



CAHIERS DU LARJE / WORKING PAPERS

N° 2021-1

LARJE

Laboratoire de Recherches
Juridique et Économique

Les effets du contexte scolaire : pairs et performances des jeunes en Nouvelle-Calédonie

Florent Sari

Université Paris Est Créteil (ERUDITE), CEET et Tepp-CNRS

Amélie Chung

(doctorant, Université de la Nouvelle-Calédonie LARJE)

Résumé : Nous analysons le rôle des pairs et de la composition des contextes scolaires dans la réussite scolaire des jeunes en Nouvelle-Calédonie, petite économie insulaire du Pacifique Sud. Nous utilisons une base de données administrative exhaustive des élèves de classe de troisième pour étudier les effets de pairs sur la réussite au Diplôme National du Brevet. Nous cherchons à déterminer dans quelle mesure la composition des classes et des établissements par les pairs peut affecter la réussite scolaire individuelle des élèves. Pour ce faire, nous estimons les effets de pairs et de composition en ayant recours à un modèle mult;niveau, et à l'approche de Mundlak pour tenir compte de l'endogénéité. Nos résultats montrent l'existence d'effets de pairs plus importants au niveau des classes qu'au niveau de l'établissement. Les élèves réussissent mieux lorsqu'ils sont au contact des élèves les meilleurs.

Abstract : Effects of the school environment on individual student achievement can be observed. In particular, the composition of schools and classrooms, by those peers who are more or less successful, is related to individual performance. We analyze the role of peers and the composition of school settings in the academic achievement of youth in New Caledonia, a small French island economy in the South Pacific. We use a full administrative database of ninth-grade students to study peer effects on success in the Diplôme National du Brevet. We seek to determine the extent to which the composition of classes and schools by peers can affect individual student academic success. To do so, we estimate peer and composition effects using a multilevel model, and the Mundlak approach to account for endogeneity. Our results show the existence of stronger peer effects at the classroom level than at the school level. Students do better when they are in contact with the best students.

Mots-clés : Effets de contexte ; Effets de pairs ; Réussite scolaire.

Keywords: Contextual effects ; educational achievement ; peer effects.

JEL Code: I21, J24, D90

Les effets du contexte scolaire : pairs et performances des jeunes en Nouvelle-Calédonie

The effects of the school context: peers and youth performance in New Caledonia

Florent Sari

Université Paris Est Créteil (ERUDITE), CEET et Tepp-CNRS ; florent.sari@u-pec.fr

Amélie Chung

Université de Nouvelle-Calédonie (LARJE) ; chung.amelie@gmail.com

Introduction

De nombreux travaux en économie de l'éducation se concentrent sur les facteurs familiaux tels que le capital humain, le statut socioéconomique ou le capital culturel au sein des familles, comme les principaux déterminants expliquant les performances scolaires des élèves (Woessmann, 2004 ; Checchi, 2006 ; Björklund et Salvanes, 2011). Il s'avère cependant que des facteurs associés aux contextes scolaires sont aussi à considérer. Les interactions sociales et les facteurs contextuels influencent les comportements individuels notamment la réussite scolaire (Manski, 1995 ; Ammermueller et Pischke, 2009)

Les effets de pairs sont des formes d'interactions entre les membres d'un même groupe qui partagent un univers, un contexte commun. Les pairs désignent les agents avec lesquels un individu est en interaction. Autrement dit, les effets de pairs correspondent aux effets issus des interactions sociales, et se rattachent forcément à un « groupe de pairs » donné. Il peut s'agir, dans le cas de l'éducation, des élèves du groupe classe, des camarades de l'établissement scolaire ou de l'internat. Par le biais de leurs caractéristiques et de leur comportement, des individus situés dans un environnement commun (lieu de travail, quartier, école, classe, résidence universitaire etc..) s'influencent mutuellement. Ces interactions affectent également les comportements scolaires et donc la performance (les résultats), l'assiduité, ou les pratiques (Sacerdote, 2001 ; Zimmerman, 2003).

Les effets de pairs jouent un rôle significatif dans le succès ou non de nombreuses politiques publiques éducatives (Brodaty, 2010). Plus précisément, les effets de pairs sont au cœur des politiques qui traitent de la composition des groupes. Leur importance est mise en avant, par exemple dans les politiques de *vouchers*, une forme de chèque-éducation permettant à des familles démunies d'envoyer leurs enfants dans des écoles privées. Les politiques de *busing* aux Etats-Unis ont également vocation à promouvoir la mixité sociale et ethnique des établissements scolaires, en déterminant des itinéraires spécifiques de bus scolaire. La constitution des groupes de niveaux dans les écoles, les politiques de sectorisation ou de concurrence entre établissements sont autant de champs de l'éducation susceptibles de les traiter. Ces politiques éducatives ont pour trait commun la modification de la composition de groupes, *a priori* définis selon certains critères.

Cette étude a pour objectif d'analyser le rôle des pairs et du contexte scolaire sur la réussite scolaire des plus jeunes dans un contexte français, en Nouvelle-Calédonie. A partir d'une base de données exhaustive de la population des élèves de 3^{eme} de ce territoire, nous analysons les effets de pairs au sein des classes, des établissements (les collèges), et leurs effets sur la réussite scolaire individuelle. Pour mesurer le rôle des effets de pairs nous utilisons les modèles multiniveaux qui permettent d'appréhender l'existence d'effets de contextes, mais également de les quantifier. Ces modèles sont utiles dans les situations où les individus partagent un contexte en commun, emboîté dans d'autres contextes communs.

La première section propose une revue de littérature sur les effets de pairs et leur rôle dans la réussite scolaire. La seconde section présente le contexte de la Nouvelle-Calédonie et les données. La troisième détaille la stratégie économétrique retenue, tandis que la quatrième section expose les résultats de notre analyse.

1. Revue de la littérature

1.1. Le rôle des effets de pairs

Les effets de pairs sont issus des interactions sociales. Des individus ayant un environnement commun s'influencent mutuellement. La notion « d'effet de pairs » est indissociable de celle du groupe de référence, le « groupe de pairs ». Celui-ci correspond aux personnes habitant dans le même voisinage ou scolarisées dans le même établissement. Le périmètre du groupe peut être modifié au cours du temps. C'est le cas par exemple pour une classe avec des cours en classe entière, et d'autres cours selon les options choisies.

Selon Monso *et al.* (2019), « *Dans le champ de l'éducation, un effet de pair correspond à une externalité par laquelle les caractéristiques des camarades d'un élève, ou leur comportement, influent sur le résultat de cet élève* ».

Les canaux par lesquels transitent les effets de pairs sont nombreux et leur identification est difficile. Cependant la littérature économique théorique permet de classer les types d'effets de pairs selon le niveau scolaire des élèves (Lazear, 2001). Dans un article de 2005, Hoxby et Weingarth proposent 8 types de situations de manifestation des effets de pairs :

- Le « Bad apple model » : C'est une forme particulière des effets de pairs (Lazear, 2001) comme une externalité négative : les élèves moins disciplinés ou les plus faibles impactent négativement le reste des élèves. Cette influence négative apparaît lorsque ceux-ci distraient la classe ou qu'ils captent une partie importante de l'attention de l'enseignant.
- Le modèle du « rayon de lumière » (*shining light*) : à l'opposé du précédent, ce modèle suppose qu'il existe une externalité positive des meilleurs éléments du groupe sur l'ensemble de l'unité. Un effet d'entraînement intervient à partir des élèves les plus brillants, et influence les autres.
- Le modèle « Individual comparison model », en opposition aux deux modèles précédents, est basé sur l'idée que la présence de bons éléments entraîne mécaniquement un découragement sur les autres et nuit à leur réussite. Par symétrie, la présence d'élèves plus faibles ferait accroître l'estime des autres et ainsi les encourage à progresser.
- Selon le « single crossing model », les élèves avec un niveau scolaire plus élevé représentent un apport pour le groupe, mais le bénéfice est encore plus notable pour les bons éléments que pour les autres.
- Le « focus model », dit de « la classe homogène », plaide en faveur de la constitution de classes homogènes. En effet les groupes de niveaux représentent un atout car le travail de l'enseignant est facilité, permet une adaptation du matériel ou du rythme sans pénaliser aucun membre du groupe.
- Par opposition, le « rainbow model » suppose que l'hétérogénéité dans un groupe bénéficie à tous les élèves puisqu'ils sont incités à interagir avec l'ensemble du spectre, voient les situations sous plusieurs angles, ce qui favorise leur réussite.
- Le « Boutique Model » indique qu'un élève semble mieux réussir s'il est entouré de membres qui lui ressemblent, par exemple ayant des capacités semblables aux siennes, ce qui faciliterait leurs interactions. Ce modèle est assez proche du Focus model, qui plaide pour la constitution de classes homogènes. Ici, l'élève a intérêt à trouver des pairs qui lui sont similaires, mais cela est également possible dans une classe hétérogène.
- Le « Subculture model » suppose qu'en présence d'une minorité d'élèves aux difficultés scolaires présumées (par exemple non locuteurs de la langue dominante, ou redoublants), la majorité peut les intégrer, ce qui peut être bénéfique pour cette minorité. Au-delà d'une certaine proportion, la minorité n'est pas intégrée correctement ce qui lui est néfaste.
- Le modèle « linéaire en moyenne » (linear-in-means model) : quelle que soit la distribution des performances des élèves, l'effet des caractéristiques ou du comportement des pairs est le même pour tous les élèves. L'influence des pairs n'agit que par leurs caractéristiques moyennes, c'est-à-dire que les effets de pairs sont pris en compte par la moyenne de la qualité du groupe.

L'effet de cette moyenne impacte linéairement, de manière homogène l'ensemble du groupe. En terme de politiques publiques, cela implique qu'une réallocation des individus dans les groupes entraîne un effet redistributif en créant des perdants et des gagnants, mais n'affecte en aucun cas l'efficacité moyenne des résultats.

Les auteurs de l'article signalent que ces modèles n'ont pas la même pertinence empirique et que leurs implications en terme de politiques publiques sont aussi très différentes.

Les théories développées plus haut et présentées par Hoxby et Weingarth (2005) (hormis celle du modèle linéaire en moyenne) font écho à celles qui expliquent les effets de voisinage (Jenck et Mayer, 1990) et repris par Brodaty avec des exemples en milieu scolaire (2010) :

- Les théories de la contagion mettent en avant l'influence du groupe de pairs comme mécanisme capable de générer des effets de voisinage. Au sein d'une classe, cela peut prendre la forme par exemple, l'entraide entre élèves
- Les théories de la socialisation collective insistent sur l'importance du rôle des modèles que sont les adultes du quartier et du contrôle social qu'ils induisent, en aidant par exemple les plus jeunes à intérioriser des normes sociales. Dans une classe, le rôle de modèle est assuré par les meilleurs ou les moins bons élèves.
- Pour les théories de la compétition en revanche, la présence de voisins favorisés peut avoir des effets néfastes sur le voisinage. Elles insistent sur la concurrence potentielle au sein d'un quartier pour réussir socialement lorsqu'il existe une certaine rareté. Dans une classe, lorsque les élèves passent un concours en fin d'année, il est possible que les interactions soient bénéfiques en assurant une élévation du niveau de chacun. Cependant, le partage d'informations fait également augmenter la probabilité de réussite des autres membres concurrents, et peut impacter négativement la réussite individuelle.

1.2. Recherches précédentes : Effets de pairs dans l'éducation

Le travail empirique pionnier qui a mis le thème des effets de pairs dans l'éducation comme un enjeu majeur est le rapport Coleman (1966) aux Etats-Unis. Ce rapport conclue à des effets de pairs massifs par les canaux de l'éducation. C'est également le cas d'autres travaux qui lui ont succédé, comme Henderson *et al.* (1978) et Evans *et al.* (1992). Bien qu'importants dans la littérature, ces travaux sont sujets à des biais dont on sait aujourd'hui tenir compte.

Nombre de ces travaux portent sur les Etats-Unis du fait de la mise en place de mesure d'efficacité du système éducatif depuis 1990, et donc de la mise à disposition de données d'évaluation exhaustives. Cela a permis de recueillir des données à partir d'évaluation standardisées des élèves au cours de leur scolarité. Burke et Sass (2013) observent que les élèves les plus faibles souffrent de la présence de très bons élèves au sein de leur groupe-classe, alors que la présence d'élèves moyens leur est bénéfique. Tout comme Hoxby et Weingarth, ils notent que la présence de camarades de bon niveau scolaire bénéficie aux élèves eux-mêmes de bon niveau. En revanche, Lavy *et al.* (2012) sur une base de données britanniques montrent qu'il est néfaste pour les élèves les plus brillants d'être dans une « trop » bonne classe, notamment pour les garçons. Ce résultat est interprété comme un effet de compétition. En 2004, Angrist et Lang mènent une recherche sur les effets du programme « Metropolitan Council for Educational Opportunity » (Metco). Ce programme consiste à transporter des élèves issus des minorités (pour la plupart noirs) habitant dans le district centre de Boston, vers des districts de banlieue à forte présence blanche. Ils observent que l'arrivée d'élèves de milieux sociaux défavorisés au sein d'école favorisées n'a pas été préjudiciable pour les élèves de cette école. Notons que ce programme est basé sur le volontariat et induit donc une forme de sélection. Guyon *et al.* (2012) étudient la réforme de 1989 en Irlande du Nord qui visait à élargir le recrutement des grammar schools (dans les pays anglophones, établissement du secondaire sélectif) en recrutant des élèves moins bons à l'école. Les résultats montrent un impact positif sur la réussite de ces

élèves, mais également que l'arrivée de ces « moins bons élèves » n'a pas affecté la réussite des autres.

Dans un contexte français, les travaux demeurent rares. Le travail de Goux et Maurin (2005) étudie le retard au collège en tenant compte de la composition du voisinage. Ils utilisent la structure aréolaire des échantillons des enquête Emploi. Toutes choses égales par ailleurs, ils observent que les enfants sont davantage en retard à l'école lorsqu'ils habitent dans un quartier où le taux de retard scolaire des autres enfants est élevé, ainsi que lorsque la composition socioéconomique du quartier est défavorable. Piketty et Valdenaire (2006) étudient la taille des classes et la ségrégation scolaire et concluent à un impact négatif de la ségrégation scolaire cependant moins important que celui de la taille des classes. Brodaty et Gurgand (2016) montrent des effets significatifs des effets de pairs et de l'influence du professeur sur la réussite scolaire dans les universités françaises. Issehnane et Sari (2013) s'intéressent aux liens potentiels entre le lieu de résidence des adolescents et leur réussite scolaire. A partir des données de l'enquête emploi permettant d'étudier le comportement de jeunes de 15 ans entre 1990 et 2002, ils confirment l'existence d'un effet négatif du contexte local sur la probabilité de redoubler, pour ceux qui vivent dans les quartiers les plus « défavorisés ». Boutchenik et Maillard (2019) s'intéressent aux effets de pairs au lycée à partir de données administratives des fichiers anonymisés d'élèves pour la recherche et les études (FAERE). Elles mobilisent des effets fixes portant sur les lycées et les cohortes, et montrent qu'une part élevée de bons élèves dans la classe profite surtout aux plus faibles et peut même s'avérer défavorable pour les autres pairs de niveau élevé.

Une partie des travaux s'intéresse également à d'autres variables comme le décrochage scolaire, les ambitions en matière d'études supérieures ou encore la consommation de drogues. Les travaux sur les effets au-delà du cadre strictement scolaire sont aussi intéressants à considérer, notamment sur l'importance des effets sociaux. Par exemple, Guo *et al.* (2015) constatent que les hommes dont le camarade de chambre universitaire a connu des épisodes d'alcoolisation sévère (*binge drinking*) au lycée, présentent quatre fois plus de chance de connaître eux-mêmes ces expériences lors de leur scolarité.

2. Contexte, données et variables

Dans cette section nous présentons le contexte, les données, les variables retenues et la stratégie empirique pour notre analyse. Nous cherchons à comprendre dans quelle mesure les effets de contexte, à savoir respectivement les effets de la composition de la classe et de l'établissement scolaire, peuvent affecter la réussite scolaire des élèves de classe de 3^{ème} en Nouvelle-Calédonie.

2.1. Le contexte de la Nouvelle-Calédonie

Petit archipel du Pacifique, la Nouvelle-Calédonie est un territoire français qui bénéficie d'un statut particulier, celui de Pays d'Outre-Mer (POM) depuis les accords de Matignon-Oudinot (1988) et de Nouméa (1998). Bien que l'économie calédonienne a bénéficié d'une prospérité économique, celle-ci repose sur des bases fragiles. Entre 1960 et 2010, le produit intérieur brut (PIB) a crû au rythme moyen de 3.9% par an, puis a ralenti entre 2012 et 2015 à 1.7%. En 2015, son PIB par habitant est au-dessus de ceux des autres territoires français, et des autres petits états insulaires du Pacifique. Il est comparable à celui de la Nouvelle-Zélande. Cette petite économie riche puise sa richesse dans l'exploitation de nickel, ressource naturelle non renouvelable, et des transferts en provenance de Métropole.

Elle accuse également un retard en termes d'éducation en comparaison des pays du même niveau de richesse. Même si la proportion de jeunes diplômés a fortement progressé, elle reste

en 2014, inférieure de 16 points à la moyenne de l'OCDE, de 13 points à la Nouvelle-Zélande, et de 22 points à l'Australie (Ris *et al.*, 2017). Par ailleurs, en 2019, 64 % des jeunes d'une même classe d'âge arrivent au bac, ils étaient 50% en 2010, et le taux de réussite à ce diplôme est de 78.5%. Ce taux est inférieur de 9.6 points à celui de la métropole, 7.4 points à la Polynésie Française, et 8.3 points à la Martinique.

Les inégalités y sont également fortement marquées, qu'elles soient sociales, géographiques ou ethniques (Ris, 2013). Un calédonien sur cinq vit sous le seuil de pauvreté en 2019, et le niveau de vie médian des ménages non Kanak est deux fois plus important que celui des ménages kanak. Cela provient du fait que les Kanak occupent des professions moins rémunératrices : seuls 4.4% des actifs occupés kanak sont des cadres et 15.3% des professions intermédiaires (respectivement 15.7% et 26.7% pour les actifs non Kanak). Alors que les kanak, population autochtone de l'archipel représentent 41% de la population en 2019, le taux d'emploi des Kanak est de 48% contre 75% pour les Européens. Ces inégalités prennent notamment leur source dans les différences de dotation en capital humain selon l'origine ethnique. En 2019, 54% des Européens ont au moins un « bac plus deux », loin devant les Métis (24%) et devant les Océaniens (9% pour les Wallisiens et Futuniens, et 8% pour les kanak). Par ailleurs, un non-natif sur deux est hautement diplômé contre un natif sur six. A cela, s'ajoutent de fortes disparités entre les trois provinces : les provinces Nord et des Iles Loyauté, où se concentrent la majorité de Kanak, accusent un retard en termes d'emploi et d'éducation en comparaison de la province Sud.

Ces observations conduisent à penser que la Nouvelle-Calédonie n'est pas totalement épargnée par le risque de malédiction des ressources naturelles. En effet, si l'existence de ressources minières apparaît comme une chance pour un territoire, l'histoire économique montre que la ressource est souvent mal gérée et peut conduire à la « malédiction des ressources naturelles » (Humphreys *et al.*, 2007, Ris *et al.*, 2017). L'exploitation de ressources naturelles est accompagné d'inégalités socio-économiques entre ceux qui détiennent les parts de la rente minière et le reste de la population et des tensions et des conflits associés peuvent conduire à la guerre civile pour le contrôle et l'appropriation des ressources.

Pour corriger ces inégalités, les gouvernements successifs depuis trente ans investissent dans des politiques de rééquilibrage, et les politiques d'éducation et de formation en font partie. Améliorer la réussite scolaire tout en réduisant les inégalités s'inscrit dans le projet éducatif de la Nouvelle-Calédonie (PENC), qui définit les enjeux et axes de l'école calédonienne. Dans ce contexte, comprendre les déterminants, comme les effets de la composition des classes et des établissements, sur la réussite scolaire est pertinent

2.2. Données et variables

Dans ce travail, nous utilisons un jeu de données d'élèves inscrits en classe de 3^{ème} en Nouvelle-Calédonie en 2018. Cette base de données provient du Vice-Rectorat (service déconcentré du ministère de l'Éducation nationale), l'institution en charge de l'enseignement secondaire. Elle fournit des informations sur les 4 198 élèves en classe de 3^{ème} pour l'année 2018. Ces élèves peuvent être inscrits dans un établissement public, dans un établissement privé ou encore inscrits au Centre national d'enseignement à distance (CNED). Pour chaque élève, nous sommes en mesure de déterminer sa classe et son collège.

Pour chaque individu, nous disposons d'informations sur ses caractéristiques sociodémographiques (lieu de naissance, âge, sexe, statut professionnel du chef de ménage, etc.) L'ensemble de données fournit également des informations sur leurs résultats scolaires. Nous observons si l'élève a obtenu ou non le Diplôme National du Brevet (DNB), s'il l'a obtenu avec la mention « très bien » (note moyenne supérieure à 16/20), la mention « bien » (note

moyenne comprise entre 14/20 et 16/20), la mention « assez bien » (note moyenne comprise entre 12/20 et 14/20) ou la mention « passable » (note moyenne entre 10/20 et 12/20).

Dans notre étude, les groupes de pairs sont les classes et les établissements scolaires du secondaire, les collèges. Il s'agit de la population totale des élèves de classe de 3^{ème} de la Nouvelle-Calédonie en 2018, et inscrits pour le DNB, soit 4198 élèves - candidats. Cette base de donnée est exhaustive car à ce moment de la vie scolaire, l'instruction est encore obligatoire (jusqu'à 16 ans). De notre échantillon de 4 198 élèves observés en 2018, nous excluons ceux qui ne sont pas inscrits dans un établissement public ou privé, c'est-à-dire les élèves inscrits au CNED. Nous ne gardons que les élèves pour lesquels nous sommes en mesure d'identifier leur école et leur classe. En effet, nous avons besoin de la classe afin de déterminer les pairs de chaque élève. Nous supposons ensuite qu'un étudiant donné interagit avec tous les autres étudiants de la classe. Finalement, avec ces restrictions et quelques informations manquantes, nous travaillons sur un échantillon de 4 000 élèves sur l'ensemble de la Nouvelle-Calédonie.

Pour mesurer la réussite scolaire des élèves, on s'intéresse aux résultats du DNB. C'est le premier niveau de diplôme des élèves français, mais c'est aussi un moment décisif dans la différenciation des formations possibles (générales, technologiques ou professionnelles) pour la suite de leur scolarité. En effet, de la maternelle à la classe de 3^{ème}, le système français propose un seul et unique parcours éducatif sans aucune différenciation des apprentissages et des attentes pour chaque élève ou étudiant. Après ce diplôme, les élèves doivent faire un choix pour la poursuite de leur carrière académique. De manière générale, les élèves qui obtiennent le DNB avec les plus hautes mentions ont la possibilité de choisir la formation générale, tandis que les autres ont moins de choix ou sont orientés directement par l'école vers des formations qui ne sont pas forcément liées aux choix des élèves.

Pour notre étude, nous choisissons donc d'examiner la réussite scolaire en créant une variable dichotomique indiquant si l'étudiant a obtenu ou non le DNB avec les plus hautes mentions, c'est-à-dire avec la mention « très bien » ou plus (note moyenne supérieure à 16/20 à la note de l'examen final). Cet indicateur prend la valeur 0 pour toutes les autres situations : les étudiants qui ont échoué et les étudiants qui ont obtenu le diplôme mais sans les plus hautes mentions.

Pour les variables indépendantes, nous utilisons des caractéristiques socio-économiques individuelles telles que le sexe, l'âge, le pays de naissance, mais aussi le statut professionnel du chef de ménage (artisan ou chef d'entreprise, cadre, profession intermédiaire, employé, ouvrier, agriculteur, retraité, inactif ou autre). Au niveau de la classe, nous intégrons des variables sur la part des réussites avec mention « très bien » au DNB, les proportions d'élèves nés ailleurs (France métropolitaine, Wallis et Futuna, Vanuatu, Polynésie française). Au niveau de l'école, nous utilisons des variables sur le statut de l'école (privé *versus* public), la part des élèves issus de deux catégories socioprofessionnelles (défavorisée, très favorisée), la part d'élèves non redoublants, et la part d'enseignants titulaires.

3. Stratégie économétrique

3.1. Mesurer les effets des environnements : le modèle multiniveau

Selon Bressoux (2010), modéliser les effets de l'environnement sur les élèves revient à traiter des données spécifiques à des niveaux d'environnement différents. Certaines données sont relatives à un niveau individuel, ce sont celles du premier niveau. D'autres variables concernent les différents environnements et constituent alors des caractéristiques globales ou agrégées, puisque ces niveaux regroupent plusieurs individus (par exemple une famille, un quartier, ou une classe). Les variables agrégées relatent le caractère non substituable des individus du groupe : changer un seul membre du groupe revient à modifier la caractéristique

agrégée de ce groupe, même si elle peut parfois être mineure. Ces variables agrégées révèlent des effets de composition, également nommés effets de contexte.

Notre objectif est de mesurer les effets de pairs au sein de la classe et de l'établissement sur la réussite scolaire des jeunes. Pour notre étude, il aurait été possible d'estimer des modèles *probit* ou *logit*. Or nous devons tenir compte du fait qu'il s'agit de données hiérarchisées et emboîtées. Nous supposons également que la réussite scolaire des jeunes de 3^{ème} peut différer selon les contextes des classes et des établissements. Ces différences peuvent être reliées aux environnements sociaux et économiques au seins desquels les élèves évoluent. Les observations partageant un même environnement (classe ou établissement) sont considérées comme corrélées. Les modèles multi-niveaux permettent de tenir compte de cette structure emboîtée et hiérarchisée de nos données, tout en intégrant l'analyse d'effets des différents environnements sur la réussite scolaire, notamment sur lesdits effets de pairs au sein des classes et établissements. En réponse à ces spécificités, nous utilisons des modèles multiniveaux (Leckie *et al.*, 2012 ; Givord et Guillerm, 2016 ; Bressoux, 2007 ; Anderson *et al.*, 2014). Ces modèles également appelées modèles hiérarchiques linéaires sont une extension de la régression multiple ordinaire. Les modèles multiniveaux présentent l'intérêt de faciliter l'étude simultanée d'au moins deux niveaux d'analyse.

Une telle stratégie économétrique permet de distinguer les effets liés à la composition des classes et établissements, de ceux proprement liés aux caractéristiques des élèves sur la réussite scolaire. Nous exploitons des données disponibles et regroupées en 3 niveaux : le niveau 1 correspond aux élèves ; le niveau 2 correspond aux classes ; le niveau 3 correspond à l'établissement scolaire. Chaque niveau apporte une dimension particulière à notre analyse. Un modèle logistique multiniveau est donc utilisé. Dans ce type de modèle, la constante n'est pas le même d'une unité à l'autre (classes, établissements), ce qui répond à notre objectif de comprendre dans quelle mesure la variable expliquée varie entre les unités.

Nous estimons la probabilité d'être admis au DNB avec la mention « Très bien » (note supérieure à 16/20). Nous modélisons la probabilité d'observer l'occurrence 1 de la variable dépendante y conditionnellement aux covariables. Dans le cas de données non groupées, nous supposons que :

$$P(y_i) = G(\beta_0 + x_i\beta)$$

Où y_i représente la variable binaire dépendante décrite plus haut et sa mesure pour le i ^{ème} individu ; x_i est un vecteur de variables indépendantes mesurées pour chaque individu i ; et G correspond à la fonction de répartition de la loi logistique.

Dans le cas de données hiérarchisées, la modélisation multiniveau émane directement de ce modèle binaire classique. Le modèle multiniveau logistique peut alors s'écrire dans notre cas :

$$P_{ijk} = (x_{ijk}, x_{jk}, x_k, \alpha_k, \alpha_{jk}) = G(\beta_0 + x_{ijk}\beta_1 + x_{jk}\beta_2 + x_k\beta_3 + \alpha_k + \alpha_{jk})$$

Où x_{ijk} , x_{jk} et x_k sont des vecteurs de variables indépendantes correspondant respectivement aux premier, second et troisième niveaux. L'indice i désigne les observations de niveau 1 (les élèves), j représente celles de niveau 2 (les classes) et k celles de niveau 3 (les établissements). Les termes α_k et α_{jk} sont des effets inobservés spécifiques au second niveau et troisième niveaux.

Le modèle peut aussi s'écrire en faisant apparaître la variable latente :

$$\begin{aligned} y_{ijk}^* &= \beta_0 + x_{ijk}\beta_1 + x_{jk}\beta_2 + x_k\beta_3 + \alpha_k + \alpha_{jk} + \varepsilon_{ijk} \\ y_{ijk} &\{ 1 \text{ si } y_{ijk}^* \geq 0, 0 \text{ sinon} \end{aligned} \quad (1)$$

Les termes $\alpha_k + \alpha_{jk} + \varepsilon_{ijk}$ correspondent aux termes inobservés qui sont décomposés en un terme individuel ε_{ijk} , un terme α_{jk} commun à tous les individus de la même classe, et α_k un terme commun à tous les individus du même établissement. Ils somment des effets des variables inobservées affectant tous les individus de la classe et du même établissement. Comme pour un modèle *logit*, ε_{ijk} représente le terme d'erreur suivant une distribution logistique.

3.2. Endogénéité

Notre analyse est exposée à différents problèmes d'endogénéité qui peuvent entraîner certains biais dans l'estimation des coefficients. On parle d'endogénéité lorsqu'une variable inobservée affectant les outputs de l'éducation, est corrélée avec une variable indépendante du modèle.

En premier lieu, il est important de noter que les modèles à effets aléatoires sont basés sur une hypothèse forte : les variables explicatives de chaque niveau ne doivent pas être corrélées avec les termes aléatoires de l'autre niveau. L'absence de biais et la convergence des estimateurs de ces modèles reposent sur cette hypothèse d'exogénéité des facteurs de régression. Cependant dans notre cas, nous pouvons imaginer que certaines caractéristiques non observables des unités aux deux niveaux supérieurs (classes et établissements), reléguées aux termes d'erreur, soient corrélées avec les caractéristiques individuelles observables x_{ijk} . Aussi, il est raisonnable de penser que les ménages (et ainsi les élèves) les plus aisés choisissent les établissements avec la meilleure réputation, par exemple. Or la réputation est un facteur inobservable dans nos données. La réussite d'un élève s'expliquerait par les caractéristiques des parents, mais également par la qualité du collège, et il serait difficile de distinguer ces deux effets. Une façon de prendre en compte ce problème est d'estimer un modèle qui tient compte explicitement de cette possible corrélation entre variables observables des élèves et caractéristiques inobservables des groupes (ici, les établissements).

Les problèmes d'endogénéité n'ont reçu que peu d'intérêt dans la littérature économétrique à propos des modèles multiniveaux. En effet la littérature théorique sur la stratification en niveaux apparaît dans les années 1990 (Epple et Romano, 1998, 2006 ; Nechyba, 2003), et celle concernant la prise en compte de l'endogénéité dans les modèles multiniveaux est très récente. Nous pouvons citer les travaux de Kim et Frees (2007), Hanchane et Mostafa (2012) ou encore Benbiga *et al.* (2012).

Pour parer aux problèmes d'endogénéité, nous retenons l'approche de Mundlak (1978) proposée par Hanchane et Mostafa (2012) puis Dieng (2017). Cette approche est généralement utilisée pour les estimations des données de panel ou les modèles multiniveaux, dont les structurations des données sont assez proches l'une de l'autre. Ils proposent la correction du problème d'endogénéité en introduisant dans l'équation à estimer, la moyenne des variables individuelles pour chaque établissement. L'introduction de moyennes individuelles au niveau de l'établissement permet de tenir compte de l'hétérogénéité non observée des variables non constantes dans la hiérarchie des unités (dans le temps, pour les données de panel). (Voir Trognon, 2003, Hanchane et Mostafa, 2012, pour une explication détaillée).

Dans la pratique, la corrélation entre les variables individuelles indépendantes et les termes aléatoires relatifs aux établissements peut être écrite de la façon suivante : $\alpha_{jk} = \bar{x}_k \beta_4 + u_{jk}$ où β_4 est non-nul et \bar{x}_k correspond à la moyenne des variables indépendantes dans l'établissement k et u_{jk} un terme d'erreur de moyenne nulle non corrélé aux variables indépendantes. Ainsi, nous introduisons ce terme dans l'estimation, en ajoutant au modèle les moyennes intra-établissements des variables individuelles. Le modèle estimé s'écrit alors de la façon suivante :

$$y_{ijk}^* = \beta_0 + x_{ijk}\beta_1 + x_{jk}\beta_2 + x_k\beta_3 + \bar{x}_k\beta_4 + \alpha_k + u_{jk} + \varepsilon_{ijk} \quad (2)$$

Il est important de noter que les modèles multiniveaux, tels que celui que nous mobilisons dans ce travail, reposent sur une hypothèse forte : les variables indépendantes de chaque niveau ne doivent pas être corrélées avec les termes aléatoires de l'autre niveau. L'absence de biais et la convergence des estimateurs de ces modèles reposent sur cette hypothèse d'exogénéité. Si l'approche de Mundlak (1978) permet de tenir compte d'une potentielle corrélation entre les caractéristiques individuelles et les termes aléatoires des établissements, elle ne permet pas de considérer les cas où les caractéristiques inobservables des élèves sont corrélées aux caractéristiques observables de la classe ou de l'établissement. Une illustration serait le cas de familles ambitieuses (caractéristique inobservable) qui chercheraient à inscrire leurs enfants dans les établissements dont le taux de réussite au DNB est le plus élevé. Toutefois, il existe en Nouvelle-Calédonie la sectorisation qui impose le collège public par rapport à l'adresse de résidence de l'élève pouvant ainsi contribuer à limiter ce phénomène.

Un autre biais pourrait découler du fait que les choix de localisation des ménages ne sont pas totalement exogènes. Ces derniers peuvent être guidés par une volonté de se rapprocher des quartiers proposant le meilleur contexte socio-économique et les meilleurs établissements scolaires. Pour pleinement tenir compte de ce phénomène, le quartier de résidence pourrait être considéré comme un niveau supplémentaire dans le modèle multiniveau. Une approche par variables instrumentales pourrait aussi être privilégié. Elle suppose néanmoins de trouver des instruments expliquant le choix du lieu de résidence mais n'ayant pas une influence directe sur la variable dépendante. Nos données ne nous permettent cependant pas d'identifier le lieu de résidence de façon aussi précise (seule la commune est renseignée) et nous ne disposons pas non plus de tels instruments.

4. Résultats

4.1. Statistiques descriptives

Le tableau 1 présente les résultats au DNB en Nouvelle-Calédonie en 2018. Le taux de réussite au DNB est de 77%, environ 3% des inscrits ne se sont pas présentés à l'examen. Les élèves refusés représentent 20 %, et la part la plus conséquente, celle des « admis » sans mention est de 26.8 %. La part des « admis avec mention » diminue en fonction des différents niveaux de mention. Ils sont 30% à réussir avec mention « bien » ou « très bien ». Les filles réussissent mieux que les garçons (42% contre 34.6%, toutes catégories « admis » confondues). Concernant la « très bonne réussite », elles sont 18.6% à obtenir une mention « bien » ou « très bien », contre 11.6% chez les garçons. La province Sud recense presque trois quarts des élèves de 3^{ème} et près de 56 % des admis toutes catégories confondues. Plus de trois quarts des élèves sont inscrits dans le secteur public. Concernant le lieu de naissance, plus de 82% des élèves natifs de NC et 8.6% de France métropolitaine. Les élèves nés en Polynésie Française (Poly.), au Vanuatu (Van.) et à Wallis et Futuna (WF) représentent respectivement moins de 1% de la population étudiée.

Tableau 1 : Caractéristiques des élèves pour les différents résultats au DNB en NC

		Absent	Admis AB	Admis B	Admis TB	Refusé	Total	
		Fréq.	137	799	726	507	814	
Genre	Filles	%	3.36	26.82	19.60	17.81	19.97	
		Fréq.	51	519	437	417	341	
		%	1.25	12.73	10.72	10.23	8.37	
		Fréq.	86	574	362	309	166	
							506	
							2003	

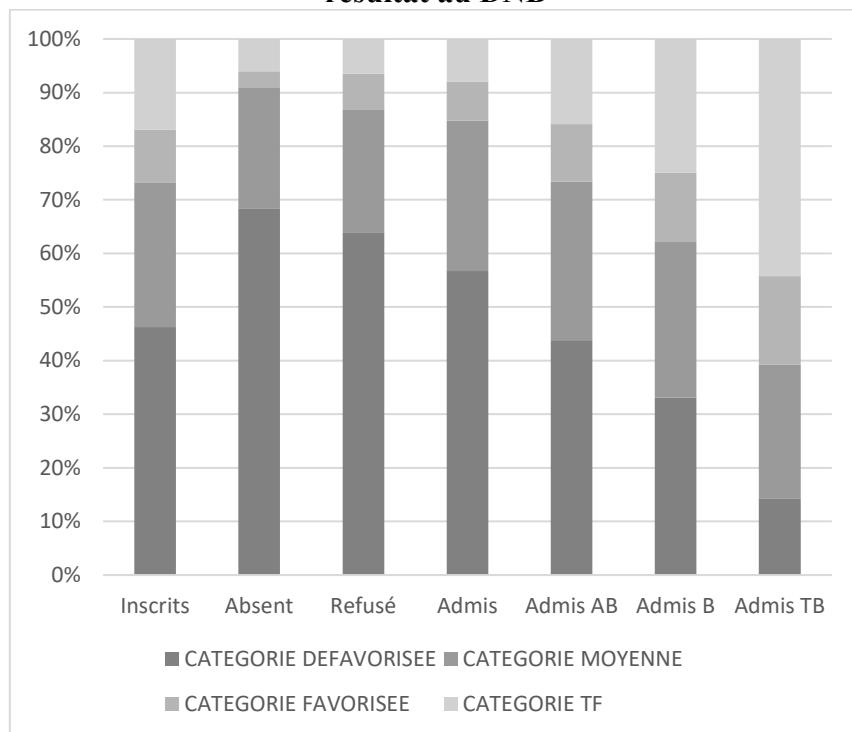
		%	2.11	14.08	8.89	7.58	4.07	12.41	49.14
Province	Nord	Freq.	29	224	163	108	43	169	736
		%	0.71	5.50	4.00	2.65	1.05	4.15	18.06
	Sud	Freq.	98	746	531	555	448	575	2953
		%	2.40	18.30	13.03	13.62	10.99	14.11	72.45
	Iles	Freq	10	123	105	63	16	70	387
		%	0.25	3.02	2.58	1.55	0.39	1.72	9.49
Statut établissement	Privé	Freq	26	293	215	170	107	163	974
		%	0.64	7.19	5.27	4.17	2.63	4.00	23.90
	Public	Fréq.	111	800	584	556	400	651	3102
		%	2.72	19.63	14.33	13.64	9.81	15.97	76.10
Lieu de Naissance	Autre	Freq	16	87	57	47	26	45	278
		%	0.39	2.13	1.40	1.15	0.64	1.10	6.82
	France	Freq	10	33	61	106	131	12	353
		%	0.25	0.81	1.50	2.60	3.21	0.29	8.66
	NC	Freq	104	947	670	556	348	744	3369
		%	2.55	23.23	16.44	13.64	8.54	18.25	82.65
	Poly.	Freq	3	6	4	8	2	5	28
		%	0.07	0.15	0.10	0.20	0.05	0.12	0.69
	Van.	Freq	0	6	3	3	0	1	13
		%	0	0.15	0.07	0.07	0	0.02	0.32
	WF	Freq	4	14	4	6	0	7	35
		%	0.10	0.34	0.10	0.15	0	0.17	0.86

Source : Vice-Rectorat de Nouvelle-Calédonie (2018).

Concernant l'origine sociale des élèves, la figure 1 présente la répartition de l'origine sociale des élèves (selon la catégorie socio-professionnelle du responsable de l'élève) aux différentes catégories de résultats du DNB en 2018. Les différentes catégories de professions du responsable ont été regroupées en quatre modalités par le vice-Rectorat : Défavorisée, moyenne, favorisée et très favorisée.

Si on n'observe que peu de différence entre la catégorie « refusé » et « admis », à mesure que le niveau de mention augmente pour le diplôme, la part d'élèves issus de la catégorie très favorisée augmente, et inversement pour la catégorie défavorisée. Les élèves issus d'un milieu très favorisé représentent 7% des refusés contre 44.3% chez les admis avec mention « très bien ». La catégorie sociale moyenne est stable entre 22% et 30% dans chaque résultat. Les élèves d'origine sociale défavorisée représentent 63.8 % des candidats refusés contre 14.3 % chez les admis avec cette mention « très bien ».

Figure 1 : Répartition des catégories socioprofessionnelles des élèves selon le résultat au DNB



Source : Vice-Rectorat de Nouvelle-Calédonie (2018).

4.2. Estimations

L’analyse économétrique consiste à estimer des équations de performances scolaires, le niveau de réussite à un diplôme, à trois niveaux. La probabilité de réussir (avec la mention « très bien ») au DNB est estimée, en tenant compte de variables attribuées à 3 niveaux différents. Le premier niveau représente les caractéristiques de l’élève et de son environnement social (niveau élève). Le second niveau fait ressortir les effets de contexte des classes avec les caractéristiques de leurs compositions (niveau classe). Le troisième niveau est relatif au contexte de l’établissement avec les caractéristiques de la composition des établissements (niveau établissement).

Tableau 2 : Résultats des estimations

	Modèle Logit	Modèles Logistiques à effets aléatoires			
		0	1	2	3
	Coeff (Std.E)	Coeff (Std.E)	Coeff (Std.E)	Coeff (Std.E)	Coeff (Std.E)
Constante	9.439*** (0.814)	-1.234*** (0.119)	17.943*** (1.372)	16.265*** (1.472)	16.900*** (1.631)
Caractéristiques de l’élève					
Homme	-0.426*** (0.048)		-0.768*** (0.083)	-0.765*** (0.083)	-0.754*** (0.084)
Age	-0.693*** (0.052)		-1.289*** (0.097)	-1.247*** (0.097)	-1.268*** (0.099)
Pays de naissance					
France	0.633*** (0.085)		1.182*** (0.147)	1.143*** (0.146)	1.108*** (0.149)

	<i>Ref</i>			
Nouvelle-Calédonie				
Wallis et Futuna	-0.169 (0.273)	-0.342 (0.499)	-0.351 (0.499)	-0.339 (0.499)
Vanuatu	0.210 (0.416)	0.540 (0.727)	0.435 (0.726)	0.470 (0.726)
Polynésie	0.388 (0.270)	0.757 (0.472)	0.677 (0.471)	0.699 (0.471)
Autre	0.146 (0.118)	0.234 (0.193)	0.194 (0.199)	0.233 (0.207)
Statut du responsable				
Agriculteur	-0.129 (0.269)	-0.286 (0.476)	-0.221 (0.469)	-0.167 (0.476)
Artisan, commerçant	-0.076 (0.101)	-0.079 (0.169)	-0.089 (0.168)	-0.113 (0.171)
Cadre	0.242*** (0.092)	0.417*** (0.155)	0.392*** (0.154)	0.412*** (0.156)
<i>Profession intermédiaire</i>	<i>Ref</i>			
Employé	-0.281*** (0.082)	-0.492*** (0.140)	-0.482*** (0.139)	-0.482*** (0.141)
Ouvriers	-0.521*** (0.082)	-0.943*** (0.142)	-0.927*** (0.140)	-0.910*** (0.142)
Retraité	-0.621*** (0.174)	-1.175*** (0.318)	-1.165*** (0.315)	-1.158*** (0.318)
Inactif	-0.597*** (0.097)	-1.213*** (0.171)	-1.159*** (0.170)	-1.102*** (0.173)
Indéterminé	-0.668*** (0.149)	-1.290*** (0.268)	-1.212*** (0.267)	-1.214*** (0.269)
Caractéristiques de la classe				
% d'élèves « très bons »	1.568*** (0.251)		2.167*** (0.530)	1.869*** (0.543)
Taille de la classe	0.017* (0.010)		0.0367* (0.020)	0.045* (0.023)
% d'élèves nés en :				
France	0.056 (0.382)		0.657 (0.804)	0.166 (0.862)
Nouvelle- Calédonie	<i>Ref</i>			
Wallis et Futuna	0.369 (1.082)		1.193 (2.190)	1.151 (2.233)
Vanuatu	-1.552 (1.902)		-0.860 (4.007)	-0.413 (4.020)
Polynésie	-1.259 (1.243)		-1.273 (2.587)	-0.979 (2.609)
Autre	0.150 (0.198)		0.600 (0.438)	0.600 (0.461)
Caractéristiques de l'établissement				
public	-0.173** (0.075)			-0.491** (0.224)
% d'élèves de PCS défavorisée	-0.002 (0.002)			-0.006 (0.007)
% d'élèves de PCS très favorisée	0.002 (0.005)			0.005 (0.015)
% d'élèves non redoublants	-0.004 (0.003)			-0.004 (0.009)
% d'enseignants titulaires	0.002 (0.002)			0.005 (0.005)
Variance inter - classe	0.534 (0.147)	0.378 (0.120)	0.178 (0.0813)	0.172 (0.079)
	0.312	0.182	0.067	0.065
Variance inter-établissement	(0.071)	(0.058)	(0.0511)	(0.054)

	0.847	0.560	0.245	0.237
--	-------	-------	-------	-------

Source : Vice-Rectorat de Nouvelle-Calédonie (2018).

Le Tableau 2 compare les résultats d'un modèle de régression logistique simple, puis les résultats des estimations logistiques multiniveaux. Le modèle 0 correspond au modèle multiniveau « vide ». Les modèles 1, 2, 3 intègrent successivement les caractéristiques relatives aux différents niveaux.

Le modèle « vide » est inconditionnel dans le sens qu'aucune variable explicative n'y est intégrée. La comparaison de la variance au niveau des groupes avant et après l'introduction de variables individuelles, de la classe et de l'établissement, permet de quantifier la part que représentent les effets dus à la décomposition de la variance dans la variabilité inter groupes. Il consiste à estimer la variance du phénomène, ici le résultat de l'élève au diplôme, en faisant une répartition de cette dernière sur les différents niveaux structurants de nos données.

Concernant la partie aléatoire, ce sont les variances résiduelles à chaque niveau de groupes et pour chaque modèle qui sont considérées. Nous avons alors une idée de la contribution de l'ajout d'un modèle par rapport au précédent sur la variance intergroupe. En premier lieu, nous observons que la variance interclasse est systématiquement supérieure à celle inter-établissement, quel que soit le modèle aléatoire et le niveau inclus (tableau 2). Concernant l'évolution de la variance inter-établissement, elle est de 0.312 pour le modèle vide et atteint 0.182 avec les variables des caractéristiques individuelles. Elle diminue ici de 42% (tableau 3), puis ne varie plus entre les modèles 1 à 3. Il existe une variance importante entre les classes, mais la prise en compte des caractéristiques individuelles, et de surcroît celle de la composition de la classe la fait diminuer considérablement. Concernant la variance entre établissement, c'est la prise en compte des caractéristiques individuelles des élèves qui la rend pratiquement nulle.

Tableau 3 : Proportion de la réduction de la variance

	Modèle 0 à modèle 1	Modèle 1 à 2	Modèle 2 à 3
	$\frac{\sigma^2_{M_0} - \sigma^2_{M_1}}{\sigma^2_{M_0}}$	$\frac{\sigma^2_{M_1} - \sigma^2_{M_2}}{\sigma^2_{M_1}}$	$\frac{\sigma^2_{M_2} - \sigma^2_{M_3}}{\sigma^2_{M_2}}$
Inter-classe	0.29	0.53	0.03
Inter-établissement	0.42	0.00	0.003

Concernant la partie aléatoire de l'analyse, le tableau 3 montre que les variances de niveau classe et de niveau établissement diminuent à mesure que le modèle évolue. Pour les classes, la variance passe de 0.534 à 0.378 entre le modèle vide au modèle incluant les variables des caractéristiques liées aux élèves (niveau 1), à 0.178 puis 0.172 lorsque sont introduites les variables de contextes classe (niveau 2) puis de l'établissement (niveau 3). L'ajout de caractéristiques de niveau 1, fait diminuer la variance résiduelle de 29%, puis l'ajout des variables de contexte classe la font diminuer de 53%. On constate que la différence de variance inter-classe comme celle inter-établissement entre les modèles 2 et 3 est quasiment nulle.

Pour l'analyse de la partie fixe, nous nous intéressons d'abord aux caractéristiques relatives aux élèves (niveau 1), avec le genre, l'âge, le lieu de naissance et le statut du responsable. D'après les résultats des modèles, être un garçon (par rapport au fait d'être une fille) et avancer dans l'âge affectent négativement et significativement la probabilité de réussir avec la mention bien ou très bien au diplôme. Concernant le lieu de naissance, les élèves nés en France métropolitaine (hors DOM-TOM) et comparativement aux élèves en Nouvelle-Calédonie, ont une probabilité positive et significative de très bien réussir au DNB. Concernant l'occupation du responsable de l'élève, avec catégorie de référence « Profession intermédiaire », les enfants

de cadres ont une probabilité significative et positive de très bien réussir. Pour les enfants d'employés, ouvriers, inactifs ou indéterminés, leurs chances sont négatives et significatives.

Au niveau 2, celui qui concerne les effets de contexte-classe, nous avons des variables sur la composition de la classe avec la part d'élèves « très bons » dans le groupe classe et les parts d'élèves en fonction des lieux de naissance, et également une variable sur la taille de la classe, mesurée par le nombre d'élèves. Le fait d'être en contact avec une proportion plus élevée d'élèves qui réussissent très bien, augmente positivement la probabilité de réussir très bien pour les pairs. On assimile ce résultat aux effets de pairs au sein des classes. La taille de la classe a également un effet positif, bien que le coefficient soit significatif à 10% seulement. Ici, partager un contexte classe avec des élèves nés hors de Nouvelle-Calédonie, quel que soit le lieu, n'a aucun effet significatif sur notre variable dépendante.

Au niveau de l'effet établissement, nous disposons des variables sur le statut du collège (privé ou public), la part d'élèves de PCS défavorisées et très favorisées, la part d'élèves non redoublants et celle d'enseignants titulaires. Finalement c'est le statut qui est le seul facteur significatif sur ce niveau : le statut public (contre privé) affecte négativement (au seuil de 5%) les chances de réussite. La composition du collège, qu'elle soit pour les élèves (être en contact avec des élèves issus de CPS défavorisées ou favorisées, et redoublants ou non) ou pour l'équipe pédagogique (le fait d'être en contact avec des enseignants titulaires) n'a pas d'effet significatif.

Nous montrons que c'est au premier niveau, celui des caractéristiques individuelles que les effets sont les plus significatifs. Le lieu de naissance est intéressant : les enfants nés en France métropolitaine et comparativement à ceux nés en Nouvelle-Calédonie, ont significativement plus de chances de mieux réussir au DNB, alors qu'ils représentent moins de 9% de la population.

S'agissant des effets de contexte (classes, établissements), on met en évidence l'existence d'effets de pairs au niveau de la classe. Partager un contexte de classe avec une part plus élevée de bons élèves fait augmenter significativement les chances de réussite de l'élève. Ce résultat est assimilé aux interactions entre élèves de la même classe. De plus, la variance interclasse diminue lorsqu'on intègre les variables de niveau 2. Ce phénomène est amplifié pour la variance inter-établissement. La variance interclasse diminue de 53% lorsque les effets classes sont ajoutés, et devient quasiment nulle pour la variabilité entre les établissements. Autrement dit, si la variance résiduelle en tenant compte seulement des caractéristiques individuelles est relativement importante, l'ajout de variables d'effets-classe vient largement diminuer la variabilité entre les classes. La variance résiduelle reste pratiquement stable après l'introduction des effets établissements. Nous pouvons dire qu'il existe un effet classe important dans la réussite scolaire en classe de 3^{ème} et, comparativement, que la variabilité des établissements est plus faible.

Maintenant, pour tenir compte de l'endogénéité, nous utilisons l'approche de Mundlak (1978). Les résultats de l'estimation de la probabilité de très bien réussir au DNB avec cette méthode sont présentés dans le tableau 4.

Il est possible que le choix d'inscription au sein d'un établissement soit soumis à des biais d'endogénéité et l'approche de Mundlak est une alternative pour en tenir compte. Des variables de moyennes au sein de l'établissement sont intégrées, et nous y observons la variance inter-établissement. Les résultats de l'estimation par cette approche confirment les résultats des estimations à effets aléatoires concernant les caractéristiques des élèves du modèle complet (le modèle 3). Il en est de même pour les caractéristiques de la classe, quoique la taille de la classe devient significative à 5%. Concernant les caractéristiques des établissements, une différence notable est que le statut public n'est plus significatif. Pour les variables moyennes par établissement, aucune n'est significative outre la part d'élèves « très bons » au DNB. Nous concluons que la composition des unités au niveau des établissements implique des effets de

pairs. Les élèves réussissent significativement mieux lorsqu'ils sont au contact des élèves les meilleurs. Cet effet existe au niveau des classes et au niveau des établissements.

Tableau 4 : Résultat de l'estimation par l'approche de Mundlak

Variables	Coeff.	Std. Error
Constante	4,402	12,421
Caractéristiques de l'élève		
Homme	-0,754***	0,085
Age	-1,289***	0,099
Pays de naissance		
France	1,116***	0,149
<i>Nouvelle-Calédonie</i>	<i>Ref.</i>	
Wallis et Futuna	-0,322	0,497
Vanuatu	0,518	0,734
Polynésie	0,723	0,473
Autre	0,242	0,214
Statut du responsable		
Agriculteur	-0,038	0,488
Artisan, commerçant	-0,087	0,173
Cadre	0,424***	0,157
<i>Profession intermédiaire</i>	<i>Ref.</i>	
Employé	-0,481***	0,141
Ouvriers	-0,911***	0,143
Retraité	-1,155***	0,319
Inactif	-1,134***	0,174
Indéterminé	-1,258***	0,273
Caractéristiques de la classe		
% d'élèves « très bons » au DNB	1,916***	0,484
Taille de la classe	0,051**	0,025
Caractéristiques de l'établissement		
Public	-0,431	0,268
% d'élèves de PCS défavorisée	-0,015	0,013
% d'élèves de PCS très favorisée	0,005	0,025
% d'élèves non redoublants	-0,002	0,011
% d'enseignants titulaires	0,011	0,007
Moyennes de l'établissement		
% d'élèves masculins	-0,8	1,603
Age	0,949	0,821
% d'élèves nés en France	-0,243	2,774
% d'élèves nés en NC	<i>Ref.</i>	
% d'élèves nés à Wallis et Futuna	0,826	10,156
% d'élèves nés au Vanuatu	-6,204	18,472
% d'élèves nés en Polynésie Française	-1,101	10,02
% d'élèves nés ailleurs	0,178	0,601
% d'agriculteurs (responsable)	-6,475	6,055
% d'Artisans, commerçants (responsable)	-3,321	3,527
% de cadres (responsable)	-3,54	3,154
% de professions intermediaries (responsable)	<i>Ref.</i>	
% d'employés (responsable)	-1,226	2,794
% d'ouvriers (responsable)	-0,073	2,181
% de retraités (responsable)	-4,808	4,394
% d'inactifs (responsables)	0,751	2,373
% de statuts indéterminés (responsable)	0,01	2,561
% d'élèves « très bons » au DNB	3,886**	1,97

Taille des classes	-0,008	0,059
Variance inter-établissement	0,206	0,095
Observations	4 000	

Source : Vice-Rectorat de Nouvelle-Calédonie (2018).

Conclusion

L'objectif de ce travail est mettre en évidence la présence d'effets de contexte liés au système scolaire. L'effet de la composition du contexte de la classe et du contexte de l'établissement sont mesurés, en plus des caractéristiques individuelles, sur la très bonne réussite des jeunes au diplôme du brevet en Nouvelle-Calédonie.

Nous montrons dans un premier temps l'importance de l'utilisation du modèle multiniveau, à effets aléatoires pour notre cas d'étude. La nature des données hiérarchisées d'une part, et la problématique de l'étude, indiquent la nécessité d'aborder l'analyse avec ces outils. L'utilisation de modèles multiniveaux à effets aléatoires permet justement de mesurer les effets des contextes, contrairement à une régression logistique simple. Pour faire face à des problèmes d'endogénéité hypothétiques dans notre analyse, nous étayons nos résultats grâce à l'approche de Mundlak. Notre étude est faite sur l'ensemble de la population scolaire de classe de 3^{ème} en 2018 en Nouvelle-Calédonie.

Nos résultats révèlent l'importance des effets de contexte, mais c'est surtout vrai pour le niveau classe. Nous montrons qu'un élève augmente ses chances d'obtenir une mention « très bien » au premier diplôme de l'enseignement français lorsqu'il partage sa classe avec des élèves qui ont eux-mêmes de fortes capacités scolaires. Cela alors même que les variables liées aux effets de composition de l'établissement comme la part d'enseignants titulaires ou d'élèves issus de catégories socio-professionnelles défavorisées n'ont pas d'effets significatifs. Cela rejoint les travaux de Burke et Sass (2013) qui trouvent des effets de pairs détectables seulement au niveau de la classe, suggérant que cette échelle semble être plus pertinente que celle de l'établissement. Hoxby et Weingarth (2005) montrent que les élèves en bas de la distribution (premier décile) des scores initiaux tirent profit de l'apport des élèves d'un niveau légèrement supérieur (deuxième ou troisième décile). Aussi, dans notre étude l'effet des variables individuelles tels que le genre et la catégorie sociale d'origine confortent les résultats d'autres recherches sur les performances scolaires. La réussite au brevet est également sensible au lieu de naissance : un enfant né en France métropolitaine a plus de chance de très bien réussir qu'un élève natif de la Nouvelle-Calédonie. Par ailleurs, il est mis en évidence que les effets de composition au niveau de l'établissement ou de la classe selon le lieu de naissance n'ont pas d'effets significatifs. Autrement dit, partager un environnement scolaire avec un élève né en France métropolitaine n'a pas d'effet sur ses camarades, alors même que ce facteur est significatif dans les caractéristiques individuelles.

Une piste intéressante aurait été d'effectuer des estimations pour chacune des 3 provinces et de les comparer pour comprendre éventuellement d'autres effets des contextes locaux. Le nombre restreint d'unités de niveau 3 (établissement) n'a pas rendu possible cette piste. Par ailleurs si la méthode utilisée nous permet de tenir compte de la structure emboîtée des données, elle ne nous permet pas d'identifier les différents types d'effets de pairs (endogènes, corrélationnels, exogènes). Une autre limite de notre travail réside dans le fait que nous n'avons pas d'information sur la procédure de composition des classes. Il est raisonnable de penser que les établissements regroupent les élèves qui réussissent le mieux, et ceux qui réussissent moins bien par ailleurs.

Ces résultats sont d'autant plus importants que les écarts de réussite scolaire se creusent en avançant dans les niveaux de diplômes. De plus, nous nous intéressons au moment où les élèves des territoires français terminent le collège unique. Aucune distinction de filière n'est faite avant, et l'orientation dans les filières professionnelles, technologiques ou générales se fait à ce moment. Le moment de la fin du collège, caractérisé par le DNB, est un moment charnière dans

les possibilités et les choix des futurs parcours scolaires. Les élèves qui réussissent très bien à ce niveau de diplôme auront beaucoup plus de choix dans les voies de poursuite. Cela est d'autant plus vrai que la Nouvelle-Calédonie souffre de congestion dans certaines filières post - 3^{ème}. Plus largement, nos résultats suggèrent qu'une composition hétérogène des classes et des établissements selon la réussite scolaire est favorable à la réussite des pairs. Nous ne disposons pas d'information quant aux modalités de composition des classes, s'il existe une directive commune réglementaire ou si les établissements peuvent le faire librement. Il serait intéressant de pouvoir en savoir davantage, pour proposer des améliorations localisées, mais également de disposer dans les bases de données d'informations sur les options (latin, sport, musique) et les langues pour affiner les résultats.

Références

- Ammermueller, A., & Pischke, J. (2009). Peer Effects in European Primary Schools : Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3), 315-348. <https://doi.org/10.1086/603650>
- Anderson, C. J., Kim, J.-S., & Keller, B. (2014). Multilevel modeling of categorical response variables. *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*, 481-519.
- Angrist, J. D., & Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5), 1613-1634. <https://doi.org/10.1257/0002828043052169>
- Benbiga, A., Hanchane, S., Idir, N., & Mostafa, T. (2012). Les déterminants des performances scolaires des élèves marocains. *Maghreb Machrek*, 211, 69-98.
- Björklund, A., & Salvanes, K. G. (2011). Education and family background : Mechanisms and policies. In *Handbook of the Economics of Education* (Vol. 3, p. 201-247). Elsevier.
- Boutchenik, B., & Maillard, S. (2019). Élèves hétérogènes, pairs hétérogènes. *Éducation & formations*, 100, 53. <https://doi.org/10.48464/halshs-02426355>
- Bressoux, P. (2007). L'apport des modèles multiniveaux à la recherche en éducation. *Education didactique*, Vol. 1(2), 73-88.
- Bressoux, P. (2010). *Modélisation statistique appliquée aux sciences sociales*. De Boeck Supérieur. <https://doi.org/10.3917/dbu.bress.2010.01>
- Brodaty, T. (2010). Les effets de Pairs dans l'Éducation : Une Revue de Littérature. *Revue d'économie politique*, Vol. 120(5), 739-757.
- Brodaty, T., & Gurgand, M. (2016). Good peers or good teachers? Evidence from a French University. *Economics of Education Review*, 54, 62-78. <https://doi.org/10.1016/j.econedurev.2016.06.005>
- Burke, M. A., & Sass, T. R. (2013). Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1), 51-82. <https://doi.org/10.1086/666653>
<https://doi.org/10.1086/680173>
- Checchi, D. (2006). *The economics of education : Human capital, family background and inequality*. Cambridge University Press.
- Coleman, J. S. (1966). *Equality of Educational Opportunity Report* (Vol. 2). US Department of Health, Education, and Welfare, Office of Education.
- Dieng, A. M. (2017). *Les performances scolaires en Afrique australe et orientale : Disparités régionales et facteurs déterminants*. Toulon.
- Epple, D., & Romano, R. E. (1998). Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *The American Economic Review*, 88(1), 33-62.
- Epple, D., Romano, R., & Sieg, H. (2006). Admission, tuition, and financial aid policies in the market for higher education. *Econometrica*, 74(4), 885-928.
- Evans, W. N., Oates, W. E., & Schwab, R. M. (1992). Measuring Peer Group Effects : A Study of Teenage Behavior. *Journal of Political Economy*, 100(5), 966-991. <https://doi.org/10.1086/261848>
- Givord, P., & Guillerm, M. (2016). Les modèles multiniveaux. In *Documents de Travail de l'Insee—INSEE Working Papers* (N° m2016-05; Documents de Travail de l'Insee - INSEE Working Papers). Institut National de la Statistique et des Etudes Economiques.

- Goux, D., & Maurin, É. (2005). Composition sociale du voisinage et échec scolaire. Une évaluation sur données françaises. *Revue économique*, 56(2), 349-361. Cairn.info. <https://doi.org/10.3917/reco.562.0349>
- Guo, G., Li, Y., Wang, H., Cai, T., & Duncan, G. J. (2015). Peer Influence, Genetic Propensity, and Binge Drinking : A Natural Experiment and a Replication. *American Journal of Sociology*, 121(3), 914-954. <https://doi.org/10.1086/683224>
- Guyon, N., Maurin, E., & McNally, S. (2012). The Effect of Tracking Students by Ability into Different Schools A Natural Experiment. *Journal of Human Resources*, 47(3), 684-721. <https://doi.org/10.3368/jhr.47.3.684>
- Hanchane, S., & Mostafa, T. (2012). Solving endogeneity problems in multilevel estimation : An example using education production functions. *Journal of Applied Statistics*, 39(5), 1101-1114. <https://doi.org/10.1080/02664763.2011.638705>
- Henderson, V., Mieszkowski, P., & Sauvageau, Y. (1978). Peer group effects and educational production functions. *Journal of Public Economics*, 10(1), 97-106. [https://doi.org/10.1016/0047-2727\(78\)90007-5](https://doi.org/10.1016/0047-2727(78)90007-5)
- Hoxby, C. M., & Weingarth, G. (2005). *Taking race out of the equation : School reassignment and the structure of peer effects*.
- Humphreys, M., Sachs, J. D., Stiglitz, J. E. (2007). *Escaping the resource curse*. Columbia University press.
- Issehnane, S., & Sari, F. (2013). Effets contextuels et effets de pairs : Quelles conséquences sur la réussite scolaire ? *Revue Economique*, 64(5), 775-804. <https://doi.org/10.3917/reco.645.0775>
- Jencks, C., & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. *Inner-city poverty in the United States*, 111, 186-222.
- Kim, J.-S., & Frees, E. W. (2007). Multilevel Modeling with Correlated Effects. *Psychometrika*, 72(4), 505-533. <https://doi.org/10.1007/s11336-007-9008-1>
- Lavy, V., Paserman, M. D., & Schlosser, A. (2012). Inside the Black Box of Ability Peer Effects : Evidence from Variation in the Proportion of Low Achievers In the Classroom. *The Economic Journal*, 122(559), 208-237. <https://doi.org/10.1111/j.1468-0297.2011.02463.x>
- Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, 116(3), 777-803. <https://doi.org/10.1162/00335530152466232>
- Leckie, G., Pillinger, R., Jones, K., & Goldstein, H. (2012). Multilevel Modeling of Social Segregation. *Journal of Educational and Behavioral Statistics*, 37(1), 3-30. <https://doi.org/10.3102/1076998610394367>
- Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard University Press.
- Monso, O., Fougere, D., Givord, P., & Pirus, C. (2019). Les camarades influencent-ils la réussite et le parcours des élèves ? *Éducation & formations*, 100, 23-52. <https://doi.org/10.48464/halshs-02426350>
- Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1), 69-85. <https://doi.org/10.2307/1913646>
- Nechyba, T. J. (2003). Centralization, fiscal federalism, and private school attendance. *International Economic Review*, 44(1), 179-204.
- Piketty, T., & Valdenaire, M. (2006). *L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français—Estimations à partir du panel primaire 1997 et du panel secondaire 1995* (p. 153 p.). Ministère de l'éducation nationale.
- Ris, C. (2013). Les inégalités ethniques dans l'accès à l'emploi en Nouvelle-Calédonie. *Economie et Statistique*, 464(1), 59-72. <https://doi.org/10.3406/estat.2013.10229>

- Ris, C., Trannoy, A., & Wasmer, É. (2017). L'économie néo-calédonienne au-delà du nickel. *Notes du conseil d'analyse économique*, 3, 1-12.
- Sacerdote, B. (2001). Peer effects with random assignment : Results for Dartmouth roommates. *The Quarterly journal of economics*, 116(2), 681-704.
- Trognon, A. (2003). L'économétrie des panels en perspective. *Revue d'économie politique*, Vol. 113(6), 727-748.
- Woessmann, L. (2004). *How Equal are Educational Opportunities? Family Background and Student Achievement in Europe and the Us*. Social Science Research Network.
- Zimmerman, D. J. (2003). Peer effects in academic outcomes : Evidence from a natural experiment. *Review of Economics and statistics*, 85(1), 9-23.

PRÉSENTATION DU LARJE (EA 3329)

Le Laboratoire de Recherches Économique et Juridique est le principal centre de recherche calédonien se concentrant sur le droit et l'environnement humain, économique et social insulaire. Le processus d'émancipation progressive du pays place la recherche juridique dans un contexte institutionnel inédit et soulève, pour l'analyse économique, des questions spécifiques en termes de viabilité des activités, comme de mesure des inégalités.

La problématique humaine a une importance essentielle en Nouvelle-Calédonie, au vu de l'histoire de son peuplement et de l'instabilité de ses statuts successifs. L'égalité des différents statuts civils, le rapport juridique complexe au sol, l'adaptation du droit du travail à la société océanienne, l'enchevêtrement des normes et des compétences juridiques, la question du droit constitutionnel calédonien, la citoyenneté ou encore le statut des peuples autochtones, forment des pistes de recherche particulières à la Nouvelle-Calédonie. Il est également temps de faire un bilan des Accords de Matignon (1988) et de Nouméa (1998). L'optique est de mesurer le rééquilibrage, qui est le critère essentiel d'évaluation de l'action publique. La voie d'un développement soutenable nécessite d'assurer l'intégration de l'ensemble des communautés au système d'échanges économiques, afin d'assurer une répartition des bénéfices de la croissance et de la rente minière. En raison de l'existence des grandes réalisations industrielles, les problématiques de la mine, du nickel, de la protection de la biodiversité terrestre et maritime comme du droit de l'urbanisme et de la construction sont fortement privilégiées.

L'autre champ d'analyse proposé concerne les relations économiques de la Nouvelle-Calédonie avec la Métropole. La question des transferts financiers, au travers de l'étude des finances publiques et de la fiscalité, est essentielle. L'émancipation prévue par l'Accord de Nouméa pose la question des ressources propres que la Nouvelle-Calédonie devra générer. L'étude des réformes fiscale et monétaire découle de ce questionnement.

L'équipe entend favoriser une recherche non seulement fondamentale, mais aussi appliquée, sur les aspects économiques et juridiques des politiques locales liées aux activités humaines et industrielles en Nouvelle-Calédonie. Dans cette optique, les membres de l'équipe favorisent les ouvertures internationales par des collaborations avec les universités anglo-saxonnes de la zone, les institutions régionales (AUF, CPS, PROE) ou les agences de moyens (CNRT, GOPS).

Le LARJE édite une série de Working-Papers (Les Cahiers du LARJE), et organise régulièrement des séminaires et des manifestations nationales et internationales. Le site Web (<http://larje.unc.nc>) est une source de documentation très riche, qui offre, par mots-clés, le libre accès aux résultats de la recherche, comme aux débats de société qui traversent la Nouvelle-Calédonie et l'Outre-mer.

© 2018 Publication du LARJE

Laboratoire de recherches juridique et économique

Tous droits réservés / All rights reserved

No part of this publication may be reproduced in any form

or by any means without the written permission

of the University of New Caledonia

Avenue James Cook – BP R4
98851 - Nouméa CEDEX - www.unc.nc

LES DERNIERS CAHIERS DU LARJE

<https://larje.unc.nc/fr/recherches/cahiers-du-larje/tous/>

- Cahier n° 2020-4 : Intégration de la population autochtone dans l'économie de marché autour de l'usine du Nord : une autre voie économique est-elle possible ?, Samuel Gorohouna
- Cahier n° 2020-3 : Pour un réexamen du triptyque « décolonisation, indépendance, souveraineté » au prisme de l'expérience néocalédonienne, Séverine Blaise, Carine David, Gerard Prinsen
- Cahier n° 2020-2 : Seeking for tipping point in the housing market: evidence from a field experiment, Sylvain Chareyron, Samuel Gorohouna, Yannick l'Horty, Pascale Petit, Catherine RIS
- Cahier n° 2020-1 : Rééquilibrage - Politique de réduction des inégalités en Nouvelle-Calédonie. Le rôle des politiques éducatives, Catherine RIS
- Cahier n° 2019-10 : Quel avenir institutionnel partagé pour la Nouvelle-Calédonie ?, Mathias Chauchat
- Cahier n° 2019-9 : Auto-détermination et géographie électorale en Nouvelle-Calédonie : cristallisation politique ou indépendance ?, Pierre-Christophe Pantz
- Cahier n° 2019-8 : L'ONU au service du processus d'émancipation de la Nouvelle-Calédonie, Caroline Gravelat
- Cahier n° 2019-7 : Le nickel, vecteur d'indépendance ou de dépendances ?, Jeanne Adrian
- Cahier n° 2019-6 : Les déterminants du comportement électoral au cours du référendum du 4 novembre 2018 en Nouvelle-Calédonie, Sylvain Brouard, Samuel Gorohouna, Anthony Tutugoro
- Cahier n° 2019-5 : La délicate adaptation de la tradithérapie kanak au système de santé calédonien : essai de synthèse sur les difficultés au plan de la coutume autochtone et du droit français, Antoine Leca
- Cahier n° 2019-4 : Intégration dans l'économie de marché autour de l'usine du nord : une nouvelle voie économique kanak ?, Samuel Gorohouna
- Cahier n° 2019-3 : Formation des équipes enseignantes, décrochage et réussite scolaires, Amélie Chung, Pierre-Yves Le Roux, Olivier Galy, Catherine Ris
- Cahier n° 2019-2 : Théorie partenariale de la gouvernance, Bernard Grand, Philippe Grill
- Cahier n° 2019-1 : Profit vs morality, results from a survey experiment on discrimination, Mathieu Bunel, Elisabeth Tovar
- Cahier n° 2018-6 : Veille et éclairage juridiques, Christine Bidaud-Garon
- Cahier n° 2018-5 : Vous ne dormirez pas chez moi ! Tester la discrimination dans l'hébergement touristique, Mathieu Bunel, Yannick l'Horty, Souleymane MBaye, Loïc du Parquet et Pascale Petit

© 2018 Publication du LARJE

Laboratoire de recherches juridique et économique

Tous droits réservés / All rights reserved

No part of this publication may be reproduced in any form

or by any means without the written permission

of the University of New Caledonia

Avenue James Cook – BP R4
98851 - Nouméa CEDEX - www.unc.nc