'élève. La proportion d'enseignants titulaires à temps plein dans chaque établissement scolaire vise à prendre en considération ce facteur présumé de la réussite. Les résultats du tableau A.1 indiquent qu'une augmentation de 1 pp de cet indicateur est associée à une hausse de 5 pp de la note à l'épreuve.

Enfin, le ministère de l'Éducation produit un indice de milieu socioéconomique pour l'établissement scolaire, qui est utilisé notamment pour le financement de programmes de soutien à la réussite. Cet indice marque le statut de défavorisation socioéconomique de l'établissement. Les résultats montrent que cette défavorisation est corrélée négativement avec la note scolaire de l'élève. En revanche, plus l'école se compose d'un nombre élevé d'élèves dont le revenu familial est élevé, plus la note en français écrit est haute. L'amplitude de la relation est particulièrement importante au secondaire. De plus, un pourcentage élevé d'élèves nés à l'étranger dans un établissement est corrélé négativement avec la note à l'épreuve. La taille des établissements scolaires, sur le plan des effectifs, joue également un rôle. Ainsi, il existe une corrélation négative : plus les effectifs sont élevés, plus les notes sont basses. Cette analyse confirme que la réussite scolaire de l'élève est tributaire d'un nombre non négligeable et indéterminé de facteurs individuels, scolaires et socioéconomiques. En effet, plusieurs facteurs sont fortement associés aux notes qu'il obtient.

Revenons maintenant au tableau 21. Compte tenu de l'effet de ces autres facteurs, les régressions par moindres carrés ordinaires suggèrent un impact négatif du redoublement scolaire sur les notes aux épreuves uniques (section A du tableau 21). Les effets estimés sont sensiblement moins importants que les écarts de notes brutes rapportés à la première ligne du tableau 21. Par exemple, l'écart statistique non ajusté pour l'épreuve de français écrit se chiffre à -10,8 pp, alors que l'estimation produite par MCO établit l'impact du redoublement à -8,3 pp. De la même façon, la différence statistique entre les résultats des

doubleurs et ceux des non-doubleurs à l'épreuve de science est estimée à -8,5 pp, alors que les résultats par MCO indiquent un écart de -5,5 pp.

La méthode de l'appariement par les scores de propension produit des estimations semblables à celles découlant des MCO. Son emploi est toutefois utile. D'une part, elle permet de consolider les résultats obtenus par le modèle paramétrique. D'autre part, elle offre la possibilité de tester l'existence d'erreurs statistiques qui pourraient être liées à de la sélection au redoublement selon des facteurs inobservables. Le test de Rosenbaum est utilisé à cette fin. Les valeurs statistiques de ce test sont rapportées au tableau 21. Il n'existe pas de barème officiellement établi qui permet de rejeter l'hypothèse de sélection selon des facteurs inobservables. Cependant, il est généralement admis que ces facteurs ne sont pas un enjeu dans les estimations si les valeurs de la statistique de Rosenbaum sont supérieures à 1,8 ou proches de ce seuil. Du coup, les valeurs du test rapportées à la section A du tableau 21 montrent que les estimations du redoublement scolaire sont probablement peu affectées par des facteurs autres que ceux spécifiés dans cette étude.

Par ailleurs, la ligne 4 indique les effets du redoublement selon l'approche des forêts aléatoires causales. Les effets estimés sur l'épreuve de français écrit sont comparables aux écarts statistiques entre les doubleurs et les non-doubleurs, à savoir -10,2 pp. La méthode permet d'estimer une baisse des résultats en mathématiques (CST) de 10,7 pp, en histoire de 12,7 pp et en anglais, langue seconde, de 4,4 pp causée par la reprise scolaire. Sauf en science, les estimations se comparent généralement à celles des autres méthodes s'appuyant sur l'HIC.

L'intérêt de cette étude est d'aborder l'estimation des effets du redoublement selon plusieurs méthodes et sources de données compte tenu de l'endogénéité présumée de la variable du redoublement scolaire rapportée par la littérature. L'une de ces méthodes est l'emploi des modèles classiques de sélection de l'échantillon, décrits à la section 5. Ces modèles offrent l'avantage de tenir compte plus ou moins efficacement des facteurs confondants inobservables pour les estimations du redoublement scolaire. La méthode permet également de vérifier si ces facteurs confondants présentent un enjeu statistique dans les estimations.

Les résultats produits par ces modèles montrent un effet atténué du redoublement scolaire pour quatre des cinq épreuves uniques (ligne 5 du tableau 21). De plus, à l'exception de l'épreuve de mathématiques (CST), ces modèles produisent les ajustements statistiques les plus importants. Ainsi, le redoublement scolaire diminue en moyenne de 7,5 pp la note en français écrit, d'environ 12,0 pp la note en mathématiques (CST) et en histoire, de 4,4 pp la note obtenue à l'épreuve d'anglais, langue seconde, et de 3,9 pp celle de l'épreuve de science. Le coefficient de l'inverse du ratio de Mills, indicateur de

l'importance de la corrélation entre le redoublement et les résultats scolaires, est négatif et significatif dans toutes les estimations. Ainsi, les facteurs inobservables qui amènent un élève à redoubler une année sont également ceux qui donnent lieu à de moins bons résultats aux épreuves uniques en ce qui le concerne. Ces facteurs confondants inobservables, contrairement aux tests de Rosenbaum rapportés plus haut, représentent bel et bien un enjeu statistique dans les estimations des effets du redoublement scolaire³¹.

³¹ Il est particulièrement intéressant de noter la forte corrélation, chiffrée à -0,9, entre les statistiques de Rosenbaum et les valeurs absolues du coefficient de l'inverse du ratio de Mills. Cette corrélation indique qu'une hausse de ce coefficient s'accompagne d'une baisse de la statistique de Rosenbaum. Dans les deux cas, il existe une indication d'un enjeu important des artifices statistiques inobservables dans les estimations que les modèles classiques de sélection de l'échantillon visent à corriger. Pour l'épreuve de science, par exemple, la statistique de Rosenbaum est la plus faible estimée, à savoir 1,8, tandis que la valeur du coefficient lié au ratio de Mills est de -0,984, soit l'effet le plus important de tous ceux rapportés à la section A du tableau. Du coup, il peut être observé que l'effet rapporté par la méthode de l'appariement par les scores de propension (-5,4) est corrigé par la méthode de sélection de l'échantillon (-3,8). En revanche, pour l'épreuve d'histoire, la statistique de Rosenbaum, chiffrée à 5, indique un enjeu relativement peu important des éléments inobservables que confirme autrement la faible valeur du coefficient du ratio de Mills relativement aux autres coefficients du ratio de Mills estimés. En conséquence, les estimations du redoublement procurées par les deux méthodes sont en fait identiques. En résumé, l'incidence des facteurs confondants inobservables sur les estimations du redoublement est rapportée de manière cohérente selon deux approches économétriques différentes. L'emploi de plusieurs méthodes économétriques dans cette étude permet bel et bien de valider les inférences statistiques sur les effets du redoublement scolaire.

Tableau 21 — Effets estimés du redoublement scolaire sur les notes aux épreuves uniques de français écrit, de mathématiques (CST), d'histoire, d'anglais, langue seconde, et de science, cohorte de la première année du primaire en 2006-2007, groupe témoin : élèves de la cohorte avant le redoublement

Section A. Redoublement au primaire et/ou au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-8,546 *	-15,155 *	-14,763 *	-6,966 *	-10,775 *
2. MCO-X	-5,480 *	-13,434 *	-12,806 *	-5,899 *	-8,313 *
3. Appariement plus proche voisin	-5,412 *	-13,352 *	-12,758 *	-6,003 *	-8,067 *
<i>Statistique Rosenbaum</i>	1,8	3,0	5,0	2,4	3,0
4. Forêts aléatoires causales	-12,937 *	-10,681 *	-12,752 *	-5,410 *	-10,203 *
5. Modèle de sélection (1 régime)	-3,857 *	-12,106 *	-12,475 *	-4,434 *	-7,504 *
<i>Lambda (inv.ratio Mills)</i>	-0,893 *	-0,777 *	-0,179 **	-0,799 *	-0,444 *

Section B. Redoublement au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-9,299 *	-16,434 *	-15,805 *	-7,109 *	-10,569 *
2. MCO-X	-5,829 *	-14,357 *	-13,458 *	-5,637 *	-7,693 *
3. Appariement plus proche voisin	-5,834 *	-14,172 *	-13,584 *	-5,712 *	-7,444 *
<i>Statistique Rosenbaum</i>	2,0	3,0	6,0	2,2	2,8
4. Forêts aléatoires causales	-13,195 *	-11,043 *	-12,991 *	-5,506 *	-10,350 *
5. Modèle de sélection (1 régime)	-3,896 *	-13,395 *	-13,226 *	-4,396 *	-6,574 *
<i>Lambda (inv.ratio Mills)</i>	-0,984 *	-0,529 *	-0,117	-0,629 *	-0,574 *

Section C. Redoublement au primaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1.Écart Reprise-sans reprise	-5,942 *	-12,043 *	-12,345 *	-6,613 *	-9,528 *
2.MCO-X	-4,042 *	-11,463 *	-11,329 *	-6,034 *	-8,225 *
3.Appariement plus proche voisin	-4,305 *	-11,373 *	-11,275 *	-6,460 *	-7,992 *
<i>Statistique Rosenbaum</i>	1,6	3,0	4,0	2,6	3,0
4.Forêts aléatoires causales	-7,163 *	-5,998 *	-7,057 *	-4,453 *	-6,917 *
5.Modèle de sélection (1 régime)	-1,757 *	-9,010 *	-10,492 *	-4,220 *	-7,366 *
<i>Lambda (inv.ratio Mills)</i>	-1,099 **	-1,265 *	-0,401 *	-0,866 *	-0,409 **

Il a été signalé, à la section 3, que certains pays appliquent la mesure du redoublement principalement au primaire, peut-être parce que son utilisation s'avère ainsi plus favorable. Cette étude visait à vérifier cette assertion. La section B du tableau 21 montre les effets du redoublement sur les élèves n'ayant redoublé qu'au secondaire, tandis que la section C se concentre sur les élèves n'ayant repris une année qu'au primaire. Une comparaison des chiffres des sections B et C révèle que le redoublement au secondaire comporte des effets délétères plus importants sur les notes de deux des cinq épreuves selon les estimations produites par la méthode de sélection de l'échantillon (ligne 5). Les résultats suggèrent que le redoublement au secondaire entraîne une baisse de 13,4 pp de la note en mathématiques comparativement à 9,0 pp pour le redoublement au primaire. En science, le redoublement au secondaire réduit la note de 3,9 pp, tandis que le redoublement au primaire entraîne une baisse de 1,7 pp. Enfin, les effets estimés du redoublement diffèrent peu selon qu'il s'agit du primaire ou du secondaire pour les épreuves d'histoire, de français écrit et d'anglais, langue seconde. En revanche, les estimations produites par les forêts aléatoires causales montrent des effets moins importants de la reprise au primaire sur les résultats à ces épreuves. On peut donc conclure que le redoublement au secondaire cause un plus grand préjudice à l'élève mais que cette mesure, qu'elle soit appliquée au primaire ou au secondaire, ne favorise pas sa réussite à long terme.

Les estimations rapportées au dernier tableau s'appuient sur la cohorte de tous les élèves inscrits à la première année du primaire en 2006-2007. Ainsi, ceux qui ont poursuivi leur scolarité normalement ont participé aux épreuves uniques de la quatrième secondaire en 2015-2016 et à celles de la cinquième secondaire en 2016-2017. Par contre, les élèves ayant repris une année ou plus ont passé les épreuves de la quatrième secondaire en 2016-2017 ou ultérieurement, tandis que leur participation aux épreuves de la dernière année du secondaire a eu lieu en 2017-2018 ou les années suivantes. Jusqu'ici, les comparaisons économétriques des notes obtenues par les deux groupes d'élèves, avec ou sans redoublement, ne portent donc pas sur le même examen et n'ont pas été effectuées avec les nouveaux camarades de classe de l'élève depuis le moment du redoublement.

Les résultats présentés dans le tableau 22 se basent sur les données concernant les élèves ayant effectué la même épreuve et constituent donc une comparaison de la cohorte entrée en première année du primaire en 2006-2007 et ayant repris une seule fois une année scolaire de même que de la cohorte entrée en première année du primaire en 2007-2008. Ainsi, le groupe témoin est formé des compagnons de classe des doubleurs au moment du test.

Tableau 22 — Effets estimés du redoublement scolaire sur les notes aux épreuves uniques de français écrit, de mathématiques (CST), d'histoire, d'anglais, langue seconde, et de science, cohorte de la première année du primaire en 2006-2007, groupe témoin : élèves de la cohorte après le redoublement

Section A. Redoublement au primaire et/ou au secondaire

	Sciences	Math	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-12,975	-10,186 *	-13,861 *	-7,996 **	-10,161 *
2. MCO-X	-10,888	-8,712 *	-11,609 *	-6,825 **	-8,181 *
3. Appariement plus proche voisin	-10,801	-8,751 *	-11,266 *	-6,779 *	-7,896 *
<i>Statistique Rosenbaum</i>	5,0	2,2	4,0	2,6	3,0
4. Modèle de sélection (1 régime)	-8,711	-6,071 *	-9,552 *	-5,458 **	-6,782 *
<i>Lambda (inv.ratio Mills)</i>	-1,136	-1,475 *	-1,076 *	-0,707 *	-0,729 *

Section B. Redoublement au secondaire

	Sciences	Math	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-14,608	-12,057 *	-15,795 *	-8,358 **	-10,681 *
2. MCO-X	-12,089	-10,106 *	-12,931 *	-6,752 **	-8,070 *
3. Appariement plus proche voisin	-11,966	-9,819 * [✓]	-12,687 * [✓]	-6,726 **	-7,822 *
<i>Statistique Rosenbaum</i>	6,0	2,4	5,0	2,6	3,0
4. Modèle de sélection (1 régime)	-9,514	-8,231 *	-10,905 *	-5,507 **	-6,631 *
<i>Lambda (inv.ratio Mills)</i>	-1,26	-0,993 *	-0,994 *	-0,604 **	-0,707 *

Section C. Redoublement au primaire

	Sciences	Math	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-11,121	-7,620 *	-11,658 *	-7,577 **	-9,487 *
2. MCO-X	-9,397	-6,664 *	-9,937 *	-6,786 **	-8,147 *
3. Appariement plus proche voisin	-9,238	-6,484 * [✓]	-9,672 * [✓]	-6,889 **	-7,951 *
<i>Statistique Rosenbaum</i>	3,0	1,8	3,0	2,8	3,0
4. Modèle de sélection (1 régime)	-7,736	-3,085 *	-8,136 *	-5,386 **	-7,263 *
<i>Lambda (inv.ratio Mills)</i>	-0,779	-1,787 *	-0,847 *	-0,652 **	-0,411 **

Ces estimations portent donc sur les épreuves de quatrième et de cinquième secondaire tenues en 2016-2017 et en 2017-2018 respectivement. Nous présentons l'ensemble des résultats, mais notre description se concentre sur les résultats des modèles de la sélection de l'échantillon, puisque ceux-ci tiennent compte des facteurs confondants inobservables³². Les chiffres indiquent peu de changements dans les résultats en ce qui concerne les épreuves de français écrit et d'anglais, langue seconde. Les effets estimés pour les mathématiques (CST) sont toutefois moins importants que ceux rapportés précédemment. Le redoublement abaisse de 8,2 pp en moyenne la note à l'épreuve pour les élèves dont la reprise s'est produite uniquement au secondaire et de 3,1 pp seulement pour les élèves n'ayant redoublé qu'au primaire. Les effets estimés sont aussi moins importants pour les résultats à l'épreuve d'histoire. Ils sont plus élevés en revanche pour l'épreuve de science de quatrième secondaire, soit 9,5 pp pour le redoublement au secondaire comparativement à 3,9 pp précédemment. En science, l'effet est aussi plus important pour les reprises ayant lieu au primaire, puisque les nouvelles estimations montrent des réductions de 7,7 pp comparativement à 1,8 pp. Il semblerait donc que, quel que soit le groupe de comparaison des doubleurs, soit les camarades au moment de l'entrée à l'école ou les camarades au moment des tests réalisés à la fin du secondaire, les résultats suggèrent que le redoublement ne permet pas aux élèves de mieux réussir aux épreuves uniques. Il est également important de se rappeler que ces résultats ne concernent pas les nombreux élèves qui abandonnent ou choisissent un parcours autre que la formation générale. Il est donc probable que ces effets soient en réalité sous-estimés.

En outre, cette étude porte sur les effets du redoublement chez les filles et les garçons pris séparément. Le tableau 23 présente les estimations pour les filles et le tableau 24, celles concernant les garçons. Seules les estimations relatives aux examens ayant eu lieu la même année sont présentées³³. La comparaison des deux tableaux montre que le redoublement induit une baisse de résultats scolaires plus importante pour les filles que pour les garçons, et cela, pour l'ensemble des épreuves uniques et sans distinction à l'égard du moment de la reprise scolaire. Cette diminution des résultats causée par la reprise scolaire est de 2 à 3 pp plus importante pour les filles que pour les garçons. Les estimations produites par le modèle de sélection de l'échantillon montrent, par exemple, que le redoublement scolaire

³² Les forêts aléatoires causales n'ont pas produites d'estimations pour ce cas particulier.

³³ Les autres estimations donnent des résultats comparables à ceux présentés dans cette section. Les tableaux sont disponibles sur demande. Il est à noter que les estimations produites par la méthode des forêts aléatoires causales concernent toutefois toutes les cohortes citées plus haut et ne restreint pas les comparaisons à une même année d'examen.

a réduit de 8,0 pp les résultats des filles à l'examen de français écrit en 2017-2018 comparativement à 5,8 pp pour les garçons³⁴.

³⁴ Bien que les tableaux ne comportent pas les écarts-types ainsi que les intervalles de confiance des paramètres estimés, les chevauchements des intervalles de confiance sont très peu fréquents de sorte que les écarts entre les filles et les garçons dans les effets du redoublement sont statistiquement crédibles.

Tableau 23 — Effets estimés du redoublement scolaire des filles sur les notes aux épreuves uniques, cohorte des filles de 2006-2007

Section A. Redoublement au primaire et/ou au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-14,531 *	-10,954 *	-15,972 *	-9,646 *	-10,759 *
2. MCO-X	-11,961 *	-9,341 *	-13,486 *	-8,552 *	-9,100 *
3. Appariement plus proche voisin	-11,989 *	-9,379 *	-13,188 *	-8,376 *	-8,770 *
<i>Statistique Rosenbaum</i>	6,0	2,4	5,0	3,0	4,0
4. Forêts aléatoires causales	-12,447 *	-10,322 *	-11,967 *	-4,709 *	-9,847 *
5. Modèle de sélection (1 régime)	-10,040 *	-6,945 *	-11,829 *	-7,628 *	-7,987 *
<i>Lambda (inv.ratio Mills)</i>	-0,979 *	-1,305 *	-0,844 *	-0,467 *	-0,565 *

Section B. Redoublement au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-16,503 *	-13,209 *	-18,388 *	-10,516 *	-10,885 *
2. MCO-X	-13,362 *	-11,129 *	-15,252 *	-9,091 *	-8,750 *
3. Appariement plus proche voisin	-13,090 *	-10,971 *	-14,840 *	-8,865 *	-8,290 *
<i>Statistique Rosenbaum</i>	6,0	2,6	6,0	3,0	3,0
4. Forêts aléatoires causales	-12,883 *	-10,905 *	-12,241 *	-4,780 *	-10,003 *
5. Modèle de sélection (1 régime)	-11,272 *	-10,006 *	-13,750 *	-8,315 *	-7,874 *
<i>Lambda (inv.ratio Mills)</i>	-0,982	-0,571 **	-0,706 *	-0,363 **	-0,414 **

Section C. Redoublement au primaire

	Sciences	Math CST	Histoire	Histoire	Français écrit
1. Écart Reprise-sans reprise	-12,705 *	-8,419 *	-13,771 *	-8,792 *	-10,632 *
2. MCO-X	-10,575 *	-7,192 *	-11,763 *	-7,926 *	-9,330 *
3. Appariement plus proche voisin	-10,444 *	-7,922 *	-11,647 *	-8,075 *	-9,286 *
<i>Statistique Rosenbaum</i>	4,0	2,2	4,0	3,0	4,0
4. Forêts aléatoires causales	-6,429 *	-3,127 *	-5,789 *	-3,935 *	-5,996 *
5. Modèle de sélection (1 régime)	-9,037 *	-3,853 *	-10,384 *	-6,855 *	-8,363 *
<i>Lambda (inv.ratio Mills)</i>	-0,715 *	-1,653 *	-0,641 **	-0,493 *	-0,446 ***

Note : Les filles du groupe de traitement sont entrées en première année du primaire en 2006-2007. Les filles du groupe témoin sont celles qui ont passé les épreuves au même moment que les filles du groupe traité.

Tableau 24 — Effets estimés du redoublement scolaire des garçons sur les notes aux épreuves uniques : cohorte des garçons de 2006-2007

Section A. Redoublement au primaire et/ou au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-11,956 *	-9,327 *	-12,105 *	-6,312 *	-8,823 *
2. MCO-X	-10,051 *	-8,181 *	-10,143 *	-5,412 *	-7,410 *
3. Appariement plus proche voisin	-9,883 *	-8,122 *	-9,654 *	-5,361 *	-7,294 *
<i>Statistique Rosenbaum</i>	4,0	2,0	3,0	2,0	2,8
4. Forêts aléatoires causales	-14,084 *	-10,411 *	-14,073 *	-6,131 *	-9,407 *
5. Modèle de sélection (1 régime)	-8,578 *	-5,618 *	-8,735 *	-4,070 *	-5,838 *
<i>Lambda (inv.ratio Mills)</i>	-0,784 *	-1,459 *	-0,752 *	-0,709 *	-0,835 *

Section B. Redoublement au secondaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-13,650 *	-10,949 *	-13,967 *	-6,410 *	-9,367 *
2. MCO-X	-11,300 *	-9,392 *	-11,448 *	-5,143 *	-7,575 *
3. Appariement plus proche voisin	-11,160 *	-9,383 *	-11,384 *	-4,901 *	-7,600 *
<i>Statistique Rosenbaum</i>	5,0	2,2	4,0	1,8	2,8
4. Forêts aléatoires causales	-14,234 *	-10,678 *	-14,283 *	-6,212 *	-9,426 *
5. Modèle de sélection (1 régime)	-9,441 *	-7,413 *	-10,293 *	-4,061 *	-5,732 *
<i>Lambda (inv.ratio Mills)</i>	-0,934 *	-1,078 *	-0,583 *	-0,538 *	-0,930 *

Section C. Redoublement au primaire

	Sciences	Math CST	Histoire	Anglais LS	Français écrit
1. Écart Reprise-sans reprise	-9,669 *	-6,741 *	-9,559 *	-6,203 *	-7,962 *
2. MCO-X	-8,232 *	-6,142 *	-8,192 *	-5,678 *	-6,956 *
3. Appariement plus proche voisin	-8,150 *	-6,048 *	-8,010 *	-5,619 *	-6,367 *
<i>Statistique Rosenbaum</i>	3,0	1,4	2,6	2,2	2,4
4. Forêts aléatoires causales	-7,660 *	-4,558 *	-7,581 *	-4,749 *	-7,661 *
5. Modèle de sélection (1 régime)	-6,806 *	-2,326 *	-6,488 *	-3,958 *	-6,259 *
<i>Lambda (inv.ratio Mills)</i>	-0,677 *	-1,927 *	-0,811 *	-0,812 *	-0,328

Note : Les garçons du groupe de traitement sont entrés en première année du primaire en 2006-2007. Les garçons du groupe témoin sont ceux qui ont passé les épreuves au même moment que les garçons du groupe traité.

6.2.2 Diplomation et sortie sans diplôme ni qualification

La section 4 a montré que les élèves de la FGJ ayant repris une année ou plus présentent des taux de sortie sans diplôme nettement plus importants que les autres élèves. Les taux de non-diplomation, toutes formations confondues, sont également beaucoup plus élevés pour les doubleurs. Le taux moyen d'obtention d'un DES est, quant à lui, très faible chez ces derniers par rapport aux élèves n'ayant pas eu à reprendre une année scolaire. En revanche, il a été indiqué à la section 2 qu'un jeune qui n'a pas atteint les objectifs des programmes d'études de l'enseignement primaire peut suivre une formation préparatoire au travail. Par ailleurs, les statistiques révèlent des taux d'obtention d'un certificat de formation à un métier semi-spécialisé nettement favorables chez les doubleurs. Cette section vise à vérifier si ces écarts dans les indicateurs de la réussite scolaire des élèves avec ou sans redoublement sont causés précisément par la reprise scolaire.

6.2.2.1 Résultats obtenus selon la modélisation combinée de la diplomation et de la sortie sans diplôme ni qualification : modèle de probit ordonné

La section 5 a présenté de façon succincte le processus décisionnel de l'élève au regard de ses choix de scolarisation. Les élèves qui anticipent des avantages sociaux ou économiques relativement importants en ce qui a trait à une formation en particulier opteront pour celle-ci et le diplôme qui lui est associé. D'autres, par exemple, jugeront trop élevées les pertes de revenu de travail pour terminer un programme de la FGJ et obtenir un DES. Pour ceux-ci, les efforts à consentir pour la scolarisation ou des difficultés ponctuelles dans le cheminement scolaire sont si importants que la seule option envisageable est de cesser hâtivement leur scolarité.

Le modèle de probit ordonné à traitement endogène, mentionné à la section 5, modélise de manière adéquate les choix de scolarisation de l'élève. Trois options sont définies : l'obtention d'un DES en sept ans, l'obtention d'un diplôme autre qu'un DES en sept ans et la sortie sans diplôme ni qualification toutes formations confondues. Le choix d'une option particulière est expliqué par le même groupe de variables que celui utilisé dans la modélisation des notes aux épreuves uniques. Le modèle est estimé par la commande « eoprobit » de la version 15 du logiciel économétrique Stata.

Le tableau 25 présente les effets estimés du redoublement sur les choix de scolarisation des élèves ayant repris une année ou plus au cours de leur cheminement scolaire. Les régressions ont été effectuées pour les données regroupées des cohortes d'élèves inscrits à la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007, ce qui porte à 264 380 le nombre d'observations utiles aux fins de l'estimation. Les trois

premières colonnes du tableau 25 concernent les résultats produits par le modèle de probit ordonné à traitement endogène³⁵. Le paramètre de traitement précisément estimé dans ce contexte est l'effet moyen du redoublement sur les *doubleurs conformistes*³⁶, c'est-à-dire les doubleurs nés au troisième trimestre de l'année. Les trois dernières colonnes montrent les estimations produites par forêts aléatoires causales, sur lesquelles a été appliqué de nouveau le modèle de probit ordonné sans le trimestre de naissance comme instrument de détermination de l'effet de traitement. Les résultats rapportés dans ces colonnes indiquent les effets moyens du redoublement scolaire sur la *totalité* des doubleurs.

Les résultats des colonnes 1 et 4 présentent les effets moyens du redoublement sur tous les élèves ayant repris une année scolaire. Les colonnes 2, 3, 5 et 6 montrent de façon spécifique ses effets sur la diplomation chez les filles et les garçons. La section A du tableau 25 concerne les effets du redoublement sur les doubleurs, peu importe le moment de la reprise. La section B indique, pour sa part, les effets observés s'il se produit au secondaire seulement et la section C, s'il se produit au primaire uniquement.

Ainsi, les résultats de la section A suggèrent un impact important du redoublement sur les choix de scolarisation, et ce, aussi bien pour les doubleurs conformistes que pour la totalité des doubleurs. Le redoublement abaisse de 41,8 pp la propension des doubleurs du premier groupe et de 38,0 pp la propension de ceux du second groupe à obtenir un DES en sept ans. Cette mesure hausse également le risque de non-diplomation pour les deux groupes de doubleurs. En revanche, la reprise soutient de manière efficace les élèves qui optent pour une formation préparatoire au marché du travail ou une formation menant à un métier semi-spécialisé. L'effet est particulièrement remarquable chez les doubleurs nés au troisième trimestre de l'année, puisqu'il augmente de 17,2 pp la probabilité d'obtenir le diplôme lié à cette formation qualifiante. Les formations professionnelles offertes aux élèves en difficulté permettent probablement d'accroître les gains anticipés de l'obtention d'un diplôme, notamment par une réduction des coûts relatifs à celle-ci et une augmentation des gains à court terme. Autrement dit, ces formations peuvent inciter plusieurs élèves à investir du temps et des efforts additionnels dans leurs apprentissages, étant donné que des gains financiers sont envisageables sur un horizon écourté. Il est à noter, par ailleurs, que le paramètre de corrélation des termes d'erreur du modèle, estimé par une maximisation de la fonction de vraisemblance, indique que les facteurs inobservables dictent en partie les décisions du redoublement et de l'obtention du diplôme.

³⁵ Les estimations ont été produites avec le logiciel Stata 15 et la commande « eoprobit » accompagnée de l'option « entreat ».

³⁶ Le terme « conformiste » est employé pour le mot anglais « complier » dans la littérature sur les effets de traitement.

Par ailleurs, la section A du tableau 25 contient des résultats qui concernent les filles et les garçons. On peut y constater que ces résultats divergent selon le paramètre de traitement analysé. Contrairement aux résultats portant sur les notes scolaires, le redoublement est moins préjudiciable à l'obtention du diplôme aux études secondaires chez les filles qu'il ne l'est chez les garçons. Les estimations, statistiquement significatives, montrent que la reprise scolaire abaisse de 35,2 pp (section A, colonne 2) la propension d'une fille à obtenir un diplôme du secondaire, alors que, pour un garçon, la diminution est de 52,1 pp (section A, colonne 3). Le redoublement est donc plus dommageable pour les garçons en ce qui concerne les choix de scolarisation au secondaire. Cette mesure hausse également de manière plus importante la propension des garçons à quitter l'école sans diplôme ni qualification. Les chiffres indiquent en effet un impact de 32,5 pp (section A, colonne 3) pour les garçons à cet égard, tandis que, pour les filles, le redoublement n'a pas d'effet statistiquement significatif sur la propension à ne pas obtenir de diplôme (section A, colonne 2). Par ailleurs, la reprise d'une année chez les garçons est perçue comme une mesure plus favorable à l'obtention d'un diplôme dans des délais écourtés, puisqu'elle accroît de 19,5 pp la probabilité d'obtenir un CFMS ou un diplôme de la FP (section A, colonne 3). Chez les filles, l'impact est estimé à 14,5 pp (section A, colonne 2).

Ces résultats qui touchent les filles et les garçons concernent uniquement les doubleurs conformistes. Si l'attention est portée aux estimations des colonnes 5 et 6, il ne fait aucun doute, à l'instar des estimations portant sur les notes aux épreuves uniques, que les effets moyens du redoublement sur les filles, peu importe leur trimestre de naissance, sont plus délétères que sur les garçons. Le redoublement entraîne en effet chez les filles une baisse de la diplomation aux études secondaires et hausse le risque de ne pas avoir de diplôme qualifiant.

De plus, les résultats rapportés plus haut montrent une image des effets moyens localisés et des effets moyens du redoublement scolaire sur la diplomation chez les doubleurs, qu'il s'agisse du primaire ou du secondaire. Les sections B et C apportent des nuances importantes sur la portée de l'impact selon les cycles scolaires. D'abord, les résultats de la section B, soit ceux qui ont trait à l'impact du redoublement n'ayant lieu qu'au secondaire, se comparent aux effets moyens discutés plus haut. Ainsi, le redoublement réduit de 54,7 pp la propension d'un doubleur conformiste à obtenir un DES, hausse de 15,8 pp la probabilité de recevoir un CFMS et augmente de 38,9 pp le risque de quitter l'école sans diplôme ni qualification. La reprise d'une année diminue de façon un peu moins importante la propension des filles conformistes à obtenir un diplôme aux études secondaires. Elle favorise par contre un peu plus les garçons pour l'obtention d'un CFMS. Quant à l'impact estimé pour la totalité des doubleurs du secondaire (colonnes 3 à 6), il est moins important,

mais influence dans le même sens leurs choix de scolarisation. Toutefois, les filles s'avèrent plus touchées que les garçons.

La section C du tableau 25 montre cependant que le redoublement, s'il n'est appliqué qu'au primaire, comporte des effets sur les choix de scolarisation des doubleurs qui diffèrent, d'une part, en amplitude et, d'autre part, selon qu'ils sont nés au troisième trimestre de l'année ou non. Ainsi, les résultats rapportés dans les colonnes 1 à 3, c'est-à-dire ceux des doubleurs conformistes, indiquent que le redoublement au primaire accroît de 25,4 pp la propension de l'élève à obtenir un DES. Cette mesure fait augmenter également de 9,8 pp la propension des élèves du secondaire à se qualifier pour un métier semi-spécialisé et abaisse de 35,2 pp la probabilité qu'ils quittent l'école sans diplôme ni qualification. Le redoublement est donc particulièrement favorable aux élèves nés au troisième trimestre de l'année.

Néanmoins, les chiffres du tableau 25 précisent que ce sont surtout les garçons conformistes qui profitent de la mesure de soutien au primaire. En effet, la reprise d'une année hausse de 45,8 pp leur propension à obtenir un diplôme au secondaire, alors que, pour les filles, l'effet est limité à 4,8 pp. La propension des filles à obtenir un CFMS n'est pas influencée par la reprise au primaire, alors qu'elle augmente de 17,7 pp celle des garçons. Enfin, le risque de non-diplomation chez les garçons est diminué de 63,5 pp en raison du redoublement au primaire, alors que, pour les filles, l'impact s'avère beaucoup plus faible.

Rappelons que les résultats présentés plus haut ne s'appliquent qu'aux doubleurs conformistes. Les colonnes 4 à 6 montrent que le redoublement scolaire désavantage de nouveau les élèves lorsque ses effets sont estimés pour la totalité des doubleurs du primaire. L'impact sur ces derniers est cependant moins élevé que celui estimé pour les doubleurs du secondaire. Le redoublement réduit de 18,5 pp la propension de l'élève à obtenir un diplôme en sept ans aux études secondaires, hausse faiblement la probabilité qu'il se qualifie pour un métier semi-spécialisé et augmente de 14,2 pp le risque de non-diplomation. Les effets sont de nouveau plus importants pour les filles que pour les garçons.

En résumé, le redoublement scolaire contribue rarement à hausser la probabilité d'obtenir un diplôme d'études secondaires de la formation générale et accentue plutôt la probabilité de quitter l'école sans diplôme ni qualification. Le seul groupe qui bénéficie de façon importante du redoublement est celui formé des garçons nés au troisième trimestre de l'année, et ce, uniquement s'il se produit au primaire.

Tableau 25 — Effets estimés du redoublement scolaire sur la diplomation et les sorties sans diplôme ni qualification, cohortes de la première année du primaire en 2003-2004, en 2004-2005, en 2005-2006 et en 2006-2007

Section A. *Variable traitement* : reprise d'une année ou plus de scolarité entre la première année du primaire et la cinquième année du secondaire

	(1)	(2)	(3)	(4)	(5)	(6)
	Effet moyen localisé (doubleurs-conformistes)			Effet moyen sur les doubleurs		
	Tous	Filles	Garçons	Tous	Filles	Garçons
Diplomation en 7 ans ES						
Effet moyen partic.	-0,418	-0,352	-0,521	-0,380	-0,389	-0,313
Écart-type	0,106	0,147	0,042	0,003	0,003	0,003
P > z	0,000	0,017	0,000	0,000	0,000	0,000
Diplomation en 7 ans (FP, CFSP)						
Effet moyen partic.	0,172	0,145	0,195	0,065	0,057	0,056
Écart-type	0,008	0,013	0,003	0,002	0,002	0,002
P > z	0,000	0,000	0,000	0,000	0,000	0,000
Aucun diplôme						
Effet moyen partic.	0,246	0,206	0,325	0,315	0,332	0,258
Écart-type	0,098	0,133	0,040	0,003	0,003	0,003
P > z	0,012	0,123	0,000	0,000	0,000	0,000
Corrélation(eY,eR)	0,356	0,457	0,172			
P > z	0,017	0,013	0,020			
N	264 380	130 145	134 235	131 344	66 641	64 702

Section B. *Variable traitement* : reprise d'une année ou plus de scolarité entre la première et la cinquième année du secondaire

	(1)	(2)	(3)	(4)	(5)	(6)
	Effet moyen localisé (doubleurs-conformistes)			Effet moyen sur les doubleurs		
	Tous	Filles	Garçons	Tous	Filles	Garçons
Diplomation en 7 ans ES						
Effets marginaux	-0,547	-0,498	-0,575	-0,310	-0,362	-0,249
Écart-type	0,034	0,066	0,035	0,004	0,004	0,004
P > z	0,000	0,000	0,000	0,000	0,000	0,000
Diplomation en 7 ans FP, CFSP)						
Effets marginaux	0,158	0,123	0,177	0,064	0,074	0,046
Écart-type	0,002	0,003	0,003	0,002	0,002	0,002
P > z	0,000	0,000	0,000	0,000	0,000	0,000
Aucun diplôme						
Effets marginaux	0,389	0,375	0,398	0,246	0,289	0,204
Ecart-type	0,035	0,064	0,036	0,003	0,003	0,003
P > z	0,000	0,000	0,000	0,000	0,000	0,000
Corrélation(eY,eR)	-0,102	-0,047	0,131			
P > z	0,147	0,700	0,116			
N	230 168	116 977	113 191	114 547	56 324	58 223

Section C. *Variable traitement* : reprise d'une année ou plus de scolarité entre la première et la sixième année du primaire

	(1)	(2)	(3)	(4)	(5)	(6)
	Effet moyen localisé (doubleurs-conformistes)			Effet moyen sur les doubleurs		
	Tous	Filles	Garçons	Tous	Filles	Garçons
Diplomation en 7 ans ES						
Effets marginaux	0,254	0,046	0,458	-0,185	-0,225	-0,147
Écart-type	0,044	0,009	0,007	0,045	0,005	0,004
P > z	0,000	0,000	0,000	0,000	0,000	0,000
Diplomation en 7 ans FP, CFSP)						
Effets marginaux	0,098	0,002	0,177	0,044	0,059	0,029
Écart-type	0,008	0,002	0,006	0,002	0,003	0,002
P > z	0,000	0,383	0,000	0,000	0,000	0,000
Aucun diplôme						
Effets marginaux	-0,352	-0,048	-0,635	0,142	0,166	0,118
Écart-type	0,036	0,007	0,006	0,003	0,004	0,003
P > z	0,320	0,000	0,000	0,000	0,000	0,000
Corrélation(eY,eR)	0,837	0,862	0,984			
P > z	0,000	0,000	0,000			
N	195 500	104 858	90 642	97 333	45 103	52 229

6.2.2.2 Résultats : sortie sans diplôme ni qualification à la FGJ

Cette section apporte un éclairage complémentaire par rapport à la section précédente. Quatre méthodes ont été mises à contribution pour estimer l'impact du redoublement scolaire sur la propension de l'élève à quitter la FGJ sans diplôme ni qualification : le modèle de probit classique, le modèle de probit avec changement de régime pour des variables discrètes, la méthode de l'appariement par les scores de propension et les forêts aléatoires causales. Tout comme dans la section précédente, les estimations portent sur les cohortes des élèves inscrits à la première année du primaire en 2005-2006, en 2006-2007 et en 2007-2008³⁷.

Il est d'abord utile de s'intéresser aux effets des facteurs explicatifs des sorties sans diplôme ni qualification afin de mieux juger des effets estimés du redoublement rapportés un peu plus loin. L'analyse porte sur les estimations produites par le modèle de probit classique, lesquelles sont signalées au tableau A.2 de l'annexe A. Les chiffres montrent que les filles ont une propension moins importante que les garçons à quitter la FGJ sans diplôme ni qualification. En revanche, être né à l'étranger est associé à une hausse de cette

³⁷ Les renseignements administratifs sur le statut longitudinal des élèves au regard de leur sortie sans diplôme ni qualification sont disponibles à partir de 2011-2012 uniquement, ce qui exclut, du coup, la prise en compte des élèves inscrits à la première année du primaire avant 2005-2006.

propension de 1 pp. L'élève risque également davantage d'abandonner la FGJ si seule la mère assume la garde des enfants ou que les parents sont sous-scolarisés. Ce risque diminue en retour pour les élèves vivant dans un secteur où les parents ont de meilleurs revenus d'emploi.

De plus, le programme Passe-Partout, initialement destiné aux familles socioéconomiquement désavantagées, vise maintenant toute famille souhaitant participer à des activités éducatives avec des tout-petits d'âge préscolaire. Les chiffres du tableau montrent que ce programme est lié à une baisse de 1,4 pp de la propension à quitter prématurément la FGJ. La fréquentation d'une école privée au secondaire diminue, pour sa part, de 3,1 pp le risque de décrochage et la participation de l'élève à un projet pédagogique particulier entraîne une baisse de 2,6 pp. Il a été rapporté plus haut que le changement d'établissement scolaire contribuait à réduire les notes aux épreuves uniques du MEQ. Or, l'effet de ce changement est également ressenti sur le cheminement scolaire de l'élève, puisqu'il augmente d'environ 2 pp la probabilité qu'il quitte la FGJ sans diplôme ni qualification.

De nouveau, la présence d'un pourcentage plus important d'enseignants titulaires occupant un poste à temps plein dans un établissement primaire est favorable à la réussite scolaire. Dans le cas actuel, cela contribue à réduire de 2,7 pp le risque de quitter la FGJ. La défavorisation socioéconomique de l'établissement scolaire, établie par une hausse de l'IMSE, augmente au contraire ce risque. Les chiffres du tableau indiquent aussi que la vétusté des immeubles scolaires contribue à accroître le risque de décrochage à la FGJ.

Le tableau 26 montre précisément les effets estimés du redoublement scolaire sur le risque de quitter la FGJ sans diplôme ni qualification. La section A de ce tableau concerne les effets estimés du redoublement scolaire pour tous les élèves ayant repris une année ou plus entre la première année du primaire et la cinquième secondaire. La première colonne présente les résultats pour tous les élèves, alors que la deuxième et la troisième colonnes rapportent les effets constatés pour les filles et les garçons. Quant à la section B, elle porte sur l'impact du redoublement s'il se produit au secondaire et la section C, s'il a lieu uniquement au primaire.

La première ligne de la section A du tableau 26 montre, à l'instar des statistiques descriptives présentées plus haut, un écart de 17,0 pp entre un doubleur et un non-doubleur pour ce qui est de la probabilité de décrochage scolaire à la FGJ. L'estimation produite par le modèle de probit classique, dont les résultats ont été rapportés précédemment, indique que la reprise scolaire augmente de 11,7 pp la propension de l'élève à quitter de façon hâtive la FGJ. Les effets délétères de la reprise sont corroborés par toutes les autres méthodes. Par contre, les forêts aléatoires causales limitent les effets estimés à 8,2 pp. De

plus, les résultats montrent que le redoublement affecte davantage la propension à l'abandon chez les garçons que chez les filles.

Les sections B et C du tableau 26 indiquent, au premier abord, que le redoublement scolaire augmente la propension d'un élève à quitter la FGJ sans diplôme ni qualification, qu'il se produise au primaire ou au secondaire. L'impact est cependant plus faible en ce qui concerne le redoublement qui a lieu au primaire. Les forêts aléatoires causales soulignent, par exemple, que le redoublement qui se produit au secondaire augmente de 8,8 pp la probabilité d'abandon, tandis que, restreint au primaire, il limite cet effet à 4,2 pp. Les estimations offertes par les autres méthodes se rallient à ces résultats, mais l'amplitude des effets est plus importante.

Un regard attentif sur les estimations produites par forêts aléatoires causales instrumentales montre que la reprise scolaire en soutien aux élèves du primaire réduit la propension à quitter la FGJ sans diplôme ni qualification plutôt que d'intensifier cette propension comme il est indiqué plus haut. La mesure réduit d'environ 4 pp la probabilité de décrochage scolaire à la FGJ. Ce résultat rejoint précisément les estimations produites par le modèle de probit ordonné et présentées plus haut dans cette section. Il s'explique principalement par la qualité de la méthode pour ce qui est d'exploiter le troisième trimestre de naissance comme source de renseignements utile à la détermination des effets du redoublement scolaire.

Tableau 26 — Effets estimés du redoublement scolaire sur les sorties sans diplôme ni qualification à la FGJ, cohortes de la première année du primaire en 2005-2006, en 2006-2007 et en 2007-2008

Section A. *Variable traitement* : reprise d'une année ou plus de scolarité entre la première année du primaire et le cinquième secondaire, unité : points de pourcentage

	Tous		Garçons		Filles	
1. Écart statistique Reprise-sans reprise	17,0		18,1		14,7	
2. Modèle PROBIT	11,7	*	12,7	*	9,9	*
3. Modèle changement endogène 2 régimes (EMP)	17,2	*	18,8	*	14,7	*
<i>Test RV, rho1=rho0=0; Prob > chi2</i>	0,03		0,14		0,24	
4. Appariement plus proche voisin (EMP)	15,3	*	16,5	*	13,0	*
<i>Statistique Rosenbaum</i>	> 6		> 6		> 6	
5. Forêts aléatoires causales	9,499	*	8,131	*	10,514	*
6. Forêts aléatoires causales instrumentales	8,229	*	7,010	*	9,004	*

Nombre d'observations	194 492		99 163		95 329	

Section B. *Variable traitement* : reprise d'une année ou plus de scolarité entre le premier et le cinquième secondaire, unité : points de pourcentage

	Tous		Garçons		Filles	
1. Écart Reprise-sans reprise	16,1		16,7		14,7	
2. Modèle PROBIT	11,1	*	12,2	*	9,9	*
3. Modèle changement endogène 2 régimes (EMP)	16,2	*	17,3	*	14,4	*
<i>Test RV, rho1=rho0=0; Prob > chi2</i>	0,06		0,30		0,43	
4. Appariement plus proche voisin (EMP)	14,4	*	15,2	*	13,1	*
<i>Statistique Rosenbaum</i>	> 6		> 6		> 6	
5. Forêts aléatoires causales	8,799	*	8,825	*	9,890	*
6. Forêts aléatoires causales instrumentales	12,337	*	11,790	*	12,618	*

Nombre d'observations	167 674		82 722		84 952	

Section C. *Variable traitement* : reprise d'une année ou plus de scolarité entre la première année et la sixième année du primaire, unité : points de pourcentage

	Tous		Garçons		Filles	
1. Écart Reprise-sans reprise	11,7		13,7		9,2	
2. Modèle PROBIT	7,5	*	9,1	*	6,1	*
3. Modèle changement endogène 2 régimes (EMP)	13,3	*	15,9	*	10,3	*
<i>Test RV, rho1=rho0=0; Prob > chi2</i>	0,00		0,00		0,30	
4. Appariement plus proche voisin (EMP)	10,7	*	12,4	*	10,5	*
<i>Statistique Rosenbaum</i>	> 6		> 6		> 6	
5. Forêts aléatoires causales	4,170	*	4,290	*	5,835	*
6. Forêts aléatoires causales instrumentales	-4,236	*	-4,255	*	-4,002	*

Nombre d'observations	146 314		68 202		78 112	

6.3 Effets du redoublement scolaire sur les indicateurs de la réussite scolaire : données du PISA

Cette section présente les estimations réalisées à partir des données du PISA. La section 6.3.1 traite des régressions par moindres carrés ordinaires des résultats aux évaluations portant sur le redoublement ou le retard scolaire. La section 6.3.2 concerne les régressions par variables instrumentales des résultats à ces évaluations. L'instrument utilisé dans ces régressions correspond au fait qu'un élève est né au cours du trimestre précédant la date d'admissibilité à l'école. Puis, la section 6.3.3 traite des régressions sous forme réduite des résultats aux évaluations portant sur le redoublement ou le retard scolaire. Enfin, dans la section 6.3.4, une approche de différence en différence est utilisée en exploitant le changement de régime observé au Québec à compter de 2008.

6.3.1 Résultats des estimations par moindres carrés ordinaires

La littérature qui tient compte du problème de sélection dans l'estimation des effets du redoublement suggère que cette mesure ne conduit pas à une amélioration de la performance scolaire à long terme (Valbuena *et al.*, 2020). La méta-analyse de Valbuena et ses collègues (2020) distingue les effets du redoublement sur la performance scolaire selon qu'il a eu lieu au primaire ou au secondaire et selon que l'effet est à court ou à long terme. Ainsi, à court terme, la littérature recensée dans cette méta-analyse présente des effets mixtes du redoublement au primaire et négatifs du redoublement au secondaire en ce qui a trait à la performance scolaire (Valbuena *et al.*, 2020). À long terme, la littérature présente des effets non significatifs ou négatifs du redoublement ayant eu lieu au primaire ou au secondaire sur la performance.

Le PISA permet d'évaluer la performance scolaire des élèves de 15 ans en mathématiques, en lecture et en science. Il est notamment demandé aux élèves répondants s'ils ont redoublé au moins une année au primaire et au moins une année au secondaire. Pour le reste de la section, nous considérerons le redoublement ayant eu lieu au primaire ou au secondaire sans faire de distinction. Toutefois, il sera important de tenir compte de cette différence dans les analyses à venir et de vérifier si les résultats persistent, étant donné que le PISA vise à évaluer la performance des élèves au secondaire, soit les effets estimés du redoublement ayant eu lieu au primaire sur la performance scolaire à long terme.

Le tableau 27 présente les régressions par moindres carrés ordinaires du résultat dans chaque matière en ce qui concerne la reprise d'au moins une année, avec un conditionnement selon un ensemble de caractéristiques liées à l'élève, à ses parents ou à

son école. La première colonne affiche la variable dépendante utilisée dans l'estimation, qui varie selon la matière considérée. La deuxième colonne indique le numéro de l'estimation, qui varie selon l'ensemble des variables incluses. Les trois colonnes suivantes présentent les coefficients estimés en ce qui a trait, respectivement, au fait d'avoir redoublé au moins une année, au fait d'être en quatrième secondaire ou non et au niveau scolaire de l'élève. La sixième colonne contient le ou les ensembles de variables utilisés dans la régression, qui sont précisés dans la dernière ligne du tableau. À travers les régressions, les variables de contrôle employées pour l'école sont la province, l'année, la taille de la ville où l'école est située, son statut privé ou non et le rapport élèves-enseignant. Les variables de contrôle utilisées pour l'élève et ses parents sont le sexe, le statut d'immigration, la langue parlée à la maison, le nombre de livres à la maison, le quartile de statut socioéconomique, l'éducation du père et de la mère ainsi que la structure familiale. La dernière colonne indique le nombre d'observations utilisées dans la régression. En effet, certaines variables comportent davantage de valeurs manquantes, ce qui réduit la taille de l'échantillon lorsqu'on tente d'effectuer un contrôle pour ces facteurs.

Les résultats affichés dans le tableau 27 révèlent que le simple fait de redoubler est associé à une performance scolaire plus faible, autant en mathématiques qu'en lecture ou en science. Ce résultat persiste même lorsqu'un conditionnement est effectué pour un ensemble de caractéristiques socioéconomiques de l'élève et de caractéristiques liées à son école. Ainsi, entre les estimations 1 à 5, on remarque que le coefficient de la variable du redoublement diminue. Dans l'estimation 1, on effectue un conditionnement uniquement pour la province, le cycle, le redoublement et le fait d'être ou non en quatrième secondaire. Ensuite, dans l'estimation 2, on ajoute la taille de la ville où est située l'école et le statut de celle-ci (privé ou public). Puis, dans l'estimation 3, on ajoute le sexe de l'enfant, son statut d'immigration, la langue parlée à la maison, le nombre de livres à la maison, le quartile de statut socioéconomique ainsi que l'éducation du père et de la mère. Enfin, on inclut la structure familiale et le rapport élèves-enseignant (estimation 4). Finalement, dans la dernière estimation, on ajoute le niveau scolaire constaté de l'élève sous forme linéaire. En observant le changement du coefficient des estimations 1 à 5, on voit qu'il existe un biais de sélection. Par contre, même lorsque le conditionnement concerne un vaste ensemble de caractéristiques, y compris le niveau scolaire de l'élève (estimation 5), on remarque que le coefficient associé à la variable du redoublement demeure largement négatif et supérieur à 40 points dans le PISA, ce qui correspond à environ une année d'acquis scolaires. Ainsi, l'effet négatif du redoublement ne peut être expliqué seulement par le fait que les élèves qui en ont fait l'objet se situent à un niveau scolaire inférieur ou présentent des

caractéristiques sociodémographiques différentes. Ces constats sont valides dans les trois domaines évalués par le PISA.

Comme nous l'avons déjà mentionné, en 2000 et en 2006, l'évaluation du PISA ne contenait pas de questions portant sur le redoublement. Les données de ces cycles d'enquête ne peuvent donc pas être utilisées dans les estimations présentées dans le tableau 27. Cependant, il est possible d'utiliser le retard scolaire des élèves comme une approximation du redoublement pour tous les cycles d'enquête du PISA. Le calcul de ce retard nécessite seulement de connaître le niveau scolaire de l'élève et son mois de naissance. De là, ce calcul peut être effectué pour tous les cycles d'enquête du PISA.

Comme le tableau 27, le tableau 28 présente les résultats des régressions par moindres carrés ordinaires du score dans chaque matière du PISA en ce qui a trait à un retard d'au moins une année, avec un conditionnement pour un ensemble de caractéristiques liées à l'élève, à ses parents ou à son école. Les ensembles de variables utilisés sont les mêmes que dans les régressions concernant le redoublement qui se trouvent au tableau 27. Les provinces de Terre-Neuve-et-Labrador, de l'Île-du-Prince-Édouard, de la Saskatchewan et de l'Alberta ont toutefois été omises des régressions étant donné que leur taux de retard scolaire diffère significativement de leur taux de redoublement, comme l'illustre la section 4.

L'effet négatif du redoublement sur la performance scolaire indique seulement que les doubleurs réussissent moins bien aux évaluations, et ce, compte tenu de facteurs observables de la performance scolaire qui ont été inclus dans les estimations. Ces estimations pourraient être biaisées par un ensemble de facteurs inobservables qui motivent à la fois le redoublement d'un élève et sa performance. La méthode par moindres carrés ordinaires indique qu'il existe une relation négative entre le redoublement et la performance scolaire, mais ne permet pas de conclure à une relation causale.

Tableau 27 — Régressions par moindres carrés ordinaires des résultats du PISA pour le redoublement

Variable dépendante	Estimation	Redoublement	Quatrième secondaire	Niveau scolaire	Variables incluses	N
Score en mathématiques	1	-78,7*** (1,92)	29,2*** (1,63)		(a)	114 304
	2	-74,9*** (2,02)	28,7*** (1,72)		(a) + (b)	110 162
	3	-62,3*** (2,23)	24,3*** (1,69)		(a) + (b) + (c)	93 249
	4	-62,6*** (2,54)	25,4*** (2,14)		(a) + (b) + (c) + (d)	53 051
	5	-52,0*** (2,19)			33,7*** (1,45) (b) + (c) + (d) + (e)	93 249
Score en lecture	1	-78,2*** (2,17)	28,1*** (1,62)		(a)	114 304
	2	-75,0*** (2,13)	27,2*** (1,62)		(a) + (b)	110 162
	3	-58,8*** (2,19)	19,4*** (1,53)		(a) + (b) + (c)	93 249
	4	-59,4*** (2,58)	19,0*** (1,93)		(a) + (b) + (c) + (d)	53 051
	5	-49,3*** (2,10)			28,6*** (1,23) (b) + (c) + (d) + (e)	93 249
Score en science	1	-77,4*** (2,21)	28,8*** (1,71)		(a)	114 304
	2	-74,5*** (2,28)	27,9*** (1,71)		(a) + (b)	110 162
	3	-59,7*** (2,41)	22,2*** (1,80)		(a) + (b) + (c)	93 249
	4	-61,7*** (2,68)	22,1*** (2,25)		(a) + (b) + (c) + (d)	53 051
	5	-49,2*** (2,33)			32,3*** (1,43) (b) + (c) + (d) + (e)	93 249

Note : Les erreurs types sont affichées entre parenthèses. * p < 0,05, ** p < 0,01 et *** p < 0,001.

Ensemble de variables incluses :

- a) Province, cycle, redoublement, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère.
- d) Structure familiale, rapport élèves-enseignant.
- e) Province, cycle, redoublement, niveau de scolarité sous forme linéaire (année scolaire).

Tableau 28 — Régressions par moindres carrés ordinaires des résultats du PISA pour le retard scolaire

Variable dépendante	Estimation	Retard scolaire	Quatrième secondaire	Niveau scolaire	Variables incluses	N
Score en mathématiques	1	-73,8*** (2,70)	16,6*** (1,94)		(a)	106 476
	2	-71,6*** (2,67)	15,4*** (2,08)		(a) + (b)	91 558
	3	-57,2*** (2,55)	16,1*** (1,76)		(a) + (b) + (c)	77 472
	4	-63,1*** (3,17)	15,4*** (2,53)		(a) + (b) + (c) + (d)	35 687
	5	-37,3*** (2,96)			31,9*** (2,10) (b) + (c) + (d) + (e)	77 472
Score en lecture	1	-79,9*** (2,36)	14,7*** (1,75)		(a)	115 294
	2	-76,5*** (2,59)	12,8*** (1,88)		(a) + (b)	91 558
	3	-56,2*** (2,44)	12,7*** (1,70)		(a) + (b) + (c)	77 472
	4	-59,2*** (3,29)	9,6*** (2,17)		(a) + (b) + (c) + (d)	35 687
	5	-39,2*** (3,00)			26,3*** (1,70) (b) + (c) + (d) + (e)	77 472
Score en science	1	-78,4*** (2,66)	13,2*** (1,85)		(a)	106 514
	2	-76,0*** (2,80)	12,1*** (1,97)		(a) + (b)	91 558
	3	-58,9*** (2,59)	12,5*** (1,90)		(a) + (b) + (c)	77 472
	4	-65,2*** (3,43)	10,8*** (2,63)		(a) + (b) + (c) + (d)	35 687
	5	-39,1*** (2,91)			28,9*** (1,77) (b) + (c) + (d) + (e)	77 472

Note : Les erreurs types sont affichées entre parenthèses. Les provinces incluses sont la Nouvelle-Écosse, le Nouveau-Brunswick, le Québec, l'Ontario, le Manitoba et la Colombie-Britannique. * $p < 0,05$, ** $p < 0,01$ et *** $p < 0,001$.

Ensemble de variables incluses :

- a) Province, cycle, redoublement, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère.
- d) Structure familiale, rapport élèves-enseignant.
- e) Province, cycle, redoublement, niveau de scolarité sous forme linéaire (année scolaire).

6.3.2 Estimations sous forme réduite

Comme il a été indiqué précédemment, des auteurs ont observé que les élèves nés dans le dernier trimestre de l'année risquent davantage de redoubler (Diris, 2017; Pedraja-Chaparro, Santin et Simancas, 2015). Si l'hypothèse selon laquelle le mois de naissance n'influence pas directement la performance scolaire est acceptée, cet indicateur pourrait être un instrument valide pour le redoublement. Il s'agit de l'approche empruntée par Diris et qui consiste en une régression du résultat scolaire selon le fait d'avoir redoublé ou non, au moyen d'un instrument qui indique si l'élève est né dans le dernier trimestre de l'année. Toutefois, étant donné que l'entrée à l'école varie d'une province à l'autre, l'instrument que nous utilisons indique plutôt si l'élève est né dans le dernier trimestre précédant la date d'admissibilité à l'école. Si cet instrument est valide, le modèle des variables instrumentales offre une meilleure estimation de l'effet du redoublement en réduisant le biais de sélection observé dans les estimations par moindres carrés ordinaires.

Pour considérer la relation entre la performance scolaire et le fait d'être né dans le dernier trimestre précédant la date d'admissibilité à l'école, on peut estimer par moindres carrés ordinaires l'effet de l'instrument sur la performance scolaire. Autrement dit, cette approche consiste à reproduire les estimations de la section précédente en incluant cette fois la variable candidate comme instrument. Le tableau 29 présente les résultats des régressions sous forme réduite du score à l'évaluation du PISA selon le fait d'avoir redoublé ou non, le fait d'être né ou non dans le dernier trimestre précédant la date d'admissibilité à l'école et les autres variables de contrôle. Bien que la relation soit significative à un seuil de 0,1 % dans les estimations 1 et 2 en mathématiques et en science, on peut noter que l'effet de l'instrument sur la performance scolaire n'est pas significatif lorsqu'on tient compte des caractéristiques socioéconomiques de l'élève (c) dans les estimations 3, 4 et 5 en mathématiques et lorsqu'on tient compte de la structure familiale et du rapport élèves-enseignant (d) dans les estimations 4 et 5 en science. Ainsi, l'instrument n'a pas une influence directe sur les notes une fois qu'on prend en compte les caractéristiques sociodémographiques.

Dans le même sens, on peut aussi étudier la relation entre l'instrument et la performance à l'évaluation du PISA en incluant le retard scolaire plutôt que le taux de redoublement. Le tableau 30 contient les résultats des régressions sous forme réduite du score à l'évaluation du PISA selon le retard scolaire, l'instrument et les autres variables de contrôle. Parmi ces résultats, la seule estimation présentant un coefficient significatif à un seuil de 5 % lié à l'instrument est la troisième pour la performance en mathématiques. Toutes les autres

estimations rapportent un coefficient lié à l'instrument qui n'est pas significatif à un seuil de 5 %.

Tableau 29 — Régressions sous forme réduite des résultats du PISA pour le redoublement
Instrument : Naissance dans le dernier trimestre

Variable dépendante	Estimation	Redoublement	Naissance au dernier trimestre	Variables incluses	N
Score en mathématiques	1	-78,5*** (1,93)	-2,4** (0,92)	(a)	114 304
	2	-74,6*** (2,02)	-2,52** (0,88)	(a) + (b)	110 162
	3	-62,2*** (2,23)	-1,8 (0,96)	(a) + (b) + (c)	93 249
	4	-62,4*** (2,54)	-1,8 (1,28)	(a) + (b) + (c) + (d)	53 051
	5	-51,9*** (2,18)	-1,5 (0,95)	(b) + (c) + (d) + (e)	93 249
Score en lecture	1	-78,1*** (2,18)	-1,3 (1,16)	(a)	114 304
	2	-74,9*** (2,14)	-1,4 (1,15)	(a) + (b)	110 162
	3	-58,6*** (2,20)	-1,59 (1,22)	(a) + (b) + (c)	93 249
	4	-59,3*** (2,58)	-1,1 (1,19)	(a) + (b) + (c) + (d)	53 051
	5	-49,2*** (2,11)	-1,3 (1,22)	(b) + (c) + (d) + (e)	93 249
Score en science	1	-77,2*** (2,21)	-2,6** (0,90)	(a)	114 304
	2	-74,2*** (2,27)	-2,7** (0,92)	(a) + (b)	110 162
	3	-59,5*** (2,40)	-2,0* (1,0)	(a) + (b) + (c)	93 249
	4	-61,5*** (2,67)	-1,4 (1,45)	(a) + (b) + (c) + (d)	53 051
	5	-49,1*** (2,32)	-1,7 (1,03)	(b) + (c) + (d) + (e)	93 249

Note : Les erreurs types sont affichées entre parenthèses. * p < 0,05, ** p < 0,01 et *** p < 0,001.

Ensemble de variables incluses :

- a) Province, cycle, redoublement, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère.
- d) Structure familiale, rapport élèves-enseignant.
- e) Province, cycle, redoublement, niveau de scolarité sous forme linéaire (année scolaire).

Tableau 30 — Régressions sous forme réduite des résultats du PISA pour le retard scolaire

Instrument : Naissance dans le dernier trimestre

Variable dépendante	Estimation	Retard scolaire	Naissance au dernier trimestre	Variables incluses	N
Score en mathématiques	1	-73,9*** (2,72)	0,2 (0,92)	(a)	106 476
	2	-71,3*** (2,70)	-0,6 (0,96)	(a) + (b)	91 558
	3	-57,2*** (2,61)	0,01 (1,06)	(a) + (b) + (c)	77 472
	4	-63,5*** (3,25)	0,9 (1,44)	(a) + (b) + (c) + (d)	35 687
	5	-36,6*** (3,05)	-1,5 (1,08)	(b) + (c) + (d) + (e)	77 472
Score en lecture	1	-80,5*** (2,46)	1,4 (1,08)	(a)	115 294
	2	-76,8*** (2,70)	0,8 (1,25)	(a) + (b)	91 558
	3	-56,2*** (2,64)	0,2 (1,38)	(a) + (b) + (c)	77 472
	4	-59,6*** (3,43)	0,8 (1,46)	(a) + (b) + (c) + (d)	35 687
	5	-38,7*** (3,23)	-1,2 (1,40)	(b) + (c) + (d) + (e)	77 472
Score en science	1	-78,3*** (2,69)	-0,1 (0,94)	(a)	106 514
	2	-75,8*** (2,84)	-0,5 (1,04)	(a) + (b)	91 558
	3	-58,7*** (2,63)	-0,3 (1,16)	(a) + (b) + (c)	77 472
	4	-65,9*** (3,58)	1,54 (1,83)	(a) + (b) + (c) + (d)	35 687
	5	-38,3*** (2,99)	-1,9 (1,18)	(b) + (c) + (d) + (e)	77 472

Note : Les erreurs types sont affichées entre parenthèses. Les provinces incluses sont la Nouvelle-Écosse, le Nouveau-Brunswick, le Québec, l'Ontario, le Manitoba et la Colombie-Britannique. * p < 0,05, ** p < 0,01 et *** p < 0,001.

Ensemble de variables incluses :

- a) Province, cycle, redoublement, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère.
- d) Structure familiale, rapport élèves-enseignant./professeur
- e) Province, cycle, redoublement, niveau de scolarité sous forme linéaire (année académique)

6.3.3 Résultats des estimations par variable instrumentale

À partir des résultats dans la section précédente, on peut penser que l'hypothèse de l'exogénéité de l'instrument est plausible. Des études ont aussi démontré que les élèves les plus jeunes d'une cohorte ont une probabilité plus élevée de redoubler (Diris, 2017; Pedraja-Chaparro, Santin et Simancas, 2015). Ce résultat suggère que le fait d'être né dans le trimestre précédant la date d'admissibilité à l'école peut constituer un instrument fort. Pour ces raisons, l'effet du redoublement sur la performance aux tests du PISA est estimé en utilisant le trimestre de naissance comme instrument.

Le tableau 31 présente les résultats des régressions par variables instrumentales des scores dans chaque matière du PISA selon le fait d'avoir redoublé au moins une année. Les ensembles de variables utilisés dans ces régressions sont les mêmes que ceux employés dans les sections précédentes avec les données du PISA. Les résultats de ces régressions indiquent que le redoublement entraîne des répercussions négatives importantes sur la performance scolaire. Toutefois, l'effet sur la performance en lecture n'est pas significatif lorsqu'on inclut le niveau scolaire en tant que variable continue dans la régression. Plus particulièrement, le fait de redoubler réduit le score à l'évaluation d'environ 100 points ou plus dans toutes les matières, selon l'ensemble de variables de contrôle utilisé. Ces résultats sont assez surprenants si l'on considère le fait qu'une différence de 100 points dans la performance en mathématiques et en lecture au Canada en 2018 correspondait à un retard d'apprentissage de plus de 32 mois. En science, la différence était de 70 points au Canada en 2018, ce qui correspond à un retard d'apprentissage de 32 mois. On peut remarquer que nos estimations manquent de précision, les écarts-types étant assez larges.

En utilisant le retard scolaire plutôt que le redoublement, il est possible d'inclure les données de 2000 et de 2006 du PISA. Par contre, les données des provinces de Terre-Neuve-et-Labrador, de l'Île-du-Prince-Édouard, de la Saskatchewan et de l'Alberta doivent être exclues, puisque leur taux de retard scolaire diffère significativement de leur taux de redoublement. Le tableau 32 présente les résultats des régressions par variables instrumentales des scores dans chaque matière du PISA pour un retard d'au moins une année. Ces résultats soutiennent l'hypothèse que les élèves accusant un retard ont une performance plus faible à l'évaluation du PISA. Ce constat est établi peu importe le domaine évalué. Selon ces estimations, il semblerait que le retard scolaire entraîne une diminution significative de la performance des élèves en mathématiques, en lecture et en science.

Tableau 31 — Régressions par variables instrumentales des résultats du PISA pour le redoublement

Instrument : Naissance dans le dernier trimestre

Variable dépendante	Estimation	Redoublement	Quatrième secondaire	Niveau scolaire	Variables incluses	N
Score en mathématiques	1	-175,3*** (36,78)	-3,0 (12,20)		(a)	114 304
	2	-177,1** (36,08)	-5,2 (11,94)		(a) + (b)	110 162
	3	-140,5*** (40,76)	0,2 (12,56)		(a) + (b) + (c)	93 249
	4	-125,1** (42,63)	4,1 (14,8)		(a) + (b) + (c) + (d)	53 051
	5	-132,2* (52,32)		10,4 (15,24)	(b) + (c) + (d) + (e)	93 249
Score en lecture	1	-130,5** (45,4)	10,6 (15,21)		(a)	114 304
	2	-134,7*** (46,11)	7,4 (15,36)		(a) + (b)	110 162
	3	-128,1* (51,86)	-2,1 (16,06)		(a) + (b) + (c)	93 249
	4	-99,6* (39,70)	5,35 (13,82)		(a) + (b) + (c) + (d)	53 051
	5	-122,1 (66,76)		7,4 (19,38)	(b) + (c) + (d) + (e)	93 249
Score en science	1	-182,4*** (36,28)	-6,1 (12,33)		(a)	114 304
	2	-182,43*** (37,91)	-7,93 (12,82)		(a) + (b)	110 162
	3	-147,8*** (44,05)	-5,1 (13,83)		(a) + (b) + (c)	93 249
	4	-110,21* (49,12)	5,6 (16,96)		(a) + (b) + (c) + (d)	53 051
	5	-144,0* (56,77)		4,7 (16,69)	(b) + (c) + (d) + (e)	93 249

Note : Les erreurs types sont affichées entre parenthèses. * p < 0,05, ** p < 0,01 et *** p < 0,001.

Ensemble de variables incluses :

- a) Province, cycle, redoublement, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère.
- d) Structure familiale, rapport élèves-enseignant.
- e) Province, cycle, redoublement, niveau de scolarité sous forme linéaire (année scolaire).

Tableau 32 — Régressions par variables instrumentales des résultats du PISA pour le retard scolaire

Instrument : Naissance dans le dernier trimestre

Variable dépendante	Estimation	Retard scolaire	Quatrième secondaire	Année scolaire	Variables incluses	N
Score en mathématiques	1	-71,8*** (11,51)	17,8* (6,93)		(a)	106 476
	2	-79,2*** (12,08)	11,0 (7,13)		(a) + (b)	91 558
	3	-57,0*** (13,86)	16,2* (7,31)		(a) + (b) + (c)	77 472
	4	-52,8** (16,86)	20,9* (9,57)		(a) + (b) + (c) + (d)	35 687
	5	-58,3*** (14,69)		22,0** (6,88)	(b) + (c) + (d) + (e)	77 472
Score en lecture	1	-63,3*** (12,94)	24,6** (7,91)		(a)	115 294
	2	-67,0*** (15,38)	18,3* (9,03)		(a) + (b)	91 558
	3	-54,1** (17,51)	13,8 (9,29)		(a) + (b) + (c)	77 472
	4	-50,0** (16,80)	14,7 (9,69)		(a) + (b) + (c) + (d)	35 687
	5	-55,2** (18,61)		18,8* (8,76)	(b) + (c) + (d) + (e)	77 472
Score en science	1	-80,11*** (11,74)	12,22 (7,039)		(a)	106 514
	2	-82,73*** (13,16)	8,219 (7,736)		(a) + (b)	91 558
	3	-63,17*** (15,36)	10,29 (8,165)		(a) + (b) + (c)	77 472
	4	-47,41* (21,17)	20,44 (11,89)		(a) + (b) + (c) + (d)	35 687
	5	-65,24*** (16,30)		16,60* (7,657)	(b) + (c) + (d) + (e)	77 472

Note : Les erreurs types sont affichées entre parenthèses. Les provinces incluses sont la Nouvelle-Écosse, le Nouveau-Brunswick, le Québec, l'Ontario, le Manitoba et la Colombie-Britannique. * p < 0,05, ** p < 0,01 et *** p < 0,001.

Ensemble de variables incluses :

- a) Province, cycle, retard scolaire, quatrième secondaire.
- b) Taille de la ville où est située l'école, école privée.
- c) Sexe, statut d'immigration de l'enfant, langue parlée à la maison, nombre de livres à la maison, quartile de statut socioéconomique, éducation du père et de la mère
- d) Structure familiale, rapport élèves-enseignant.
- e) Province, cycle, retard scolaire, niveau de scolarité sous forme linéaire (année scolaire).

6.3.4 Changement de régime : méthode des doubles différences

Dans cette section, nous exploitons le changement de régime observé au Québec en 2008³⁸. Comme nous l'avons vu dans la section 4, les taux moyens de reprise d'une année changent à partir de 2008-2009, particulièrement au primaire. Étant donné ce changement, il est possible d'utiliser une approche de différence en différence, puisque le taux de redoublement a largement diminué après 2008 suivant une modification des pratiques au Québec. Puisque le domaine principal d'évaluation du PISA était la lecture en 2000 et les mathématiques en 2003, et que la science n'a été évaluée en détail qu'en 2006, nous menons l'analyse uniquement pour la lecture et les mathématiques, les domaines couverts à au moins deux reprises avant le changement de régime au Québec. La stratégie adoptée est relativement simple : nous comparons l'évolution des scores au Québec avec celle observée ailleurs au Canada pour déterminer si le changement de régime a influencé les résultats des jeunes. Par contre, pour que cette stratégie fonctionne, il est essentiel que les pratiques de redoublement soient demeurées relativement stables dans les autres provinces. Nous remarquons cependant que les pratiques de redoublement ont aussi évoluées ailleurs au Canada (tableau 14). Ainsi, pour nous assurer d'avoir un groupe témoin stable, nous conservons uniquement les provinces du Québec (groupe traité), de l'Ontario, de la Saskatchewan et de la Colombie-Britannique (groupe témoin) d'abord. Puis, nous concentrons notre attention sur le Québec et l'Ontario exclusivement.

Le modèle suivant est estimé :

$$Y_{ipc} = \alpha + \beta^{DID} * POST_t * QC_i + \theta X_i + \gamma_p + \delta_c + u_{ipc}$$

où Y_{ipc} est le score de l'étudiant i vivant dans la province p durant le cycle d'enquête c . La variable dichotomique $POST_t$ est égale à 1 si l'élève est observé après 2006 et la variable dichotomique QC_i est aussi égale à 1 si l'élève vit au Québec. Nous incluons des effets fixes de province et d'année d'enquête. Puisque notre modèle comprend des effets fixes d'année d'enquête, il peut capter les changements dans le temps, communs à toutes les provinces. Ainsi, le coefficient β^{DID} capte l'effet de la réforme, à supposer que l'hypothèse des tendances communes soit respectée. Nous incluons aussi un ensemble de contrôles sociodémographiques, soit en ce qui concerne le niveau scolaire, la province, le cycle d'enquête, le statut d'immigration, la langue parlée à la maison, le nombre de livres à la maison, le statut socioéconomique, le niveau d'éducation des parents, l'emplacement de

³⁸ Source : <https://ici.radio-canada.ca/tele/entree-principale/2016-2017/segments/chronique/15823/redoubler-annee-scolaire>.

l'école (ville ou autre) et le statut de celle-ci (privé ou public). Les variables manquantes n'ont pas été imputées.

Nos résultats suggèrent que la baisse du taux de redoublement observée au Québec n'est pas associée à une diminution de la performance des élèves. Le coefficient β^{DID} , qui capte l'effet de la réforme sous certaines conditions, n'est pas différent de zéro en lecture, tant pour les garçons que pour les filles. De plus, le choix du groupe témoin ne change pas ce résultat. Ainsi, il semblerait que la performance des élèves en lecture n'ait pas été affectée par la réforme en moyenne. En mathématiques, le portrait est différent. Quel que soit le groupe témoin, la performance des élèves du Québec s'est améliorée après la réforme, passant de 8,4 à 8,6 points selon le PISA. L'amélioration semble exclusivement concentrée chez les garçons. La performance de ces derniers est ainsi passée de 12,3 à 14,7 points en mathématiques après la réforme. On se rappellera que, dans le PISA, 40 points représentent environ une année. Ainsi, on parle d'une amélioration de près de 25 % d'une année d'apprentissage. Bien entendu, il est possible que d'autres changements spécifiques au Québec se soient produits en même temps que la réforme portant sur le redoublement. Cependant, ces résultats laissent tout de même présager que le changement de régime n'a pas eu d'effets négatifs sur la performance des élèves et pourrait même avoir été bénéfique. En résumé, ces résultats apportent des évidences empiriques supplémentaires suggérant que le redoublement n'est peut-être pas la pratique la plus efficace pour favoriser la réussite des jeunes.

7 Analyse coût-avantage

Les résultats de cette étude montrent que le redoublement, appliqué comme mesure de soutien à la réussite scolaire des élèves en difficulté au Québec, ne produit pas les effets escomptés. En réalité, pour une majorité d'élèves, la mesure s'avère préjudiciable aux notes scolaires et à l'obtention d'un diplôme aux études secondaires tout en incitant un certain nombre à abandonner leur scolarité. Cette section vise à quantifier les effets socioéconomiques du redoublement à partir des estimations produites dans cette étude.

Pour quantifier ces effets, il est important de comprendre les conséquences du redoublement non seulement à court terme, mais aussi à moyen et à long terme. En effet, au-delà des conséquences à court terme sur le parcours éducatif, le redoublement peut également mener à des conséquences négatives à long terme tant sur le plan économique que sur le plan social. Par exemple, puisque les individus ayant connu un ou plusieurs épisodes de redoublement durant leur scolarité risquent davantage de quitter l'école sans diplôme ni qualification, il y a fort à parier qu'ils recevront aussi de plus faibles revenus

d'emploi que les autres à l'âge adulte (Segal, 2013). Cette baisse des revenus affecte l'individu lui-même, mais se répercute aussi négativement sur la richesse collective et les rentrées fiscales pour le financement des services publics. De plus, il est probable que les personnes ayant redoublé demandent plus souvent, à moyen et à long terme, l'aide financière de dernier recours, puisque les adultes sans diplôme ni qualification éprouvent de plus grandes difficultés à obtenir et à conserver un emploi sur le marché du travail que les autres. Bref, le redoublement occasionne des coûts socioéconomiques individuels et collectifs qui s'ajoutent au coût financier de la mesure présenté dans l'introduction de cette étude.

L'analyse coût-avantage de cette section s'appuie sur des calculs comptables et statistiques des avantages socioéconomiques de la formation qualifiante. La modélisation est effectuée au moyen des données de l'enquête Relance produite annuellement par le MEQ depuis 2001. Elle s'appuie également sur les taux d'imposition moyens par tranche de revenus au Québec et les effectifs de l'aide de dernier recours par palier de scolarisation. L'enquête Relance contient des renseignements sur le revenu et le statut d'emploi des diplômés pour l'année suivant l'obtention du diplôme. Des renseignements sont ainsi disponibles pour les diplômés des formations qualifiantes suivantes : diplôme d'études professionnelles (DEP), attestation de spécialisation professionnelle (ASP), diplôme d'études collégiales (DEC) de la formation technique, attestation d'études collégiales (AEC), baccalauréat et maîtrise.

Le modèle d'analyse coût-avantage permet d'estimer les gains liés à la diplomation pour ce qui concerne les revenus et les taux d'emploi, les recettes fiscales pour le gouvernement du Québec et les dépenses relatives à l'aide de dernier recours. L'estimation des gains pour un diplôme en particulier repose sur une comparaison du revenu moyen d'une cohorte de diplômés et d'un revenu contrefactuel si le diplôme n'avait pas été obtenu. Il est présumé que le revenu contrefactuel des diplômés est celui qu'ils auraient reçu avec un diplôme requérant moins d'années de scolarité. Par exemple, le revenu contrefactuel des titulaires d'un baccalauréat correspond au revenu des détenteurs d'un diplôme d'études collégiales de la formation technique, à supposer que cette formation a été suivie par la totalité des bacheliers. Le tableau 33 montre la composition des facteurs contrefactuels pour chaque palier de diplomation. Ainsi, le facteur contrefactuel des titulaires d'un DEP est composé de 35,0 % d'individus sans diplôme ni qualification, de 32,5 % de détenteurs d'une ASP et de 32,5 % de titulaires d'un DEC de la formation technique. De la même façon, le facteur contrefactuel des titulaires d'une maîtrise regroupe uniquement des bacheliers.

La comparaison des indicateurs (revenu, recettes fiscales, dépenses relatives à l'aide de dernier recours) peut être effectuée pour un groupe d'âge particulier en ce qui concerne les

diplômés de l'enquête Relance afin de limiter les biais statistiques dans le calcul des avantages. Par exemple, une comparaison des revenus des titulaires d'un DEP âgés de 15 à 24 ans et du revenu contrefactuel de personnes du même âge réduit l'effet de l'expérience de travail dans le calcul des écarts. Par contre, cette approche repose sur le postulat que les différences de revenu liées à l'expérience de travail sont identiques, peu importe l'âge des diplômés. Or, nous savons que le rendement de l'éducation augmente avec l'âge, particulièrement au regard de l'enseignement universitaire (Statistique Canada, *Education Indicators in Canada: Report of the Pan-Canadian Education Indicators Program*, novembre 2021, catalogue n° 81 582 X).

Tableau 33 — *Groupes d'âges et de diplômés pour la formation des groupes témoins*

Diplôme	Groupe d'âge	Groupe témoin
DEP	15 à 24 ans	Sans diplôme (35%), ASP (32,5%), DECT (32,5%)
ASP	25 à 34 ans	Sans diplôme (27%), DEP (73%)
Collégial technique	15 à 24 ans	DEP (60%), ASP (20%), AEC(20%)
AEC	25 à 34 ans	DEP (34%), ASP (33%), DECT (33%)
Baccalauréat	Tous les diplômés	DECT (100%)
Maîtrise	Tous les diplômés	Baccalauréat (100%)
Doctorat	Tous les diplômés	Maîtrise (100%)

Le modèle d'analyse C-B rapporte les gains à long terme observés pour les diplômés et la collectivité, c'est-à-dire les avantages retirés par les diplômés au cours de leur vie active sur le marché du travail. Ce modèle tient compte également des taux d'emploi pour chacun des secteurs de diplomation. Les rentrées fiscales pour le gouvernement provincial ont été estimées au moyen des taux moyens d'imposition par tranche de revenus. De plus, le modèle calcule les coûts unitaires pour chacun des programmes de formation qualifiante, les coûts d'opportunité pour l'élève (pertes de revenu durant la scolarisation) et ceux liés au matériel scolaire. Les coûts et les avantages sont rapportés dans un tableau de type Kaldor-Hicks.

Le calcul des pertes socioéconomiques et des gains liés au redoublement scolaire porte sur la cohorte des élèves de la première année du primaire en 2006-2007. Une première étape établit un compte des élèves touchés par la mesure selon les effets estimés dans cette étude. L'attention est portée sur les élèves sans diplôme ni qualification et ceux ayant

obtenu un CFMS en raison de la reprise d'une année³⁹. Les chiffres sont rapportés au tableau 34. La colonne 1 du tableau indique les effets qui concernent la non-diplomation et la colonne 2, ceux qui sont liés à l'obtention d'un CFMS. Les effets estimés du

Tableau 34 — *Calcul du nombre d'élèves touchés par le redoublement scolaire (non-diplomation : FGJ, FGA et autres; diplomation : CFMS), cohorte de la première année du primaire en 2006-2007*

<i>Paramètres</i>	(1) <i>Non diplomation</i>	(2) <i>Diplomation CFMS</i>
1. Doubleurs (n) : 20 218		
<i>Impact moyen sur les participants</i>		
2. Points de pourcentage	24,6	17,2
3. Nombre d'élèves (ligne 1. x ligne 2.)	4 974	3 477

redoublement qui sont affichés dans le tableau sont ceux produits par le modèle de probit ordonné. La ligne 2 indique 24,6 pp pour l'effet du redoublement sur la probabilité de ne pas obtenir un diplôme. Le nombre de non-diplômés qui découle de la reprise scolaire est estimé à 4 974. Le même exercice comptable établit à 3 477 le nombre d'élèves qui ont obtenu un CFMS en raison de la reprise d'une année.

Un premier calcul des coûts socioéconomiques porte sur les doubleurs qui ont décroché du secondaire ou qui n'ont pas obtenu de diplôme qualifiant en raison du redoublement. Pour évaluer ces coûts, il est nécessaire de spécifier le diplôme qu'auraient obtenu ces élèves s'ils n'avaient pas redoublé. Dans ce cas-ci, nous posons comme hypothèse que la totalité des non-diplômés auraient obtenu un diplôme d'études professionnelles compte tenu du fait que ce diplôme est plus facilement accessible pour ces jeunes et qu'il est lié à une baisse de certains coûts (coûts d'opportunité, matériel scolaire)⁴⁰.

L'analyse coût-avantage est présentée au tableau 35 conformément à une comptabilité Kaldor-Hicks. Les calculs portent sur une période de vie active fixée à 40 années. La colonne 1 du tableau 35 indique que les pertes à long terme au regard des revenus d'emploi

³⁹ Les résultats de la section 6 indiquent que le redoublement scolaire est lié à une baisse de la diplomation aux études secondaires, de sorte qu'il est présumé que les élèves touchés par la mesure quittent l'école sans diplôme ni qualification ou optent pour une formation et un diplôme de substitution. Par ailleurs, aucune estimation des pertes socioéconomiques associées à la performance aux examens uniques du MEQ n'a été produite en raison du manque de renseignements pour cet exercice.

⁴⁰ L'enquête Relance ne recense pas le statut d'emploi et les salaires des diplômés ayant terminé une formation menant à un métier semi-spécialisé.

des décrocheurs se chiffrent à 2,0 milliards de dollars une fois prises en considération les réductions d'impôt sur le revenu (216,0 millions de dollars) et les gains d'opportunité au regard des revenus (45,8 millions de dollars). Les pertes fiscales sont estimées à 332,2 millions de dollars pour le gouvernement du Québec. À ce montant s'ajoute le coût de la formation des décrocheurs durant tout leur parcours au primaire et au secondaire, chiffré à 381,4 millions de dollars⁴¹. Au total, le coût à long terme du redoublement scolaire pour le gouvernement s'établit à 654,7 millions de dollars compte tenu du coût direct de la mesure, à savoir 57,2 millions de dollars pour la cohorte 2006-2007.

De plus, plusieurs décrocheurs demanderont, à un moment ou l'autre, une aide financière de dernier recours. À ce titre, les dépenses additionnelles pour la société québécoise sont estimées à 669,1 millions de dollars, toujours selon une approche liée à un facteur contrefactuel⁴². Des actes criminels pourraient également être commis durant le parcours de vie des non-diplômés. Le coût social de ces actes se chiffre à 51,7 millions de dollars. Dès lors, les pertes financières, économiques et sociales totalisent 3,4 milliards de dollars seulement pour les élèves de la cohorte 2006-2007 n'ayant pas reçu de diplôme qualifiant en raison du redoublement scolaire.

⁴¹ Le calcul est établi par la multiplication de la dépense par élève (12 411 \$), du nombre de décrocheurs (3 073) et du nombre d'années de scolarité avant que l'élève quitte l'école sans diplôme ni qualification (10).

⁴² Le ministère du Travail, de l'Emploi et de la Solidarité sociale publie le nombre de prestataires de l'aide financière de dernier recours et la durée moyenne de celle-ci par palier de diplomation. Le calcul des dépenses supplémentaires pour une aide de dernier recours fournie aux décrocheurs ayant redoublé s'appuie précisément sur ces renseignements.

Tableau 35 – Analyse coût-avantage du redoublement scolaire pour les élèves inscrits à la première année du primaire en 2006-2007 et ayant quitté l'école sans diplôme ni qualification en raison du redoublement scolaire, tableau Kaldor-Hicks

	Décrocheurs	G. Québec	Société (dépenses)	Total
	(1)	(2)	(3)	(4)
Bénéfices				
Économique (revenu)	-2 279 502 543 \$			-2 279 502 543 \$
Sociaux				
Aide dernier recours			669 102 480 \$	669 102 480 \$
Criminalité			83 750 246 \$	83 750 246 \$
Transferts				
Impôt provincial	216 117 508 \$	-216 117 508 \$		
Coûts				
Coût de formation PPS		343 251 027 \$		343 251 027 \$
Coût d'opportunité	-45 823 984 \$			-45 823 984 \$
Matériels scolaires	-2 361 956 \$			-2 361 956 \$
Redoublement scolaire		57 208 505 \$		57 208 505 \$
Bénéfices nets				
Absolues	-2 015 199 095 \$	-616 577 040 \$	752 852 726 \$	-3 384 628 861 \$

Source : Calculs des auteurs à partir de données du MEQ, de l'enquête Relance, des taux moyens d'imposition et des données colligées sur les sites du ministère du Travail, de l'Emploi et de la Solidarité sociale ainsi que du ministère de la Sécurité publique.

Le portrait coût-avantage diffère toutefois pour les 3 477 élèves qui ont obtenu un certificat de formation à un métier semi-spécialisé en raison du redoublement scolaire. Le tableau 36 rapporte des gains économiques à long terme estimés à 1,4 milliard de dollars dans le cas de ces élèves⁴³. Pour le gouvernement, l'investissement s'avère toutefois déficitaire en ce qui concerne les recettes fiscales. Cet investissement conduit néanmoins à une diminution de 465 millions de dollars des dépenses de programmes, compte tenu principalement d'un nombre moins élevé d'individus qui doivent demander une aide financière de dernier recours. Au total, les gains socioéconomiques des travailleurs, au

⁴³ L'enquête Relance n'offre pas de renseignements sur les individus ayant obtenu uniquement un certificat de formation à un métier semi-spécialisé ou ceux qui ont trouvé un emploi à l'issue d'une formation préparatoire au marché du travail. Le calcul des gains économiques permet donc de présumer que le revenu d'emploi de ces travailleurs correspond à 70 % de celui des travailleurs qui ont obtenu un diplôme d'études professionnelles. De plus, il est supposé que la totalité des élèves abandonneraient leur scolarité si les formations préparatoires au travail et celles menant à un métier semi-spécialisé n'existaient pas.

terme de leur vie active, se chiffrent à 1,9 milliard de dollars du fait d'une reprise scolaire motivée par des perspectives de qualification à très court terme offertes par les formations menant à un métier semi-spécialisé.

Tableau 36 – *Analyse coût-avantage du redoublement scolaire pour les élèves inscrits à la première année du primaire en 2006-2007 et ayant terminé une formation préparatoire au travail ou ayant obtenu un certificat de formation à un métier semi-spécialisé en raison du redoublement scolaire, tableau Kaldor-Hicks*

	Individus (FPT, CFMS)	G. Québec	Société (\$ épargnés)	Total
	(1)	(2)	(3)	(4)
Bénéfices				
Économique (revenu)	1 593 452 019 \$			1 593 452 019 \$
Sociaux				
Aide dernier recours			413 379 960 \$	413 379 960 \$
Criminalité			51 741 959 \$	51 741 959 \$
Transferts				
Impôt provincial	-151 073 698 \$	151 073 698 \$		
Coûts				
Coût de formation (FPT, CFSP)		101 510 412 \$		101 510 412 \$
Coût d'opportunité	32 032 568 \$			32 032 568 \$
Matériels scolaires	1 651 090 \$			1 651 090 \$
Redoublement scolaire		68 881 050 \$		68 881 050 \$
Bénéfices nets				
Absolues	1 408 694 663 \$	-19 317 764 \$	465 121 919 \$	1 854 498 818 \$

Source : Calculs des auteurs à partir de données du MEQ, de l'enquête Relance, des taux moyens d'imposition et des données colligées sur les sites du ministère du Travail, de l'Emploi et de la Solidarité sociale ainsi que du ministère de la Sécurité publique.

9 Conclusion

Le cheminement éducatif des élèves éprouvant des difficultés d'apprentissage inattendues ou récurrentes est une préoccupation importante des intervenants en éducation au Québec. Pour appuyer ces élèves, le *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire* prévoit la prolongation de l'un des cycles du primaire ou du secondaire, ce qui leur permet de poursuivre l'acquisition des connaissances scolaires. Ainsi, le redoublement compris comme la prolongation des

apprentissages a lieu, en principe, l'année qui suit la fin d'un cycle à consolider. Dans les faits, les statistiques montrent que le redoublement scolaire se produit à tous les ordres d'enseignement, principalement depuis la réforme de 2008-2009, qui limite à un le nombre de reprises pour le primaire ou le secondaire.

Une analyse descriptive a montré que les élèves qui reprennent une année scolaire présentent des particularités socioéconomiques et des résultats scolaires forts différents de ceux des autres élèves. D'abord, ils obtiennent de moins bonnes notes aux épreuves uniques. Les statistiques descriptives présentées dans cette étude permettent en effet d'observer que les jeunes ayant redoublé une seule fois ont des résultats aux épreuves à peine plus élevés que la note de passage. Ceux qui ont redoublé plus d'une fois échouent, quant à eux, à la plupart des épreuves. De plus, les doubleurs sont beaucoup moins nombreux que les autres élèves à obtenir un diplôme d'études secondaires et ont une propension importante à cesser leurs études sans diplôme ni qualification. Une analyse des données du PISA a également indiqué qu'ils présentent de moins bons scores aux tests de mathématiques, de lecture et de science.

Tant les données administratives du ministère de l'Éducation que celles des enquêtes du PISA montrent que les doubleurs sont issus, en général, de familles moins nanties sur le plan socioéconomique. Les parents de ces élèves sont moins scolarisés et ont des revenus d'emploi moins élevés. Ces élèves sont aussi plus nombreux à vivre dans une famille monoparentale que les autres jeunes. En général, on trouve une moins grande quantité de livres à la maison dans le cas des doubleurs. En outre, la participation de ces derniers à des projets éducatifs est beaucoup moins fréquente tout comme leur fréquentation des écoles privées. Par ailleurs, les doubleurs d'âge préscolaire et ceux des communautés défavorisées de l'île de Montréal ont participé, dans une proportion plus importante à la maternelle 4 ans à demi-temps comparativement aux autres enfants. Enfin, les statistiques descriptives de cette étude ont montré un plus fort pourcentage d'élèves avec un handicap physique ou intellectuel dans le groupe des doubleurs.

Les écarts dans la réussite scolaire entre les doubleurs et les non-doubleurs peuvent ainsi s'expliquer par des différences socioéconomiques et des événements inattendus et délétères. Ils peuvent aussi découler du redoublement lui-même si cette mesure comporte des effets avérés sur le cheminement et le succès scolaires. Cette étude visait donc à déterminer si la reprise d'une ou de plusieurs années scolaires comporte des effets sur la réussite éducative dans un contexte où les élèves qui redoublent présentent d'emblée des différences scolaires et socioéconomiques importantes.

Une recension sommaire des écrits sur le redoublement scolaire nous apprend que cette mesure, de façon générale, nuit aux chances d'obtenir un diplôme chez les élèves concernés

et augmente leur propension à abandonner le secondaire. Des études ont permis de relever néanmoins une hausse de la fréquentation scolaire et une réduction du nombre de suspensions si le redoublement se produit à la fin du primaire. Du point de vue des effets psychologiques et comportementaux, plusieurs recherches notent une augmentation du nombre de troubles du comportement ou de l'attention de même que du stress et de l'anxiété. D'autres ont plutôt révélé un effet favorable sur l'image de soi de l'élève à court terme, mais cet effet se dissiperait rapidement selon certaines recherches.

La présente étude doit donc tenir compte de plusieurs difficultés non négligeables quand vient le temps d'évaluer les effets du redoublement scolaire. Le recours au redoublement comme mesure de soutien à la réussite scolaire s'explique par plusieurs facteurs, dont ceux cités plus haut. Pour déterminer ses effets, il faut tenir compte de ces autres facteurs influençant la réussite. De plus, le redoublement scolaire peut s'expliquer par des facteurs difficilement mesurables ou observables, notamment les aptitudes scolaires des élèves, leurs relations avec leurs camarades de classe et certains événements inattendus jugés défavorables à la réussite éducative. Les estimations concernant le redoublement scolaire peuvent donc être entachées d'erreurs statistiques si ces facteurs sont exclus des évaluations économétriques.

Deux sources de données ont été mobilisées afin de minimiser la portée des biais statistiques. La première fait référence aux bases de données administratives du ministère de l'Éducation, qui regroupent des renseignements historiques sur le parcours scolaire des élèves. Ces bases de données totalisent un nombre important d'observations. En revanche, elles rassemblent très peu de renseignements socioéconomiques sur les élèves. La seconde porte sur les enquêtes du PISA, qui, bien que riches sur le plan des données socioéconomiques, ne concernent que les élèves âgés de 15 ans. Toutefois, ces enquêtes comportent un nombre beaucoup moins important d'observations. L'emploi des données administratives et d'enquête a contribué à valider les inférences statistiques par une comparaison des estimations produites à partir de chacune de ces sources.

Cette étude a fait appel à plus d'une méthode économétrique et profité des dernières recherches menées dans le secteur de l'apprentissage machine pour assurer des estimations statistiquement fiables. Les modèles classiques de la sélection de l'échantillon, de l'appariement par les scores de propension, de l'approche des variables instrumentales, du modèle de probit ordonné et des forêts aléatoires causales sont autant de méthodes qui ont été mises à profit pour prendre en considération les particularités statistiques de la reprise scolaire. Par ailleurs, les écrits sur le redoublement scolaire révèlent que les élèves les moins âgés de leur cohorte ont une propension plus importante à redoubler une année scolaire que les autres membres de celle-ci. Les analyses présentées dans cette étude ont

permis d'observer un phénomène identique au Québec, en particulier pour les élèves qui redoublent plus d'une fois. Les méthodes économétriques utilisées ont permis de tirer parti de cette information pour contribuer à la fiabilité statistique des estimations.

Les résultats de cette étude sont sans équivoque : le redoublement ne favorise pas la réussite scolaire des élèves en difficulté et contribue même à compromettre leur succès ainsi que leurs chances d'obtenir un diplôme d'études secondaires. L'étude relève toutefois une exception : la reprise d'une année accroît la diplomation chez les élèves qui envisagent de poursuivre ou qui poursuivent une formation préparatoire au marché du travail ou un programme menant à l'obtention d'un certificat de formation à un métier semi-spécialisé. Il est également estimé que la reprise d'une ou de plusieurs années de scolarité est liée à une baisse des notes aux épreuves uniques d'environ 5 à 12 points de pourcentage selon le moment de la reprise ou la discipline scolaire et selon que des filles ou des garçons sont assujettis à la mesure. En outre, il est estimé que le redoublement scolaire affecte davantage les filles que les garçons en ce qui concerne les résultats aux épreuves uniques du MEQ.

Les estimations indiquent aussi que le redoublement entraîne une baisse de 43 points de pourcentage pour la propension de l'élève à obtenir un diplôme d'études secondaires en 7 ans, et ce, davantage chez les garçons. De plus, le redoublement fait augmenter de 25 points de pourcentage la probabilité qu'un élève quitte l'école sans diplôme ni qualification, peu importe la formation poursuivie. De nouveau, l'effet estimé est plus important pour les garçons. Exceptionnellement, le redoublement scolaire est lié à une hausse de 18 points de pourcentage quant à la propension de l'élève à obtenir un certificat de formation menant à un métier semi-spécialisé. Enfin, la mesure s'avère profitable aux doubleurs du primaire nés précisément au troisième trimestre de l'année, puisqu'elle hausse leurs chances d'être diplômés aux études secondaires et réduit leur risque de n'obtenir aucun diplôme qualifiant.

Une attention particulière a été portée, au cours de l'étude, aux effets du redoublement sur le décrochage scolaire à la formation générale des jeunes. Les effets estimés étaient positifs pour l'ensemble des méthodes économétriques utilisées. Ainsi, le redoublement fait passer de 8 à 15 points de pourcentage le taux de décrochage scolaire de la totalité des doubleurs à la FGJ. Par contre, en ce qui concerne spécifiquement les doubleurs du primaire nés au troisième trimestre de l'année, les méthodes faisant appel à des forêts aléatoires causales instrumentées ont permis d'estimer un effet favorable du redoublement menant à une réduction du décrochage scolaire à la FGJ.

Par ailleurs, les résultats obtenus avec les données du PISA confirment que la reprise scolaire est une mesure délétère pour la réussite éducative des élèves. Ainsi, les scores aux épreuves de mathématiques, de lecture et de science sont réduits de 50 à 180 points en

moyenne si l'élève a redoublé avant l'âge de 15 ans, et ce, compte tenu des autres facteurs affectant la réussite scolaire. Enfin, une analyse coût-avantage montre que le redoublement scolaire est une mesure coûteuse pour la société québécoise. Une analyse de type Kaldor-Hicks permet en effet d'estimer que, pour une seule cohorte d'élèves, le redoublement entraîne une baisse à long terme des revenus anticipés sur le marché du travail de 2,0 milliards de dollars, une diminution des recettes du gouvernement du Québec de 616,5 millions de dollars et une hausse de 752,8 millions de dollars pour les dépenses gouvernementales allouées à l'aide de dernier recours et à la gestion de la criminalité. La rentabilité socioéconomique du redoublement est toutefois positive en ce qui a trait aux élèves pour qui la reprise d'une année scolaire a favorisé une formation qualifiante menant à un métier semi-spécialisé.

Bibliographie

- Alexander, K. L., Entwisle, D. R. et Dauber, S. L. (2003). *On the Success of Failure: A Reassessment of the Effects of Retention in the Primary School Grades* (2nd ed.). Cambridge, England : Cambridge University Press.
- Allen, C. S., Chen, Q., Willson, V. L. et Hughes, J. N. (2009). Quality of design moderates effects of grade retention on achievement: A meta-analytic, multi-level analysis. *Educational Evaluation and Policy Analysis*, 31(4), 480-499.
- Anderson, G. E., Jimerson, S. R. et Whipple, A. D. (2005). Student ratings of stressful experiences at home and school: Loss of a parent and grade retention as superlative stressors. *Journal of Applied School Psychology*, 21(1), 1-20.
- Athey, S. (2015). Machine learning and causal inference for policy evaluation. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 5-6. <https://doi.org/10.1145/2783258.2785466>
- Athey, S. et Imbens, G. W. (2019). Machine learning methods that economist should know about. *Annual Review of Economics*, 11, 685-725.
- Athey, S et Wager, S. (2019). Estimating treatment effect with causal forests: An application. 902.07409.
- Babcock, P. et Bedard, K. (2011). The wages of failure: New evidence on school retention and long-run outcomes. *Education Finance and Policy*, 6(3), 293-322.
- Bellani, D. et Ortiz-Gervasi, L. (2021). Parental time preferences and educational choices. The role of children's gender and social origin. *Cahier de recherche*, 10.13140/RG.2.2.32226.25284.
- Bonvin, P., Bless, G. et Schuepbach, M. (2008). Grade retention: Decision-making and effects on learning as well as social and emotional development. *School Effectiveness and School Improvement*, 19(1), 1-19.
- Borland, J., Wilkins, R. et Tseng, Y. P. (2005). *Experimental and Quasi-Experimental Methods of Microeconomic Program and Policy Evaluation*. Melbourne Institute of Applied Economic and Social Research, 40 pages.

- Byrd, R. S., Weitzman, M. et Auinger, P. (1997). Increased behavior problems associated with delayed school entry and delayed school progress. *Pediatrics*, 100(4), 654-661.
- Caliendo, M. et Hujer, R. (2005). The microeconomic estimation of treatment effects: An overview. *AstA Advances in Statistical Analysis*, 90(1), 199-215.
- Carneiro, P., Hansen, K. et Heckman, J. J. (2003) Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44(2), 361-422.
- Cham, H., Hughes, J. N., West, S. G. et Hee Im, M. (2015). Effect of retention in elementary grades on grade 9 motivation for educational attainment. *Journal of School Psychology*, 53(1), 7-24.
- Chen, J. et Hsiang, C. W (2019). Causal random forest model using instrumental variable quantile regression. *Econometrics*, 7(4). <https://doi.org/10.3390/econometrics7040049>
- Cocks, B., Picchio, M. et Baert, S. (2018). Modeling the effects of grade retention in high school. *Journal of Applied Econometrics*, novembre, 403-424.
- Compas, B. E. (1987). Stress and life events during childhood and adolescence. *Clinical Psychology Review*, 7(3), 275-302.
- Dehejia, R. H. et Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1), 151-161. <https://doi.org/10.1162/003465302317331982>
- Demagnet, J. et Van Houtte, M. (2016). Are flunkers social outcasts? A multilevel study of grade retention effects on same-grade friendships. *American Educational Research Journal*, 53(3), 745-780.
- Diris, R. (2017). Don't hold back? The effect of grade retention on student achievement. *Education Finance and Policy*, 12(3), 312-341.
- Ehmke, T., Drechsel, B. et Carstensen, C. H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. *Studies in Educational Evaluation*, 36(1-2), 27-35.
- Eide, E. R. et Showalter, M. H. (2001). The effect of grade retention on educational and labor market outcomes. *Economics of Education Review*, 20(6), 563-576.
- Festinger, L (1954). A theory of social comparison processes. *Human Relations*, 7: 117.
- Fruehwirth, J. C., Navarro, S. et Takahashi, Y. (2016) How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. *Journal of Labor Economics*, 34(4), 979-1021.
- García-Pérez, J. I., Hidalgo-Hidalgo, M. et Robles-Zurita, J. A. (2014). Does grade retention affect students' achievement? Some evidence from Spain. *Applied Economics*, 46, 1373-1392.
- Gary-Bobo, R. J., Goussé, M. et Robin, J. M. (2016). Grade retention and unobserved heterogeneity. *Quantitative Economics*, 7(3), 781-820.
- Gonzalez-Betancour, S. M. et Lopez-Puig, A. J. (2016). Grade retention in primary education is associated with quarter of birth and socioeconomic status. *PLoS ONE*, 11(11), e0166432. DOI : 10.1371/journal.pone.0166431.
- Goos, M. (2013). *Grade Retention. The Role of the National Educational Policy and the Effects on Students' Academic Achievement, Psychosocial Functioning, and School Career*. Web.
- Greene, J. P. et Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, 2(4), 319-340.

- Gouvernement du Québec (2020a). *Loi sur l'instruction publique* (RLRQ, chapitre I-13.3). Québec : Éditeur officiel du Québec.
- Gouvernement du Québec (2020b). *Régime pédagogique de l'éducation préscolaire, de l'enseignement primaire et de l'enseignement secondaire; Loi sur l'instruction publique* (RLRQ, chapitre I-13.3, a. 447). Québec : Éditeur officiel du Québec.
- Heckman, J. J., Ichimura, H. et Todd, P. (1998). *Matching as an econometric evaluation estimator*. *Review of Economic Studies*, 65(2), 261-294.
- Heckman, J. J., Lalonde, R. J. et Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics, volume 3*. Édité par A. Ashenfelter et D. Card. Elsevier Science.
- Heckman, J. J. et Navarro, S. (2007). Dynamic discrete choice and dynamic treatment effects. *Journal of Econometrics*, 136(2), 341-396.
- Heckman, J., Pinto, R. et Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Reviews*, 103(6), 2052-2086.
- Heckman, J., Tobias, J. L. et Vytlacil, E. (2001). Four parameters of interest in the evaluation of social programs. *Southern Economic Journal*, 68(2), 210-223.
- Hirano, K., Imbens, G. W. et Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161-1189.
- Holmes, C. T. (1984). Effect size estimation in meta-analysis. *Journal of Experimental Education*, 52, 106-109.
- Hong, G. et Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27(3), 205-224.
- Hu, L. C. et Hannum, E. (2020). Red flags: Grade retention and student academic and behavioral outcomes in China. *Children and Youth Services Review*, 113, 104896.
- Hughes, J. N., West, S. G., Kim, H. et Bauer, S. S. (2017, November 9). Effect of early grade retention on school completion: A prospective study. *Journal of Educational Psychology*. Advance Online Publication. <http://dx.doi.org/10.1037/edu0000243>.
- Hwang, S. H. J. et Cappella, E. (2018). Rethinking early elementary grade retention: Examining long-term academic and psychosocial outcomes. *Journal of Research on Educational Effectiveness*, 11(4), 559-587.
- Jacob, B. A. et Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, 86(1), 226-244.
- Jacob, B. A. et Lefgren, L. (2007). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, 1(3), 33-58.
- Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review*, 30(3), 420-437.
- Jimerson, S., Carlson, E., Rotert, M., Egeland, B. et Sourie, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology*, 35, 3-25.
- Jimerson, S. R., Ferguson, P., Whipple, A. D., Anderson G. E. et Whipple, A. D. (2002). Exploring the association between grade retention and dropout: A longitudinal study examining socio-emotional, behavioral, and achievement characteristics of retained student. *The California School Psychologist*, 7, 51-62.

- Kretschmann, J., Vock, M., Lüdtke, O., Jansen, M. et Gronostaj, A. (2019). Effects of grade retention on students' motivation: A longitudinal study over 3 years of secondary school. *Journal of Educational Psychology, 111*(8), 1432-1446.
- Krier, J. (Année inconnue). *Grade Retention in Elementary Schools: Policies, Practices, Results, and Proposed New Directions*. National Center for Mental in Schools at UCLA.
- Lamote, C., Pinxten, M., Van Den Noortgate, W. et Van Damme, J. (2014). Is the cure worse than the disease? A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept. *Educational Studies, 40*(5), 496-514.
- Linden, A. et Yarnold P. R. (2016). Combining machine learning and matching techniques to improve causal inference in program evaluation. *Journal of Evaluation in Clinical Practice, 22*(6), 868-874.
- Lokshin, M. et Sajaia, Z. (2004). Maximum likelihood estimation of endogenous switching regression models. *The Stata Journal, 4*(3), 282-289.
- Lorence, J., Dworkin, A. G., Toenjes, L. A., Hill, A. N., Rotherham, A. et Shepard, L. A. (2002). Grade retention and social promotion in Texas, 1994-99: Academic achievement among elementary school students. *Brookings Papers on Education Policy, 5*, 13-67.
- Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics. *Econometric Society Monographs, 3*.
- Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics, 94*(2), 596-606.
- Mariano, L. T. et Martorell, P. (2013). The academic effects of summer instruction and retention in New York City. *Educational Evaluation and Policy Analysis, 35*(1), 96-117.
- Mariano, L. T. et Martorell, P. (2017). The causal effects of grade retention on behavioral outcomes. *Journal of Research on Educational Effectiveness, 11*(2), 192-216.
- Mariano, L. T., Martorell, P. et Berglund, T. (2018). *The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools*. RAND Corporation, Santa Monica, Calif.
- Marsh, H. W. et Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?. *Journal of Personality and Social Psychology, 47*(1), 213-231.
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T. et Lichtenfeld, S. (2017). Long-term positive effects of repeating a year in school: Six-year longitudinal study of self-beliefs, anxiety, social relations, school grades, and test scores. *Journal of Educational Psychology, 109*(3), 425-438.
- Mathys, C., Véronneau, M.-H. et Lecocq, A. (2019). Grade retention at the transition to secondary school: Using propensity score matching to identify consequences on psychosocial adjustment. *The Journal of Early Adolescence, 39*(1), 97-133. <https://doi.org/10.1177/0272431617735651>
- McCormick, M. P., Neuhaus, R., Horn, E. P., O'Connor, E. E., White, H. S., Harding, S., Cappella, E. et McClowry, S. (2019). Long-term effects of social-emotional learning on receipt of special education and grade retention: Evidence from a randomized trial of "INSIGHTS". *Grantee Submission, AERA Open, 5*(3), 1-21.
- McCoy, A. R. et Reynolds, A. J. (1998). *Grade Retention and School Performance: An Extended Investigation*.

- Ministère de l'Éducation (2006). *Programme de formation de l'école québécoise : éducation préscolaire, enseignement primaire, version approuvée*. Québec : Gouvernement du Québec.
- Ministère de l'Éducation, du Loisir et du Sport (2006). *Le cheminement scolaire de l'élève : questions et éléments de réponse ; principales références dans les encadrements ministériels*. Québec : Gouvernement du Québec.
- Nagin, D. S., Pagani, L., Tremblay, R. E. et Vitaro, F. (2003). Life course turning point: The effects of grade retention on physical aggression. *Development and Psychopathology*, 15, 343-361.
- Nguyen, A. et Taylor, J. (2003). Post-high school choices: New evidence from a multinomial logit model. *Journal of Population Economics*, 16, 287-306.
- Özek, U. (2015). Hold back to move forward? Early grade retention and student misbehavior. *Education Finance and Policy*, 10(3), 350-377.
- Pagani, L., Tremblay, R., Vitaro, F., Boulerice, B. et McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology*, 13(2), 297-315. DOI : 10.1017/S0954579401002061.
- Pedraja-Chaparro, F., Santin, D. et Simancas, R. (2015). Determinants of grade retention in France and Spain: Does birth month matter?. *Journal of Policy Modeling*, 37(5), 820-834.
- Peixoto, F., Monteiro, V., Mata, L., Sanches, C., Pipa, J. et Almeida, L. S. (2016). « To be or not to be retained... That's the question! » Retention, self-esteem, self-concept, achievement goals, and grades. *Frontiers in Psychology*, 7. Article 1550.
- Range, B. G., Pijanowski, J., Holt, C. R et Young, S. (2012). The perceptions of primary grade teachers and elementary principals about the effectiveness of grade-level retention. *Professional Education*, 36(1).
- Roderick, M. (2020). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31(4) (1994): 729-59. Accessed April 7, 2020. www.jstor.org/stable/1163393
- Rosenbaum (2002), P.R. *Observational Studies*. Springer Series in Statistics, 375 p.
- Santarossa, G. et Blais, P. (2018). *Les effets de Passe-Partout sur le développement cognitif et non cognitif des enfants et leur transition à la maternelle cinq ans : une évaluation multi-traitement par appariement sur les scores de propension*. Cahier de recherche, ministère de l'Éducation du Québec.
- Schwerdt, G., West, M. R. et Winters, M. A. (2017). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. *Journal of Public Economics*, 152, 154-169.
- Segal, C. (2013). Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association*, 11(4), 743-779.
- Stearns, E., Moller, S., Blau, J. et Potochnick, S. (2020). Staying back and dropping out: The relationship between grade retention and school dropout. *Sociology of Education*, 80(3) (2007): 210-40. Accessed April 7, 2020. www.jstor.org/stable/20452707
- Therriault, G., Bader, B. et Lapointe, C. (2011) Redoublement et réussite scolaire : une analyse du rapport au savoir. *Revue des sciences de l'éducation*, 37(1).
- Todd, P. E. et Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113, 3-33.

- Uysal, S. D. (2010). The effect of grade retention on school outcomes: An application of doubly robust estimation method. (CMS Discussion Paper, 1). Konstanz : Universität Konstanz, Center for Quantitative Methods and Survey Research (CMS). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-430024>
- Valbuena, J., Mediavilla, M., Choi, A. et Gil, M. (2020). Effects of grade retention policies: A literature review of empirical studies applying causal inference. *Journal of Economic Surveys*, 35(2), 408-451.
- Winship, C. et Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25, 659-706.
- Wu, W., West, S. G. et Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology*, 102(1), 135-152. DOI : 10.1037/a0016664.

Annexe A

Résultats détaillés des régressions économétriques

Tableau A.1 — Résultats des régressions économétriques pour l'épreuve unique de français écrit

Variable	Ecart des moyennes des notes		MCO		Modèle sélect. Éch. 1 régime			
					Équation Y		Équation de sélection	
	(1)		(2)		(3)		(4)	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
Reprise p1 à s4	-10,775	-62,807	-8,313	-48,386	-7,504	-25,615		
<i>Caractéristiques personnelles</i>								
Filles			4,968	46,163	4,996	46,198	-0,197	-12,584
Né à l'étranger			0,295	1,200	0,317	1,293	-0,109	-2,930
Élèves avec handicap			2,032	5,168	1,818	4,060	0,887	20,224
Né au 3ième trimestre							0,238	13,944
<i>Environnement familial</i>								
Revenu moyen (P1 à S5)			0,106	4,850	0,108	5,234	-0,026	-5,883
Mère monoparentale (P1 à S5)			-1,183	-3,783	-1,204	-3,612	0,088	2,162
Mère et/ou père sans diplôme (P1 à S5)			-0,388	-3,103	-0,397	-3,116	0,044	2,496
<i>Programmes et parcours scolaires</i>								
Fréquentation école privée (S1 à S5)			2,863	19,509	2,913	20,490	-0,389	-15,516
PPP (S1 à S5)			2,867	22,541	2,949	23,694	-0,613	-28,198
Passe-partout			-0,173	-0,998	-0,171	-0,987	0,007	0,266
Mater. 4 ans demi-temps			-1,380	-5,665	-1,411	-5,354	0,161	5,220
Changement d'école (P1 à P6)			-0,141	-1,287	-0,172	-1,574	0,213	13,454
Changement d'école (S1 à S5)			-2,538	-21,544	-2,592	-21,290	0,328	20,319
<i>Environnement scolaire</i>								
% nb. Enseignant tit. Temps plein (P1 à P6)			4,667	6,863	4,651	6,790	0,149	1,536
Années ancienneté ens. Tit. (P1 à P6)			-0,036	-1,878	-0,033	-1,676	-0,023	-8,028
IMSE (P1 à S5)			-0,098	-6,576	-0,098	-6,504	-0,006	-2,501
% né à l'étranger (P1 à P6)			-6,504	-6,734	-6,600	-6,768	0,615	4,462
% né à l'étranger (S1 à S5)			-3,118	-3,641	-3,316	-3,843	1,068	8,961
Année construction immeuble (P1 à S5)			-0,033	-7,895	-0,034	-8,009	0,001	2,197
Revenu familial moyen école (P1 à P6)			0,068	1,729	0,065	1,674	0,032	5,190
Revenu familial moyen école (S1 à S5)			0,169	5,549	0,178	5,837	-0,076	-11,135
694 <= Effectifs école < 1006			-0,413	-2,597	-0,391	-2,403	-0,144	-6,239
1006 <= Effectifs école < 1346			-0,101	-0,638	-0,089	-0,547	-0,115	-5,080
1346 <= Effectifs école			-1,092	-6,471	-1,064	-6,246	-0,221	-8,805
Région Ile de Montréal			1,256	7,423	1,281	7,261	-0,202	-8,120
Constante	74,4073	1248,484	132,359	15,728	132,537	15,863	-2,952	-2,337
<i>Paramètres modèle sélect. Échantillonnale</i>								
Lambda							-0,444	-3,924
rho							-0,037	-4,111
sigma							11,988	249,750
Test de Wald équations indépendantes rho = 0 (khi2(1))							15,420	
N	50 407		50 407		50 407			
R2	0,073		0,153					

Tableau A.2 — Résultats des estimations par probit, décrochage de la FGJ, cohortes de 2005-2006, de 2006-2007 et de 2007-2008

Variables	Coefficient	Effets marginaux		t-stat
		$\sum_i \Phi_i$	$\Phi(\bar{X})$	
	(2)	(3)	(4)	(5)
Variable traitement : reprise p1 à s5	0,972	0,117	0,109	83,354
Caractéristiques personnelles				
Fille	-0,200	-0,021	-0,015	-19,342
Né à l'étranger	0,087	0,010	0,007	3,723
Élèves avec handicap	-0,031	-0,003	-0,002	-1,711
Environnement familial				
Mère monoparentale (P1 à S5)	0,147	0,017	0,013	6,483
Mère et/ou père sans diplôme (P1 à S5)	0,107	0,011	0,008	9,826
Revenu familial moyen (P1 à S5)	-0,032	-0,003	-0,002	-11,534
Programmes et parcours scolaire				
Programme passe-partout (4 ans)	-0,135	-0,014	-0,009	-8,383
Maternelle 4 ans demi-temps	-0,003	0,000	0,000	-0,185
Fréquentation école privée (S1 à S5)	-0,335	-0,031	-0,021	-17,474
PPP (S1 à S5)	-0,273	-0,026	-0,019	-18,410
Changement d'école (P1 à P6)	0,205	0,021	0,016	19,348
Changement d'école (S1 à S5)	0,154	0,017	0,012	14,760
Environnement scolaire				
Ratio nb. enseignant tit. temps plein (P1 à P6)	-0,250	-0,027	-0,019	-4,125
Années d'ancienneté ens.tit. (P1 à P6)	-0,004	0,000	0,000	-2,081
IMSE (P1 à S5)	0,015	0,002	0,001	11,564
% né à l'étranger (P1 à P6)	-0,213	-0,023	-0,016	-2,366
% né à l'étranger (S1 à S5)	0,006	0,001	0,000	0,075
Année de construction immeuble (P1 à S5)	-0,003	0,000	0,000	-7,535
Revenu familial moyen école (P1 à P6)	0,040	0,004	0,003	9,822
Revenu familial moyen école (S1 à S5)	0,004	0,000	0,000	1,344
694 <= Effectifs école < 1006	-0,109	-0,011	-0,008	-7,747
1006 <= Effectifs école < 1346	-0,106	-0,011	-0,008	-7,318
1346 <= Effectifs école	-0,098	-0,010	-0,007	-6,213
Constante	4,378			5,129
Nombre d'observations	194 492			