



**HAL**  
open science

## Les camarades influencent-ils la réussite et le parcours des élèves ? Une revue de littérature sur les effets de pairs dans l'enseignement primaire et secondaire

Olivier Monso, Denis Fougère, Pauline Givord, Claudine Pirus

### ► To cite this version:

Olivier Monso, Denis Fougère, Pauline Givord, Claudine Pirus. Les camarades influencent-ils la réussite et le parcours des élèves ? Une revue de littérature sur les effets de pairs dans l'enseignement primaire et secondaire. 2019. hal-03456953

**HAL Id: hal-03456953**

**<https://hal-sciencespo.archives-ouvertes.fr/hal-03456953>**

Preprint submitted on 30 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SciencesPo

LABORATOIRE INTERDISCIPLINAIRE  
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

Avril 2019, n°86

## Les camarades influencent-ils la réussite et le parcours des élèves ?

*Une revue de littérature sur les effets de pairs  
dans l'enseignement primaire et secondaire*

**Olivier MONSO**

MENJ-DEPP (sous-direction des synthèses), Sciences Po (OSC, LIEPP)  
[olivier.monso@education.gouv.fr](mailto:olivier.monso@education.gouv.fr)

**Denis FOUGERE**

CNRS, Sciences Po (OSC, LIEPP)  
[denis.fougere@sciencespo.fr](mailto:denis.fougere@sciencespo.fr)

**Pauline GIVORD**

Insee (SSP Lab), Crest  
[pauline.givord@m4x.org](mailto:pauline.givord@m4x.org)

**Claudine PIRUS**

MENJ-DEPP (bureau des études statistiques sur les élèves)  
[claudine.pirus@education.gouv.fr](mailto:claudine.pirus@education.gouv.fr)

[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)

© 2019 by the authors. All rights reserved.

Comment citer cette publication:

Olivier MONSO, Denis FOUGERE, Pauline GIVORD, Claudine PIRUS, Les camarades influencent-ils la réussite et le parcours des élèves ?, *Sciences Po LIEPP Working Paper* n°86, 2019-04-04.

# **Les camarades influencent-ils la réussite et le parcours des élèves ?**

## **Une revue de littérature sur les effets de pairs dans l'enseignement primaire et secondaire**

Olivier Monso, Denis Fougère, Pauline Givord, Claudine Pirus

### **Résumé**

*En éducation, les effets de pairs résultent des différents types d'interactions entre élèves, au sein d'une même classe ou d'un même établissement. Toutefois, caractériser la nature et mesurer l'influence de ces interactions pose des problèmes méthodologiques substantiels. Ce document vise à présenter les difficultés relatives à la mesure des effets de pairs en éducation, ainsi que les résultats des recherches qui leur ont été consacrées dans l'enseignement primaire et secondaire. Au sein d'un établissement, les élèves sont influencés par la composition socio-économique et le niveau scolaire de leurs pairs. Les élèves de milieu défavorisé, ou en difficulté scolaire, y sont en général plus sensibles. En raison de tels effets, les phénomènes de ségrégation peuvent aggraver les inégalités scolaires. Les résultats des recherches relatives aux effets de pairs ne sont toutefois pas convergents.*

**Mots-clés :** effets de pairs, économétrie, méthodes d'évaluation

## Introduction

Mesurer et comprendre l'influence qu'ont nos collègues de travail, nos amis, nos voisins, sur nos comportements et notre parcours socio-professionnel, est une thématique qui concerne plusieurs champs disciplinaires, et qui a fait l'objet de nombreuses recherches<sup>1</sup>. Par ailleurs, cette thématique soulève des problèmes de méthodologie statistique importants. En effet, s'il est facile de mettre en évidence une corrélation entre le comportement d'une personne (en termes de consommation, de parcours scolaire...) et celui de ses collègues, amis ou camarades de classe, il est bien plus difficile d'en déduire une causalité, notamment parce que la formation des groupes de pairs est rarement le fait du hasard. Comprendre cette difficulté, et y apporter des réponses méthodologiques, est toutefois nécessaire en raison des enjeux de politique publique. Le domaine de l'éducation illustre pleinement la nature de ces enjeux : aux politiques et aspirations visant à promouvoir la mixité sociale et scolaire font écho les interrogations récurrentes quant à leur efficacité pour les élèves concernés. La thématique des effets de pairs évoque également la question de la constitution de classes de niveau dans les établissements, notamment dans les collèges.

Ce document actualise des travaux de synthèse antérieurs consacrés aux effets de pairs en éducation (Brodaty, 2010 ; Sacerdote, 2011). Il doit, tout d'abord, permettre d'exercer un regard critique sur la littérature relative aux effets de pairs. Il doit, ensuite, permettre de dégager les principales convergences et divergences des résultats empiriques, ainsi que les points qui sont à ce jour peu éclairés par ces recherches, notamment dans le cas français. Nous commencerons par définir et présenter les principaux types d'effets de pairs. Puis nous aborderons les difficultés méthodologiques liées à leur identification et nous présenterons les principales méthodes statistiques utilisées pour mesurer ces effets. Enfin, nous présenterons un aperçu des principaux résultats des recherches internationales conduites sur ce thème, en mettant plus particulièrement l'accent sur les quelques travaux français disponibles.

---

<sup>1</sup> Ce document tire parti des échanges tenus dans le cadre du groupe de travail « effets de pairs et de composition » entre la DEPP et l'Insee entre 2016 et 2019. Nous remercions ses participants et en particulier Cédric Afsa (alors sous-directeur des synthèses à la DEPP), Linda Ben Ali, Paul-Olivier Gasq, Fabrice Murat (DEPP), ainsi que Béatrice Boutchenik et Sophie Maillard (Insee), qui ont animé des séances de ce groupe. Ce projet a bénéficié du soutien apporté par l'ANR et l'État au titre du programme d'Investissements d'avenir dans le cadre du labex LIEPP (ANR-11-LABX-0091, ANR-11-IDEX-0005-02).

## I. Les effets de pairs : de quoi parle-t-on ?

Les effets de pairs correspondent aux effets résultant des interactions entre individus : par l'intermédiaire de leurs caractéristiques et leur comportement, des personnes situées dans un environnement commun (habitant dans le même quartier, scolarisées dans la même classe...) s'influencent mutuellement. La notion d'« effets de pairs » est donc inséparable de celle de « groupe de pairs », qui, dans ce document, correspond en général aux élèves faisant partie de la même classe ou du même établissement scolaire<sup>2</sup>.

### 1.1 Définition des effets de pairs en éducation

Dans le champ de l'éducation, un effet de pairs correspond à l'influence des caractéristiques des camarades, ou de leur comportement, sur le résultat et le comportement des élèves. Cette définition, assez large, inclut selon Sacerdote (2011) des interactions de nature très diverse :

- les effets transitant par les interactions entre élèves, par exemple si des élèves plus performants font bénéficier les autres élèves de leurs acquis ;
- les effets transitant par les interactions avec l'enseignant ; par exemple, la présence d'élèves plus performants peut conduire ce dernier à accroître ses exigences ;
- ou encore les effets transitant par les familles ; c'est par exemple le cas si les parents d'élèves exercent une influence sur le fonctionnement de l'établissement.

Depuis l'article pionnier de Manski (1993), on distingue traditionnellement deux grands types d'effets de pairs :

- les « **effets endogènes** » (*endogenous effects*), correspondant à la façon dont le comportement d'un individu est influencé par celui de ses pairs ; de tels effets résultent des interactions directes entre pairs ;
- les « **effets exogènes** » (ou contextuels) (*exogenous/contextual effects*), correspondant à la façon dont le comportement d'un individu est influencé par les caractéristiques propres de ses pairs, leur origine sociale par exemple.

Pour Manski (1993), ces effets de pairs doivent être distingués de ce qu'il nomme les « **effets corrélés** ». Ces derniers traduisent le fait qu'au sein d'un groupe de pairs, les individus sont soumis au même environnement, qui exerce une influence sur eux, ou partagent des caractéristiques communes qui ne relèvent pas à proprement parler d'effets de pairs. Ainsi, des élèves d'une même classe de collège ont les mêmes enseignants, et ces derniers peuvent avoir un effet conjoint sur la

---

<sup>2</sup> Certains travaux font un compromis entre ces deux choix en considérant comme pairs de l'élève tous les élèves scolarisés dans le même niveau, le même établissement et la même année que lui (par exemple les élèves de sixième d'un collège donné, une année donnée).

progression de l'ensemble des élèves. Ces effets corrélés peuvent également recouvrir des effets de sélection dans les classes ou les établissements. Par exemple, si les élèves sont sélectionnés dans un établissement en fonction de leurs résultats scolaires précédents, les niveaux scolaires à l'issue de la scolarité des élèves de cet établissement seront sans doute très corrélés. On aurait néanmoins tort de considérer cette corrélation comme résultant des interactions entre élèves. L'existence d'effets corrélés traduit la difficulté pour l'économiste à identifier les effets de pairs séparément des autres mécanismes exerçant une influence sur la scolarité.

Ces notions sont courantes dans les travaux des économistes. Ce document traite essentiellement des recherches sur les effets de pairs menées par les économistes. Toutefois, elles s'inscrivent dans les travaux plus larges des sciences de l'éducation sur les effets de pairs et de contexte. Les notions utilisées diffèrent selon les disciplines mais aussi le contenu des phénomènes qu'elles désignent. Un éclairage sur ces notions est présenté dans l'encadré 1.

Les effets endogènes et exogènes sont liés ; toutefois, leurs implications pour les politiques éducatives diffèrent. Les effets exogènes dépendent directement de la composition de la classe ou de l'établissement. Ils ont donc une importance toute particulière pour les politiques de mixité sociale ou scolaire. Les caractéristiques exogènes des pairs sont de nature diverse : si la plupart des travaux privilégient le niveau scolaire des pairs (être entouré de camarades de niveau<sup>3</sup> plus ou moins élevé), d'autres s'intéressent à leurs caractéristiques socio-économiques (élèves de milieu social plus ou moins favorisé, d'origine étrangère...). Or, les résultats et interprétations portant sur ces différentes formes de mixité ne se recoupent pas forcément, car les dimensions en question, même si elles sont assez fortement corrélées (origine sociale et niveau scolaire par exemple), ne se définissent pas de la même façon.

De leur côté, les effets endogènes traduisent l'importance du mécanisme appelé « **multiplicateur social** » (Manski, 1993) : les pairs influencent leurs camarades et, dans un second temps, ils bénéficient eux-mêmes de la hausse générale du niveau de la classe, par un effet boule de neige. L'existence d'effets endogènes traduit le fait que tout facteur améliorant la réussite de certains élèves (par exemple, une intervention éducative ciblée) se transmet aux autres élèves, et son impact final, sur l'ensemble des élèves, est supérieur à celui qui serait constaté en l'absence d'interactions (voir l'annexe 1).

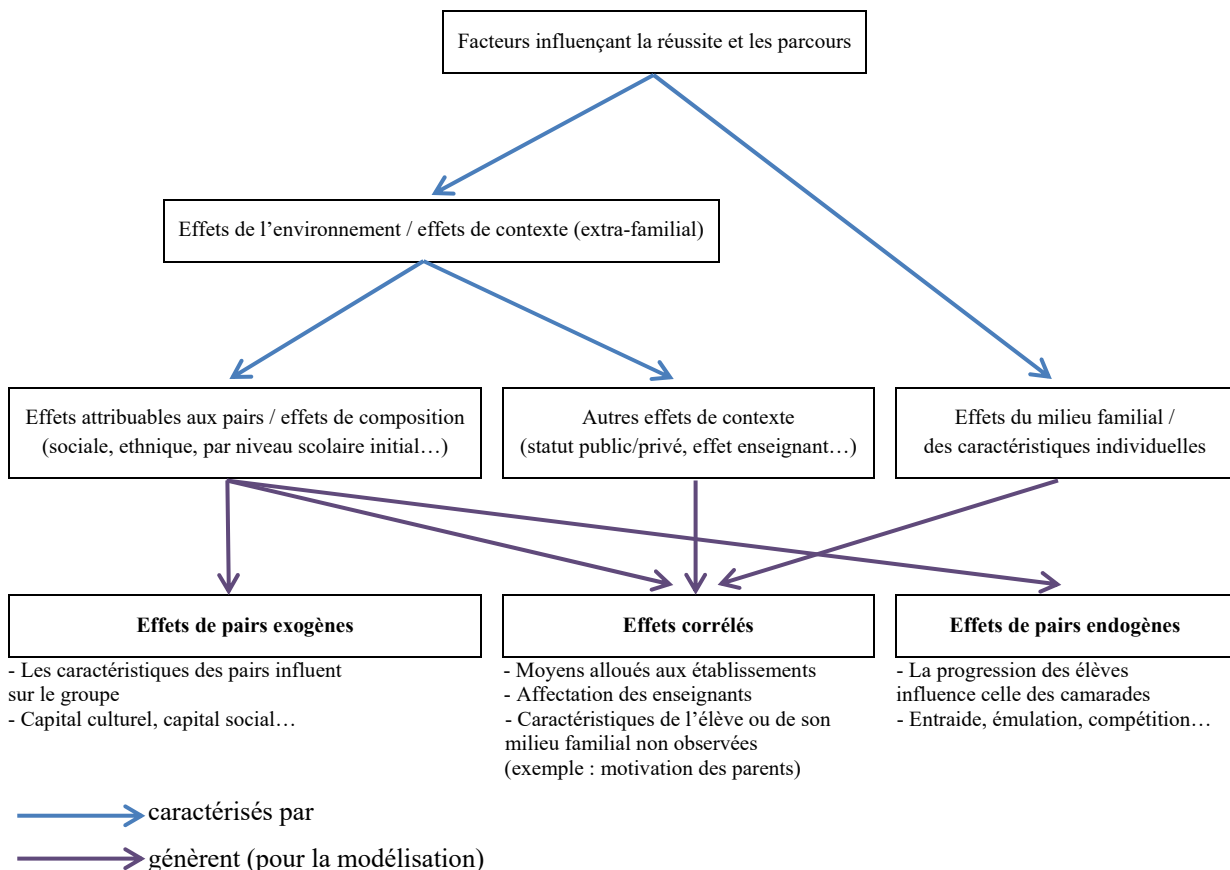
---

<sup>3</sup> Dans le cadre des effets de pairs exogènes, c'est le niveau initial des pairs qui importe, avant interactions avec les autres élèves. Ce niveau est par exemple mesuré suivant des tests réalisés en début d'année ou l'année précédente.

## Encadré 1. Effets de pairs, effets de contexte, effets de composition

Le vocabulaire employé pour étudier les facteurs influençant la réussite et les parcours des élèves est d'une grande diversité, qui renvoie notamment aux choix faits par les chercheurs de différentes disciplines. Cette diversité prête parfois à confusion. La figure 1 ci-dessous peut nous éclairer sur les principaux mécanismes qui entrent en jeu.

Figure 1. Effets de contexte, effets de composition, effets de pairs



Ainsi, l'expression « effets de contexte » désigne l'ensemble des déterminants de la réussite et du parcours de l'élève qui ne sont liés ni à ses caractéristiques individuelles (aptitudes innées, état de santé...), ni à son environnement familial, mais s'inscrivent dans son cadre de vie et de scolarisation. Ils comprennent aussi bien l'influence de ses camarades ou de ses voisins de quartier (effets attribuables aux pairs) que la qualité de ses enseignants, ou encore les caractéristiques de son établissement (taille, statut public ou privé...).

L'expression « effets de composition » désigne une forme particulière d'effets de contexte relative aux camarades qui sont en relation avec ce même élève, le plus souvent dans son établissement ou sa classe. Ces effets prennent en compte les variations du contexte éducatif induites par la composition des établissements et des classes (notamment socio-économique ou en termes de niveau scolaire). Comme le soulignent Dupriez, Monseur et van Campenhout (2012), l'expression « effets de composition » a une acception plus large que celle d'effets de pairs. La taille des classes, le niveau d'expérience des enseignants, parce qu'ils sont en partie liés à la composition des classes (du fait des politiques d'allocation des ressources entre établissements et au sein des établissements), engendrent des effets de composition, mais sont en général exclus par les économistes du champ des effets de pairs (Sacerdote, 2011). Ils constituent des effets corrélés au sens de Manski (1993), c'est-à-dire des effets résultant des caractéristiques des établissements ou des classes qui, sans être des effets de pairs, sont corrélés à la composition du groupe d'élèves accueillis ; ils rendent plus difficile l'identification des effets de pairs.

Du fait de l'existence d'effets endogènes, les individus sont affectés par le résultat de leur groupe en même temps qu'ils y contribuent : c'est le problème de « **réflexion** » (*reflection problem*) soulevé par Manski (1993). Il rend plus complexe l'identification séparée des effets endogènes et exogènes dès lors que ces deux dimensions sont présentes dans l'analyse (voir l'annexe 1).

De façon plus générale, ces deux types d'effets sont difficiles à distinguer, à la fois d'un point de vue méthodologique et de l'interprétation de ces effets. On ne peut guère concevoir l'existence d'effets exogènes sans qu'elle soit associée à des effets endogènes : si un élève tire parti de la proximité de camarades d'un milieu social plus favorisé, c'est sans doute en partie parce que ces camarades ont une meilleure réussite scolaire. En revanche, on peut concevoir l'existence d'effets endogènes pour ce même élève indépendamment des caractéristiques des pairs. Par exemple, une nouvelle méthode d'apprentissage peut avoir des effets immédiats d'ampleur modeste, mais qui seront amplifiés au cours du temps car l'amélioration des résultats de certains élèves bénéficie à leurs camarades (émulation, meilleur climat de classe...). L'étude des effets endogènes conserve donc un intérêt propre, car elle permet notamment de quantifier un canal de diffusion des politiques éducatives, passant par les interactions entre élèves.

En raison de la difficulté à dissocier ces effets, plusieurs études se limitent à l'estimation des seuls effets liés à la composition du groupe de pairs, sans faire apparaître explicitement d'effets endogènes. Ce choix permet une interprétation plus directe des résultats au regard des politiques de mixité scolaire et sociale. Dans ce cas, on parle d'effets exogènes en forme réduite.

La suite de cette revue de littérature est consacrée aux travaux ayant fait ce choix, et donc privilégié les effets exogènes. Ces travaux sont confrontés à deux types de questions méthodologiques qui relèvent, d'une part, des hypothèses relatives à la façon dont l'analyste spécifie ces effets, et, d'autre part, de la possibilité de les identifier de façon robuste.

## **1.2 Spécification des effets de pairs**

Les travaux sur les effets de pairs doivent prendre en compte la variété des configurations que ces derniers peuvent prendre. De prime abord, on peut supposer que le résultat d'un élève dépend, de façon linéaire, des résultats moyens obtenus par ses camarades, ou encore de leurs caractéristiques moyennes. Ces hypothèses définissent la formalisation du modèle dit « linéaire en moyenne » (présentée dans les annexes 1 et 2). Ce modèle a l'avantage de la simplicité, mais peut ne pas rendre compte de manière pertinente de tous les types d'effets de pairs. C'est d'autant plus dommageable que ce modèle ne permet pas de répondre à certaines des questions importantes du débat public sur l'école : par exemple, faut-il favoriser ou au contraire restreindre la possibilité pour les établissements de constituer des classes de niveau, ou encore faut-il accroître la mixité sociale dans les



établissements ? Sous les hypothèses du modèle linéaire en moyenne, les gains obtenus par les uns d'une réallocation compenseront, par construction, exactement les pertes obtenues par les autres élèves. Pour dépasser ces hypothèses trop restrictives, les recherches prennent généralement en compte :

- des effets dits « **non-linéaires** » : les effets de pairs sont alors supposés dépendre du niveau des pairs les plus faibles, les plus forts, ou encore de l'hétérogénéité des niveaux des élèves présents dans la classe, plutôt que de leur valeur moyenne ;
- des effets dits « **hétérogènes** », par exemple si les effets de pairs sont plus prononcés pour les élèves les plus faibles ou les plus forts.

Introduire des effets non-linéaires ou hétérogènes permet de tester un spectre plus vaste de modèles théoriques. Hoxby et Weingarth (2005) ont par exemple proposé huit types de modèles dans lesquels les effets de pairs peuvent se manifester. Ils incluent dans leur typologie le modèle linéaire en moyenne, mais aussi des modèles prenant en compte la non-linéarité, l'hétérogénéité des effets, ou les deux dimensions à la fois<sup>4</sup>.

Tous ces modèles n'ont sans doute pas la même pertinence empirique. Par ailleurs, leurs implications en termes de politique éducative diffèrent fortement. En présence d'élèves de niveau scolaire variable au sein d'un établissement, une question habituellement posée est de savoir s'il faut regrouper les élèves en fonction de leur niveau scolaire (*tracking*), ou au contraire favoriser des classes de niveaux hétérogènes. Afin d'apporter une réponse à ce type de questions, dans des travaux tels que ceux de Hoxby et Weingarth (2005), ou de Boutchenik et Maillard (2018), la formalisation des effets de pairs prend en compte le fait que les classes sont hétérogènes du fait des disparités de niveaux en leur sein, et le fait que ces compositions de classes affectent différemment les élèves selon qu'ils sont d'un niveau scolaire plus ou moins faible.

## II. Comment identifier les effets de pairs ?

Comme nous l'avons vu, un grand nombre de travaux font le choix d'estimer des effets exogènes<sup>5</sup>, autrement dit, les effets que la composition socio-économique ou scolaire d'une classe ou d'un établissement ont sur les résultats des élèves. Les problèmes d'identification restent importants, en raison de l'existence d'effets corrélés (Manski, 1993). Ces derniers traduisent l'influence des variables liées à la fois à la composition du groupe de pairs et aux résultats des élèves. Leur présence

<sup>4</sup> Pour une présentation plus exhaustive des modèles, on peut se référer à Fougère *et alii* (2018) ou plus directement à Hoxby et Weingarth (2005), Epple et Romano (2011) et Sacerdote (2011).

<sup>5</sup> Ou plus précisément exogènes en forme réduite, puisque ce type de modèle inclut la mesure des effets endogènes, sans les identifier pour autant séparément (voir *supra* et annexe 1).

est donc une source de biais statistique lorsqu'il s'agit d'estimer des effets de pairs. Deux types de biais peuvent être distingués.

### 2.1 Identifier les effets de pairs : un exercice difficile

Le premier type d'effets corrélés résulte du fait que les élèves font face, dans leur classe ou dans leur établissement, à un environnement commun, dont les spécificités ne sont pas toutes observables. Cette partie inobservée est présumée corrélée aux variables observées du modèle. C'est par exemple le cas si le système d'affectation des enseignants favorise l'affectation d'enseignants peu expérimentés dans les établissements les moins attractifs. *A contrario*, les établissements plus attractifs et plus favorisés bénéficient alors d'enseignants plus expérimentés. Dans cet exemple, l'effet de pairs associé au fait d'être scolarisé avec des élèves de meilleur niveau scolaire est alors biaisé (à la hausse) par l'absence de prise en compte de l'effet de la qualité des enseignants.

Le second type d'effets corrélés provient du fait que les élèves partagent des caractéristiques communes, et que ces caractéristiques sont présumées corrélées aux variables observées du modèle. Il s'agit d'un « **effet de sélection** ». C'est le cas si le fait, pour un élève, d'être scolarisé dans une classe ou un établissement donné, est associé à des caractéristiques particulières de cet élève ou de sa famille, par exemple des aspirations particulières en matière de parcours scolaire ou encore la constitution de classes de niveau<sup>6</sup>. Dès lors qu'une partie de ces informations reste inobservée et se trouve dans le résidu individuel de l'équation devant être estimée, l'impact associé au fait d'être dans une classe de bon niveau scolaire est à nouveau biaisé.

Comme l'illustrent ces exemples, ainsi que plusieurs travaux empiriques, la suspicion de biais affectant la mesure des effets de pairs tend plutôt à suggérer que ces effets pourraient être surestimés<sup>7</sup> : une estimation naïve de ces effets, ne tenant pas compte du biais d'endogénéité, risque de surévaluer fortement l'importance des pairs sur les résultats des élèves. Certains mécanismes occasionnent toutefois un biais en sens inverse. Par exemple, la mise en œuvre d'une politique d'allocation des moyens ciblée sur les établissements en difficulté, si elle est efficace, améliore la situation des élèves qui y sont scolarisés. Elle réduit par conséquent la corrélation observée entre les résultats des élèves et la composition de la classe ou de l'établissement.

<sup>6</sup> Cette hypothèse est corroborée par des travaux empiriques. Selon Ly et Riegert (2015), la moitié des collèges français tiendraient compte du niveau des élèves lorsque les classes sont constituées (« classes de niveaux », ou s'en approchant).

<sup>7</sup> Ces travaux confrontent une telle estimation naïve à ceux issus de méthodes d'identification plus poussées (voir par exemple Ammermueller et Pischke, 2009 ; Fougère *et alii*, 2017 ; Gibbons et Telhaj, 2016). Ils mettent en général en évidence que les effets de pairs sont massivement surestimés lorsqu'on ne tient pas suffisamment compte des caractéristiques des élèves et d'environnement jouant, à la fois, sur la composition de leur classe ou établissement, et sur la réussite scolaire.

Plusieurs réponses à ces problèmes d'effets corrélés sont envisageables. La plus évidente, et l'une des plus courantes, consiste à tenir compte de caractéristiques observables des élèves, en nombre suffisamment important, et suffisamment pertinentes, qui sont corrélées à la composition de leur classe ou de leur établissement. Certaines bases de données scolaires, tels que les panels de la DEPP, peuvent être ainsi mobilisées car elles fournissent une description fine de l'environnement familial et scolaire : diplômes et revenus des parents, compétences cognitives et non-cognitives de l'enfant en début de scolarité, caractéristiques et localisation de l'établissement (voir par exemple, Davezies, 2005 ; Fougère *et alii*, 2017). Une autre illustration est fournie dans les travaux menés à partir des évaluations internationales *PISA*, qui, grâce aux questionnaires proposés aux élèves et aux établissements, recueillent des données d'une grande variété, dont certaines de nature subjective<sup>8</sup> ; ces données ont été mobilisées dans plusieurs recherches (par exemple, Rangvid, 2007 ; Schneeweis et Winter-Ebmer, 2007).

Toutefois, cette solution est souvent perçue comme insuffisante, car les sources de biais sont potentiellement nombreuses et difficiles à mesurer. Deux autres grands types de méthodes sont fréquemment adoptées pour essayer de réduire ces biais<sup>9</sup>.

## 2.2 Les méthodologies fondées sur des sources de variation exogène

La première approche consiste à trouver une source de variation exogène affectant la composition des établissements et des classes, autrement dit, un événement qui influence cette composition sans avoir d'effet direct sur les résultats des élèves.

La source exogène la plus évidente est celle fournie par une procédure expérimentale, c'est-à-dire la situation où les élèves et leurs enseignants sont affectés de façon aléatoire aux classes. Par exemple, Duflo, Dupas et Kremer (2011) ont conduit une expérimentation contrôlée dans 121 écoles primaires kenyanes. Dans une moitié des écoles tirées au sort, les élèves ont été répartis entre deux classes de CP en fonction de leur score obtenu à un test passé en début d'année scolaire, c'est-à-dire dans des classes de niveaux (ces élèves constituent ce qu'il est habituel d'appeler le groupe de traitement). Dans l'autre moitié, les élèves ont été affectés aléatoirement entre deux classes de CP de niveaux hétérogènes (ces élèves constituent le groupe de contrôle). Dans chaque école, les enseignants ont été affectés aléatoirement à chaque classe de CP.

---

<sup>8</sup> Par exemple la perception des enseignants par les chefs d'établissement et les élèves (motivation, qualité des relations avec les élèves...).

<sup>9</sup> Un autre type de méthodes, moins fréquemment utilisées, est basé sur la comparaison des variances des résultats entre les classes (Graham, 2008). En effet, l'existence d'effets de pairs engendre un surcroît de variance interclasses qu'il est possible d'utiliser pour les identifier.

À défaut de cette possibilité, les économètres s'appuient parfois sur des expériences dites « naturelles », encore appelées « quasi-expériences ». Il s'agit de situations où la composition des établissements et des classes a varié d'une façon qui peut être considérée comme exogène, autrement dit, sans lien avec les caractéristiques inobservées des élèves et des classes. Ces situations ont pu être créées par un événement exceptionnel. Par exemple, Imberman, Kugler et Sacerdote (2012) ont étudié l'impact de l'arrivée massive, dans des écoles du Texas et de Louisiane, d'élèves évacués de la Nouvelle-Orléans suite aux ouragans Katerina et Rita en 2005. Ces élèves ont été placés dans les écoles et les classes d'une manière qui peut être considérée comme aléatoire. Étant majoritairement issus de milieux sociaux défavorisés, leur arrivée a engendré une modification de la composition des écoles d'accueil. Cette variation exogène est utilisée pour identifier des effets de pairs. Hoxby et Weingarth (2005) ont pour leur part exploité un changement de politique de réaffectation des élèves entre les établissements scolaires dans le Comté de Wake (Caroline du Nord). Les nouveaux critères déterminant l'affectation des élèves aux établissements (par exemple, la prise en compte du revenu des parents et non plus d'un critère ethnique) ont modifié la composition des établissements scolaires ; cette nouvelle procédure d'affectation est là encore considérée comme une source de variation exogène.

Dans d'autres situations, les économètres ne s'appuient sur aucun événement exceptionnel, mais isolent des situations où l'affectation des élèves aux classes est supposée être aléatoire, ce qui peut encore s'assimiler à une forme d'expérience naturelle. Par exemple, Ly et Riegert (2016) s'intéressent aux effets des pratiques de composition des classes en seconde générale et technologique sur le parcours des élèves français. Ils se restreignent à des situations très spécifiques (1 % de leur échantillon de départ) dans lesquelles deux élèves d'un même lycée sont originaires d'un même collège, présentent des niveaux scolaires et des caractéristiques sociodémographiques similaires (âge, sexe, origine sociale, niveau scolaire mesuré par leurs notes à l'examen du diplôme national du brevet). Dans ce cas, on peut présumer que le chef d'établissement ne dispose pas d'autres informations sur ces élèves, et la répartition qu'il opère entre les classes de son lycée (en choisissant de laisser ces deux élèves dans la même classe ou de les séparer) peut être considérée comme aléatoire, conditionnellement à ces caractéristiques.

La régulation des systèmes de choix d'établissements par des procédures centralisées offre également des sources d'identification. Dans ce type de système, qui est utilisé par exemple dans l'académie de Paris, les élèves sont amenés en fin de collège à classer les lycées en fonction de leurs préférences (voir par exemple, Fack et Grenet, 2016, pour une description de la procédure). Les algorithmes utilisés visent à affecter l'ensemble des élèves en fonction de leurs préférences – mais également en fonction de celles des lycées. Compte tenu des contraintes de place, certains lycées parmi les plus

réputés ne peuvent accueillir toutes les candidatures. Les élèves sont affectés selon des règles de priorité, qui dépendent en particulier de leurs résultats scolaires antérieurs, et d'autres critères (par exemple le fait d'être boursier à Paris). Cette règle crée mécaniquement des discontinuités dans les procédures d'affectation : un bon niveau scolaire augmente la probabilité d'être affecté dans un très bon lycée, mais à un certain seuil, la capacité d'accueil du lycée est atteinte. En pratique, des élèves ayant des résultats presque identiques à la fin du collège vont se retrouver, pour certains, dans un lycée élitiste – et d'autres dans des lycées d'un niveau moyen plus faible. On peut alors mesurer les effets liés à la fréquentation de pairs d'un très bon niveau, en comparant les résultats d'élèves dont les résultats sont initialement très proches. Dorénavant, certains côtoient des pairs de niveau plus faible, les autres sont scolarisés avec de meilleurs élèves, tout ceci en raison de la part d'arbitraire caractérisant la procédure d'affectation (être juste au-dessus ou juste en dessous du seuil d'éligibilité). Ce type d'identification, dit « régression avec discontinuité », a été fréquemment utilisé dans la période récente du fait de l'utilisation de ce type d'affectation dans les établissements. C'est le cas des travaux de Landaud, Ly et Maurin (2018) pour les lycées parisiens, d'Abdulkadiroğlu, Angrist et Pathak (2014) et de Dobbie et Fryer (2014) pour ceux de Boston et New-York, mais également d'Hoekstra, Mouganie et Wang (2018) en Chine, Pop-Eleches et Urquiola (2013) en Roumanie, Jackson (2010) pour Trinidad et Tobago, Lucas et Mbiti (2014) au Kenya.

L'existence de variables instrumentales est une autre source de variation exogène. La procédure consiste à utiliser pour estimer un modèle linéaire un instrument qui est corrélé à la variable explicative (la composition des établissements ou des classes) mais qui n'est pas corrélé au terme d'erreur de l'équation (voir l'annexe 2). Toute la difficulté réside dans le choix de cette variation exogène permettant de limiter l'influence d'éventuels facteurs inobservés. Davezies (2005) et Goux et Maurin (2007) utilisent par exemple le mois de naissance comme variable instrumentale<sup>10</sup>. En s'appuyant sur la relation entre le niveau scolaire de l'élève et son mois de naissance, ils observent que plus la part d'élèves appartenant au groupe des pairs nés en début d'année est importante, plus le niveau scolaire moyen de ce groupe est élevé. Les recherches utilisant cet instrument font l'hypothèse que cette corrélation traduit uniquement celle existant entre les acquis scolaires de chaque camarade de l'élève et son mois de naissance, et non l'influence de caractéristiques inobservées. Le choix d'une école et la constitution des classes sont ainsi supposés sans lien direct avec le mois de naissance.

Ces méthodes sont confrontées à deux limites. La première est qu'il est difficile de trouver une variation véritablement exogène de la composition d'une classe ou d'un voisinage. Même dans le cas d'une expérimentation, la validité de cette hypothèse dépend étroitement des conditions de sa mise

---

<sup>10</sup> Ces travaux, sur données françaises, ont appliqué cette méthodologie à la mesure des effets de pairs endogènes, mais on pourrait tout aussi bien l'appliquer à la mesure d'effets exogènes.

en œuvre<sup>11</sup> (Fougère, 2010). La seconde limite de ces méthodes est que, dans de nombreux cas, l'identification provient d'une situation très particulière, un événement exceptionnel, une expérimentation sur un territoire ciblé, ou encore lorsqu'on se restreint à un sous-échantillon des données. Dans tous ces cas, il est possible que les résultats dépendent de la situation particulière qui a été prise en compte, sans garantie de pouvoir généraliser les résultats à l'ensemble de la population.

### 2.3 Les modèles incluant des effets fixes

Le second type d'approche consiste à introduire dans le modèle des effets fixes, représentant soit les élèves, soit les enseignants ou bien les établissements. Ces effets fixes sont supposés capter les effets de sélection. L'estimation vise alors à les neutraliser, par exemple en ne considérant que les variations de résultat pour un même élève au cours du temps, ou encore entre les classes d'un même établissement (voir par exemple, Ammermueller et Pischke, 2009 ; Hanushek, Kain et Rivkin, 2009 ; Burke et Sass, 2013 ; Gibbons et Telhaj, 2016). Ce type de méthodes est relativement fréquent dans les travaux de recherche, compte tenu de la difficulté à trouver des expériences naturelles (voir le modèle avec des effets d'établissement qui est présenté dans l'annexe 2).

Plusieurs travaux s'appuient sur l'estimation d'effets fixes spécifiques aux établissements. Par exemple, Ammermueller et Pischke (2009) utilisent les données de l'évaluation internationale en lecture *Progress in International Reading Literacy Study (PIRLS)* dans six pays européens. Le score en lecture en fin de quatrième année obligatoire (CM1 pour la France) est mis en relation avec l'origine sociale des camarades de classe, cette variable étant construite à partir du nombre de livres possédés par les familles. Lorsque l'on se limite aux écoles ayant au moins deux classes, l'identification des effets de pairs résulte de la variation de la composition sociale des classes au sein d'une même école. Les auteurs font donc l'hypothèse que la répartition des élèves et des ressources entre les classes est aléatoire, avec pour argument que la constitution de classes de niveaux serait peu fréquente en primaire. Ils justifient cette hypothèse par un examen statistique de la composition des classes et de la répartition des ressources<sup>12</sup>. La recherche de Boutchenik et Maillard (2018) constitue un autre exemple d'application d'effets fixes d'établissement ; cette étude a pour originalité d'exploiter des échantillons de grande taille (Fichiers anonymisés d'élèves pour la recherche et les études, Faere) composés d'établissements observés pendant plusieurs années. Enfin, Gibbons et

---

<sup>11</sup> On pense par exemple à la perte d'informations (attrition...), aux refus de participer à l'expérimentation ou encore à l'effet de divers comportements, par exemple si les parents changent l'enfant d'école lorsqu'ils ne sont pas satisfaits de l'affectation qu'il a reçue. Toutes ces situations font qu'une affectation des élèves effectuée sur une base *a priori* aléatoire ne l'est pas forcément *a posteriori*.

<sup>12</sup> Cet examen concerne, d'une part, la corrélation entre les caractéristiques sociodémographiques de l'élève et l'affectation à une classe de l'école, et d'autre part, entre les caractéristiques des classes et l'allocation des ressources. Dans la plupart des cas, les tests ne rejettent pas l'hypothèse d'indépendance.

Telhaj (2016), utilisant des données longitudinales britanniques, tiennent compte des effets fixes associés aux établissements secondaires, mais aussi aux écoles primaires que les élèves ont préalablement fréquentées. En effet, au sein d'un même collège, les élèves proviennent généralement d'écoles primaires différentes et la probabilité d'avoir fréquenté la même école primaire que tous ses camarades de collège est très faible, ce qui permet d'identifier ces deux types d'effets fixes.

Les modèles à effets fixes sont particulièrement adaptés à l'utilisation de données longitudinales, comme l'illustre également la recherche de Burke et Sass (2013). Leur étude porte sur un échantillon d'élèves scolarisés en Floride dans trois niveaux d'enseignement (primaire, collège, lycée) ; elle permet de mesurer l'évolution temporelle de leur performance. Les effets fixes spécifiques aux élèves permettent de capter une part importante de l'hétérogénéité inobservée affectant les résultats scolaires. Ils sont également utilisés pour définir le niveau d'aptitude moyen des pairs, selon un processus d'estimation itératif. Cette recherche prend aussi en compte des effets fixes spécifiques aux enseignants, en s'appuyant sur le fait qu'un même enseignant fait cours dans plusieurs classes. Ce faisant, les auteurs montrent l'importance de tenir compte de cette seconde dimension. Lorsque les effets de pairs sont estimés sans effets fixes d'enseignant, l'influence du niveau scolaire des pairs est surestimée car les enseignants les plus performants sont plus vraisemblablement affectés aux classes de meilleur niveau scolaire.

Comme on peut le constater, les modèles diffèrent en fonction du type d'effets fixes pris en compte (élèves, enseignants, établissements), le nombre de types d'effets fixes dépendant étroitement de la nature des données disponibles. Or, les résultats sont sensibles aux dimensions introduites dans l'analyse. Enfin, la mise en œuvre de ces méthodes s'appuie sur des hypothèses assez fortes (voir l'annexe 2). Ces hypothèses, si elles peuvent être confortées par certains tests statistiques, ne peuvent pas être validées dans leur intégralité. Cette diversité de méthodes, et la complexité de l'identification des effets de pairs, incite donc à une grande prudence dans la lecture des résultats issus de ces recherches.

### III. Quels sont les principaux résultats empiriques ?

Nous ciblons notre revue de littérature sur les recherches qui portent sur des élèves scolarisés dans le premier et le second degré, et qui abordent la question des effets de composition sociale et scolaire sur la réussite et le parcours des élèves<sup>13</sup>.

#### 3.1 La composition scolaire et socio-économique des camarades a une influence sur les résultats scolaires

La plupart des recherches des économistes concluent que la composition socio-économique ou scolaire des pairs a une influence sur les résultats scolaires des élèves.

Ces effets semblent plus forts à l'échelle de la classe que de l'établissement<sup>14</sup>, comme l'indique la recherche de Burke et Sass (2013) qui propose une comparaison entre ces deux niveaux d'analyse. Dans la plupart des travaux empiriques, les effets de pairs transitent par le niveau scolaire des camarades de classe. Les caractéristiques socio-économiques des pairs ont également une influence. Toutefois, ces deux dimensions sont fortement corrélées. Plusieurs recherches concluent que c'est le niveau scolaire des pairs qui explique avant tout les progressions des élèves ; les caractéristiques socio-économiques des pairs n'apportant qu'un supplément d'information faible, voire non significatif (Hoxby et Weingarth, 2005 ; Gibbons et Telhaj, 2016). D'autres auteurs, comme Hanushek, Kain et Rivkin (2009) mettent toutefois en avant que certaines variables socio-économiques conservent un intérêt propre pour comprendre les effets de pairs, notamment les mécanismes liés à la ségrégation ethnique<sup>15</sup>. Toutefois, les variables de contexte socio-économique présentes dans ces travaux sont souvent en nombre limité.

Lorsque les effets de pairs sont considérés dans leur globalité, c'est-à-dire lorsque ceux-ci sont supposés agir en moyenne de la même façon sur tous les élèves (comme dans le modèle « linéaire en moyenne »), leur ampleur apparaît souvent modeste, mais elle s'accompagne d'une forte hétérogénéité (Sacerdote, 2011). Les quelques études françaises sur la question mettent également en avant des effets modérés de la composition des pairs sur les résultats scolaires. À partir des données du panel d'élèves entrés au CP en 1997, Davezies (2005) constate que le fait d'être scolarisé dans une

<sup>13</sup> Ce choix exclut de notre synthèse les travaux portant sur l'enseignement supérieur. Nous n'abordons pas non plus les recherches sur les effets de voisinage ou de quartier. Les enjeux méthodologiques pour identifier ces effets sont en grande partie similaires à ceux développés dans ce document (Vallet, 2005).

<sup>14</sup> Ce constat fait écho à des travaux français en sciences de l'éducation, constatant que les effets de contexte associés aux classes sont généralement supérieurs à ceux associés aux établissements, en matière de progression des acquis des élèves (Duru-Bellat et Mingat, 1988 ; Bressoux, 2008).

<sup>15</sup> Ces auteurs trouvent qu'une plus forte proportion d'élèves afro-américains dans l'établissement est associée à de moindres résultats scolaires, mais seulement pour les élèves afro-américains. Ce phénomène est difficile à comprendre si l'on mobilise seulement le niveau scolaire des pairs. Il peut s'expliquer par les différents types d'interactions au sein de chaque groupe ethnique.



école parmi les plus favorisées socialement, plutôt que dans les écoles les plus défavorisées<sup>16</sup>, augmente le score de l'élève en sixième d'environ 5 points sur 100 en mathématiques, et de 2 à 3 points en français aux évaluations nationales de début de sixième. L'effet reste modeste si l'on considère qu'il s'applique aux cinq années de la scolarité élémentaire. Piketty (2004) et Piketty et Valdenaire (2006) obtiennent un ordre de grandeur similaire sur les scores obtenus aux évaluations nationales de CE2. Ils jugent cet effet faible au regard de celui d'une réduction de la taille des classes. Selon eux, « une taille de classe réduite de deux élèves a plus d'effet sur la réussite scolaire que de passer d'une classe avec 0 % d'enfants (de milieux) favorisés à 100 % d'enfants (de milieux) favorisés ».

Fougère *et alii* (2017) trouvent que la composition des classes de collège a des effets modestes et parfois non-significatifs, qu'ils soient mesurés selon la part de camarades de classe étrangers, de milieu social très favorisé, ou ayant redoublé. Enfin, en terminale générale et technologique (GT), Boutchenik et Maillard (2018) estiment l'influence sur les notes au baccalauréat du niveau scolaire des camarades de classe, ce dernier étant mesuré par leurs notes aux épreuves écrites du diplôme national du brevet (DNB). Pour l'ensemble des élèves, une augmentation d'un point de la note moyenne des camarades de terminale au DNB (exprimée de 1 à 100) n'améliore que de 0,083 point le résultat de l'élève considéré.

Les estimations des effets de pairs publiées dans la littérature scientifique se distinguent donc avant tout par leur forte hétérogénéité. Ils sont assez souvent d'ampleur modeste, et ce dernier constat est notamment vérifié dans les rares travaux français sur le sujet, à l'école, au collège et au lycée.

### **3.2 Les élèves d'origine défavorisée ou de faible niveau scolaire semblent être plus sensibles aux effets de pairs**

Toutefois, plusieurs recherches suggèrent que ce résultat en moyenne masque la diversité des effets pour des élèves de niveaux différents. Dans plusieurs études, les élèves de niveaux plus faibles, ou dont le contexte familial est moins favorable à la réussite scolaire, sont les plus sensibles à la composition de leur établissement ou de leur classe. C'est notamment le cas de l'étude de Davezies (2005) sur données françaises. Être scolarisé dans une école parmi les plus favorisées socialement, plutôt que parmi les plus défavorisées, produit un impact deux à trois fois supérieur pour des élèves qui sont parmi les plus faibles que pour ceux les plus performants à l'entrée au CP<sup>17</sup>. Dans la recherche

<sup>16</sup> Le classement des écoles est le résultat d'une typologie appliquée à la composition sociale de leurs élèves de CM2, cette variable n'étant pas disponible pour une année antérieure. Le groupe des écoles socialement les plus favorisées et les moins favorisées représente chacun un tiers des élèves.

<sup>17</sup> Les élèves les plus faibles correspondent aux élèves des deux premiers déciles des scores mesurés à l'entrée au CP, les élèves les plus forts aux deux derniers déciles.

menée par Boutchenik et Maillard (2018) sur les effets de la composition scolaire des classes de terminale, cette asymétrie est plus marquée encore. L'effet d'une hausse du niveau scolaire des pairs est cinq fois supérieur pour les élèves initialement les plus faibles<sup>18</sup>.

Les recherches internationales confirment en partie ce constat : l'influence des pairs sur les élèves les plus faibles est souvent soit du même ordre, soit plus forte que pour les autres élèves. À l'aide des données danoises des évaluations *PISA* sur des élèves de 15 ans, Rangvid (2007) constate que la composition sociale de l'établissement a un effet plus fort sur les élèves les plus faibles en lecture, mais pas en mathématiques. Avec les données autrichiennes de *PISA*, Schneeweis et Winter-Ebmer (2007) trouvent un résultat similaire pour les élèves d'origine sociale défavorisée. Dans une étude consacrée à ce type d'effets dans six pays européens, Ammermueller et Pischke (2009) trouvent toutefois que les effets de la composition sociale de la classe sur les résultats en lecture ne sont pas différenciés selon le milieu social de l'élève dans le premier degré.

Par ailleurs, plusieurs recherches consacrées à l'effet du niveau scolaire des pairs soulignent qu'être scolarisé dans un établissement ou dans une classe dont le niveau moyen est plus élevé bénéficie aux élèves les plus faibles, mais jusqu'à un certain point seulement. Pour Hoxby et Weingarth (2005), en fin d'enseignement élémentaire et au collège, les 10 % d'élèves situés au bas de la distribution des scores initiaux bénéficient de l'apport d'élèves qui sont d'un niveau légèrement supérieur au leur, mais pâtissent d'une élévation trop forte du niveau de la classe. Burke et Sass (2013) parviennent à un résultat similaire : une hausse de la part des très bons élèves (i.e., ceux situés dans le dernier quintile de la distribution des scores) est préjudiciable aux élèves les plus faibles (ceux dont les scores sont situés dans le premier quintile), alors qu'elle bénéficie aux élèves moyens. Avec un échantillon de collégiens britanniques, Lavy, Silva et Weinhardt (2012) obtiennent des constats différents selon le genre de l'élève. Les filles semblent bénéficier de la présence de très bons élèves dans la classe, surtout si elles sont d'un niveau scolaire initialement faible, alors que la présence de camarades de très bon niveau n'a guère d'effet sur les garçons<sup>19</sup>.

Lorsque l'on considère les effets de pairs affectant les élèves d'un bon niveau scolaire, les résultats sont très divergents. Hoxby et Weingarth (2005) et Burke et Sass (2013) trouvent que la présence de camarades de haut niveau scolaire bénéficie le plus nettement aux élèves qui sont eux-mêmes d'un haut niveau scolaire. Dans la recherche menée par Boutchenik et Maillard (2018), la présence dans

---

<sup>18</sup> Les élèves les plus faibles correspondent au premier quartile des notes aux épreuves écrites du DNB, les élèves les plus forts au dernier quartile.

<sup>19</sup> Toutefois, les études différenciant les effets par genre peuvent être tributaires des caractéristiques spécifiques des systèmes scolaires. Si Jackson (2010), à Trinidad et Tobago, observe que les filles sont celles qui bénéficient le plus d'un accès à une meilleure école, Hoekstra, Mouganie et Wang (2018), en Chine, font le constat inverse.

la classe de bons élèves ne bénéficie pas forcément à ces mêmes bons élèves ; il est même désavantageux pour eux de se retrouver dans une très bonne classe.

D'autres travaux ont examiné l'effet d'une scolarisation dans des lycées très sélectifs, dont le recrutement se fait à partir des résultats scolaires antérieurs. Dans le cas des grandes villes américaines, Abdulkadiroğlu, Angrist et Pathak (2014) n'observent pas d'effet positif sur les élèves acceptés « à la marge » dans ce type de lycées « élitistes », pourtant très demandés, plutôt que dans d'autres lycées. Ce résultat est également obtenu par Lucas et Mbiti (2014) dans le cas du Kenya. Les études qui trouvent des effets positifs concluent que ces effets semblent plutôt liés à la qualité des enseignants que ce type d'établissements est capable d'attirer qu'à l'effet des pairs (Hoekstra, Mouganie et Wang, 2018, pour le cas des lycées chinois, ou Jackson, 2010, pour ceux de Trinidad et Tobago).

Certains travaux se sont appuyés sur une modification des seuils d'admission. L'étude de Guyon, Maurin et McNally (2012) est consacrée à la réforme mise en œuvre en 1989 en Irlande du Nord. Cette réforme a favorisé l'accès aux *grammar schools*, établissements d'enseignement secondaire sélectifs mais non élitistes, pour des élèves d'un niveau scolaire plus faible. Les auteurs ont trouvé que les élèves ayant bénéficié de cette réforme ont vu leurs résultats progresser. De plus, les élèves ayant un niveau suffisant pour entrer en *grammar school* avant la réforme ne semblent pas avoir été affectés par l'arrivée d'élèves « un peu moins bons ». Ces constats doivent toutefois être interprétés dans le cadre particulier d'une réforme concernant uniquement des élèves dont les résultats sont situés « à la marge » de l'admission.

D'autres études ont été consacrées à des politiques d'inclusion d'élèves de milieux sociaux défavorisés dans des écoles favorisées. Par exemple, la recherche menée par Angrist et Lang (2004) est consacrée aux effets du programme *Metropolitan Council for Educational Opportunity* (Metco) conçu pour des élèves de l'enseignement élémentaire et des collégiens. Ce programme consistait à transporter (*busing*) des élèves issus des minorités (essentiellement afro-américaines) habitant le district central de Boston en vue de les scolariser dans des districts de banlieue à forte majorité blanche. Ce programme reposait sur la base du volontariat. Il ciblait par conséquent des élèves auto-sélectionnés. Au total, le programme n'a que peu modifié la composition des classes d'accueil. Dans ce cas très particulier, Angrist et Lang (2004) observent que l'arrivée d'élèves de milieux sociaux plutôt défavorisés dans des écoles plus favorisées n'a pas été préjudiciable aux élèves de ces écoles<sup>20</sup>.

---

<sup>20</sup> De façon très schématique, on peut considérer que la classe « type » de ce dispositif a une vingtaine d'élèves, dont deux élèves noirs, et l'effet « Metco » y porte plus ou moins sur l'ajout d'un élève noir. Les résultats pourraient donc s'interpréter en lien avec le modèle dit « de la sous-culture » (Hoxby, Weingarth, 2005), dans lequel des élèves d'une minorité ne représentent pas une perturbation pour les autres à partir du moment où leur proportion n'a pas atteint un seuil critique.

Signalons enfin que de nombreux travaux ont trait à des sujets connexes, par exemple aux effets de la proportion d'élèves d'origine étrangère sur les résultats scolaires des collégiens. Nous renvoyons le lecteur à l'article de Fougère *et alii* (2017) pour une présentation de cette littérature. Ce sujet a fait l'objet de deux recherches récentes utilisant des données françaises (Cebolla-Boado, 2007, et Fougère *et alii*, 2017). Dans ces deux articles, la corrélation négative brute que l'on observe entre la concentration d'enfants étrangers dans les classes et les scores des élèves de ces classes, ou leurs notes, en fin de collège est annulée dès lors que des méthodologies visant à limiter les biais de sélection (ou d'endogénéité) sont utilisées. Ainsi, Fougère *et alii* (2017) trouvent des effets modérés négatifs en début de collège. En français, ils sont plus élevés lorsque la proportion de camarades étrangers dans la classe dépasse un certain seuil ; ils sont également plus élevés pour les élèves étrangers, ce qui permet de penser que ces élèves sont les premiers perdants à la ségrégation dont ils font l'objet.

Dans la plupart de ces recherches, ce sont donc les élèves aux acquis les plus fragiles, ou dont l'environnement familial est le moins favorable à la réussite scolaire, qui sont les plus sensibles à la composition sociale et scolaire de leur classe. Ils bénéficient de l'entourage de camarades de niveau scolaire supérieur même si dans certains cas, une hétérogénéité trop forte leur est préjudiciable. Ce constat se retrouve notamment dans plusieurs recherches sur données françaises.

### **3.3 Toutefois, en matière d'orientation, fréquenter des camarades d'un meilleur niveau scolaire est parfois un désavantage**

Une partie des recherches relatives aux effets de pairs va au-delà des stricts acquis cognitifs, et explore d'autres dimensions, comme l'orientation, le décrochage ou l'influence des camarades de lycée sur les choix d'orientation dans l'enseignement supérieur. L'effet d'entraînement que les meilleurs élèves, ou ceux d'origine sociale plus favorisée, pourraient avoir sur les autres, n'a alors rien d'évident.

Ainsi, dans l'étude de Davezies (2005) portant sur l'enseignement élémentaire, la probabilité de réaliser un parcours scolaire sans redoublement ni orientation vers l'enseignement spécial est supérieure dans les écoles socialement les plus défavorisées, alors que les résultats scolaires y sont plus faibles. Ce constat semble s'expliquer par le fait que les enseignants adaptent leurs exigences au niveau de la classe, y compris en matière de décision de passage en classe supérieure.

En début de lycée, Ly et Riegert (2016) constatent que plus le niveau scolaire des camarades de classe de seconde GT est élevé (le critère étant ici les notes obtenues aux épreuves écrites du DNB en fin de troisième), plus la probabilité de redoubler en fin d'année est élevée, et plus la probabilité d'être orienté en voie générale est faible. Cette situation pourrait traduire le fait que le niveau scolaire atteint

par ces élèves en fin de seconde est plus faible, c'est-à-dire que leurs pairs ont un effet négatif sur leur réussite scolaire. Toutefois, elle est plus vraisemblablement causée par l'ajustement des exigences des conseils de classe et des aspirations des élèves. Pour un même niveau scolaire, les décisions du conseil de classe en matière de redoublement et d'orientation seraient plus sévères. De plus, les aspirations de l'élève seraient elles-mêmes influencées par sa position relative dans la classe, et donc par le niveau scolaire de ses camarades.

Cette même recherche met en évidence que le fait de conserver, en seconde, des camarades de la classe de troisième, est associé à une probabilité plus faible de redoubler, à une orientation plus fréquente en première générale, à l'accès et à l'obtention plus fréquente du baccalauréat trois ans après. En matière de parcours scolaire au lycée général et technologique, il semble donc que l'important ne soit pas d'être entouré de camarades ayant un bon niveau scolaire, mais plutôt de retrouver des camarades de classe du collège. Cet effet est plus marqué pour des élèves en difficulté scolaire ou de milieu social défavorisé.

De tels mécanismes peuvent, par ailleurs, différer selon le genre de l'élève. Landaud, Ly et Maurin (2018) mettent en évidence qu'en seconde générale, la fréquentation de camarades de niveau scolaire élevé influence l'orientation des filles, qui se dirigent alors moins souvent vers une première scientifique. Ces auteurs avancent l'hypothèse que les filles, mises en concurrence avec des élèves de niveau élevé, reverraient leurs ambitions à la baisse, ce qui ne serait pas le cas des garçons.

Au moment de l'orientation vers l'enseignement supérieur, le lien entre décisions d'orientation et composition de la classe ou de l'établissement n'est *a priori* pas évident : comme précédemment, il fait intervenir les caractéristiques, le niveau scolaire et les aspirations des pairs, mais aussi les politiques d'établissement. Il n'existe, à notre connaissance, pas de recherche quantitative sur données françaises permettant de mesurer des effets de pairs *stricto sensu* dans l'orientation post-baccalauréat. Toutefois, une étude sur les données françaises du panel d'élèves entrés au collège en 1995 conclut que les élèves qui sont dans un lycée de composition sociale favorisée envisagent des études supérieures plus longues (Nakhili, 2005). Cette corrélation n'est pas négligeable : à âge, origine sociale et migratoire, mention et série au baccalauréat donnés, un élève scolarisé dans un lycée socialement très favorisé envisage en moyenne des études d'un semestre plus longues qu'un élève scolarisé dans un lycée très défavorisé<sup>21</sup>. Ces élèves ont également une probabilité plus élevée

---

<sup>21</sup> La population d'étude est constituée des élèves de terminale GT, qui ont obtenu le baccalauréat en 2002, sans redoubler dans l'enseignement secondaire. 21 % de ces élèves fréquentent un lycée de type très favorisé et 14 % un lycée de type très défavorisé. En revanche, il n'est pas possible ici d'exclure de possibles effets de sélection : le choix de la scolarisation dans un lycée socialement favorisé pouvant aussi traduire une volonté déjà présente de poursuivre une scolarité longue.

d'envisager une classe préparatoire<sup>22</sup>, même si cet effet est modeste au regard du poids prépondérant que représente, dans ce choix d'orientation, le niveau scolaire de l'élève. Les mécanismes sous-jacents incluent le rôle de l'équipe pédagogique<sup>23</sup>, notamment du corps enseignant (information et motivation des élèves) et la convergence des aspirations scolaires entre pairs.

Ce résultat fait écho à celui obtenu sur données américaines par Bifulco, Fletcher et Ross (2011), avec une méthodologie permettant d'isoler plus finement des effets de pairs. Ces auteurs trouvent qu'être scolarisé avec des camarades d'un milieu social plus favorisé réduit la probabilité de quitter le lycée sans être diplômé et que la probabilité d'accéder à l'université est accrue sans pour autant que les résultats scolaires soient significativement affectés.

Toutefois, exploitant des données suédoises, Jonsson et Mood (2008) ont montré qu'être entouré d'élèves d'un haut niveau scolaire au lycée a un effet négatif sur l'ambition d'entrer à l'université, pour des élèves de niveau moyen. Dans la recherche de Dobbie et Fryer (2014), fréquenter un « lycée d'élite » à New York, à savoir un lycée dont le recrutement se fait sur le niveau scolaire initial, a peu d'effet sur le choix d'accéder à l'enseignement supérieur ou sur la qualité de ce dernier.

Dans ces travaux, incluant ceux sur données françaises, côtoyer des pairs d'un meilleur niveau scolaire ou social est parfois un désavantage, notamment au moment des décisions de redoublement et d'orientation. Ce serait lié à une exigence plus grande de l'institution scolaire lors de ces décisions, et aux ambitions scolaires revues à la baisse. Néanmoins, ces constats sur l'orientation dépendent des logiques institutionnelles. Au-delà des seules recherches sur les effets de pairs, soulignons que les effets ambigus, ou paradoxaux, de la composition des établissements et des classes sur l'orientation et les dimensions non cognitives des élèves sont également illustrés par les travaux de sociologues. Felouzis (2003) a mis en évidence qu'une forte concentration d'élèves d'origine étrangère dans les collèges d'Aquitaine est, au-delà d'un certain seuil, associée, à la fois, à des notes plus faibles au DNB, et à un accès plus fréquent en seconde générale et technologique<sup>24</sup>. Par ailleurs, Fouquet-Chauprade (2013), dans une recherche portant sur les élèves de six collèges ségrégués franciliens, trouve que le niveau de bien-être, pour des élèves d'origine étrangère, est plus élevé lorsque la

---

<sup>22</sup> Une partie de cet effet passe par le fait que ces lycées ont plus souvent une classe préparatoire qui y est implantée. Par ailleurs, cette corrélation se retrouve lorsque l'on considère l'accès effectif aux classes préparatoires, comme le montre une autre étude réalisée sur les mêmes données (Caille et Lemaire, 2009).

<sup>23</sup> Cette hypothèse est corroborée par plusieurs recherches qualitatives, dont celle menée à l'échelle d'une trentaine de lycées généraux et technologiques franciliens (van Zanten, 2015). La composition sociale et scolaire des élèves accueillis, ainsi que le type d'établissement, apparaissent comme des variables déterminantes dans la façon dont sont présentées aux élèves les différentes filières post-baccalauréat.

<sup>24</sup> Ce résultat fait également écho à l'existence de pratiques de sur-notation dans des collèges accueillant des publics en difficulté scolaire (Murat, 1998).

proportion de camarades de classe d'origine étrangère est elle-même élevée<sup>25</sup>. D'autres travaux des sciences de l'éducation font état de résultats des effets de composition sur le bien-être qui peuvent, là encore, diverger de ceux portant spécifiquement sur les résultats scolaires. Par exemple, Belfi *et alii* (2011) soulignent qu'il est positif, pour les filles, d'être dans une classe non-mixte. En effet, il y a moins de rivalité dans les classes non-mixtes et les filles s'y sentent beaucoup plus à l'aise, notamment dans les matières scientifiques. Toutefois, elles n'en tireraient pas forcément d'intérêt en termes de résultats scolaires (Hattie, 2002). Une recherche a mis en évidence que les lycéens américains de milieu social défavorisé ne tirent pas d'avantage net en termes de résultats scolaires lorsqu'ils fréquentent un établissement favorisé (Crosnoe, 2009). Par contre leur sentiment de bien-être se dégrade.

Enfin, d'autres sphères que la seule réussite scolaire doivent être également prises en compte. En effet, plusieurs études mettent en avant des effets « sociaux » importants. Par exemple, au Danemark, Larsen et Kristensen (2017) constatent que les jeunes entrant au lycée professionnel ont une probabilité plus grande de commettre un délit lorsqu'ils côtoient, dans leur établissement, des camarades ayant déjà commis un délit durant leur scolarité au collège. Pour les jeunes ayant déjà commis au moins deux délits, la probabilité de récidive est fortement réduite lorsque la proportion de camarades au passé délinquant est plus faible. D'autres travaux montrent aussi que les comportements et la sociabilité des élèves sont soumis à des effets de pairs significatifs. Ces travaux ayant trait au bien-être et aux comportements des élèves sont toutefois plus rares que ceux portant sur le niveau scolaire, et nous ne connaissons pas de recherche sur données françaises qui y soit consacrée. De tels travaux devraient donc être développés afin d'aboutir à des constats plus robustes.

## Conclusion

La thématique des effets de pairs exige une grande prudence méthodologique. Cette prudence est avant tout la conséquence des difficultés soulevées par leur mesure. Pour proposer une identification non biaisée de ces effets à l'aide de données qui ne proviennent pas d'une expérimentation aléatoire, il est nécessaire de démontrer que ces données et la méthodologie qui leur est appliquée échappent aux critiques usuelles sur la formation non aléatoire des groupes de pairs.

Ces difficultés expliquent sans doute en partie que l'ampleur des effets diffère très sensiblement d'une étude à l'autre. Cette absence de consensus peut être mise en regard des niveaux de scolarité sur

---

<sup>25</sup> Inversement, la situation la plus défavorable, en termes de bien-être, est celle d'un élève qui n'est pas d'origine étrangère mais qui est entouré d'élèves d'origine étrangère.

lesquels ces recherches portent : l'influence des pairs n'est sans doute pas la même selon que l'élève est scolarisé au primaire, au collège ou encore à l'université.

Ces effets peuvent également dépendre des caractéristiques du système scolaire considéré, en particulier, du nombre d'élèves par classe – il est plus facile de gérer l'hétérogénéité d'une classe de 15 que de 30 élèves – ou encore de la formation des enseignants à la différenciation pédagogique, et aux incitations qu'ils reçoivent pour la mettre en œuvre (Duflo, Dupas et Kremer, 2011).

Dans l'ensemble, et de manière assez conforme à l'intuition, les élèves ayant de bons résultats scolaires, ou issus de milieux socio-économiques favorisés, ont plutôt un effet d'entraînement sur l'ensemble des élèves. Inversement, la concentration d'élèves en difficulté scolaire, ou issus d'un environnement familial moins propice à la réussite scolaire, est un facteur pénalisant les performances scolaires, surtout pour ces groupes d'élèves. Il en résulte que les phénomènes de ségrégation peuvent aggraver les inégalités scolaires.

Ces résultats sont toutefois complexes à interpréter sur un plan normatif. Tout d'abord, les recherches ne permettent pas de déterminer un niveau « optimal » de mixité dans les établissements et les classes. De plus, la plupart d'entre elles laissent penser qu'il y aura toujours des gagnants et des perdants à une politique de déségrégation. Dans les recherches des économistes, les effets de pairs liés à la composition sociale et scolaire apparaissent souvent faibles au regard des enjeux de politique éducative, cette conclusion rejoignant un constat fait plus généralement dans les sciences de l'éducation (Hattie, 2002). Enfin, en matière de parcours des élèves et de bien-être, l'intérêt de côtoyer des élèves d'un niveau social ou scolaire plus élevé n'a rien d'évident, et pourrait même être préjudiciable dans certaines situations.

Les travaux récents traitant de ce sujet à l'aide de données françaises sont encore rares. Compte tenu de l'importance du thème de la mixité dans le débat public, ce manque doit être comblé par de nouvelles recherches. Celles-ci devraient également permettre de mieux comprendre les liens entre les effets de pairs et les autres dimensions du contexte scolaire. Il serait notamment souhaitable de les développer en direction des pratiques et compétences des enseignants, tout en prenant en compte le vécu et les attentes des élèves. Le développement de ces recherches implique aussi la construction et la mise à disposition de bases de données très riches (Thrupp, Lauder et Robinson, 2002).



## Annexe 1. Effets exogènes et endogènes, multiplicateur social

### 1.1. Formalisation des effets exogènes et endogènes

Cette présentation reprend la formalisation proposée par Manski (1993) dans laquelle le résultat d'un élève est supposé dépendre de façon linéaire des résultats de ses camarades et de leurs caractéristiques (modèle dit « linéaire en moyenne »). Le modèle est le suivant :

$$y = \alpha + \beta E[y|c] + E[x|c]\gamma + x\lambda + u \quad (1)$$

Dans cette équation :

- $y$  est la variable de résultat (par exemple, le score ou la note d'un élève) ;
- les variables  $c$  caractérisent le groupe des pairs, qu'on considèrera ici être la classe ;
- les variables  $x$  sont les caractéristiques de l'élève (par exemple, son milieu social...) ;
- et  $u$  correspond aux variables inobservées affectant la réussite, qu'on suppose pour l'instant non corrélées au fait d'appartenir à une classe, donc aux variables  $c$ , ce qui mathématiquement se traduit par  $E[u|c] = 0$ . L'expression  $E[u|c]$  représente l'espérance mathématique du terme d'erreur  $u$  conditionnellement à la valeur prise par la variable  $c$ .

Le terme  $\beta E[y|c]$  correspond à la prise en compte des effets endogènes (i.e., aux effets de la réussite moyenne des camarades sur l'élève considéré).  $E[y|c]$  désigne l'espérance mathématique de la variable de résultat  $y$  parmi les élèves de la classe  $c$ . Par hypothèse, en présence d'effets endogènes positifs, la valeur du paramètre  $\beta$  est comprise entre 0 et 1 ( $0 < \beta < 1$ ).

Le terme  $E[x|c]\gamma$  correspond à la prise en compte des effets exogènes (qui sont les effets induits par les caractéristiques des pairs).

## 1.2. L'effet de multiplicateur social

À partir de l'équation (1), on déduit le résultat attendu pour un élève donné, compte tenu de son appartenance à la classe  $c$  :

$$E[y|c] = \alpha + \beta E[y|c] + E[x|c]\gamma + E[x|c]\lambda \quad (2)$$

ce qui implique

$$E[y|c] = \frac{\alpha}{1-\beta} + E[x|c] \frac{\gamma+\lambda}{1-\beta} \quad (3)$$

formule que l'on peut réinjecter dans l'équation (1) pour obtenir :

$$y = \frac{\alpha}{1-\beta} + E[x|c] \frac{\gamma+\beta\lambda}{1-\beta} + x\lambda + u \quad (4)$$

Cette écriture fait apparaître un effet de « multiplicateur social », c'est-à-dire que les effets des modifications des caractéristiques des pairs sont amplifiés par le jeu des interactions.

Supposons que  $x$  corresponde au niveau de diplôme de ses parents, lequel a une influence positive sur ses performances scolaires, par exemple du fait que des parents diplômés peuvent plus facilement aider leurs enfants pour des devoirs à la maison. Pour simplifier, on peut considérer qu'il n'y a que deux types de familles, les « diplômées » ( $x = 1$ ) et les « non diplômées » ( $x = 0$ ). Le coefficient  $\lambda$  dans le modèle quantifie l'écart de performance entre un élève dont les parents sont diplômés et un élève dont les parents ne le sont pas, en l'absence d'effets de pairs. Supposons maintenant qu'on souhaite comparer ces mêmes deux élèves, en présence d'une ségrégation complète entre les classes : l'élève dont les parents sont diplômés est dans une classe « favorisée », dans laquelle tous les élèves ont des parents diplômés ; l'autre élève est dans une classe où aucun élève n'a de parents diplômés. Dans le cas où des effets de pairs existent, et si ceux-ci sont uniquement exogènes, c'est-à-dire si  $\beta = 0$  et  $\gamma \neq 0$  (dans la suite on suppose  $\gamma > 0$ , cas le plus vraisemblable), l'élève dans la classe favorisée bénéficie aussi du fait que ses camarades ont un capital culturel élevé – et cela se traduit désormais par un écart de niveau  $\lambda + \gamma$  par rapport à son camarade qui se trouve dans la classe non favorisée. L'avantage initial de l'élève ayant des parents diplômés est augmenté de celui d'être entouré d'élèves similaires (effets exogènes).

Enfin, en présence d'effets de pairs exogènes et endogènes, c'est-à-dire lorsque  $\gamma \neq 0$  et  $\beta \neq 0$  (dans la suite on suppose  $0 < \beta < 1$  et  $\gamma > 0$ ), le fait que les élèves dans la classe favorisée ont une meilleure réussite influence aussi la progression de chacun de leurs camarades (par exemple par le biais de phénomènes d'émulation ou au contraire de découragement). L'écart de score individuel entre un élève de parents diplômés dans une classe dont tous les camarades ont aussi des parents

diplômés et un élève non diplômé dont les camarades sont identiques est alors de  $\frac{\gamma+\beta\lambda}{1-\beta} + \lambda = \frac{\gamma+\lambda}{1-\beta}$ . Cet écart est supérieur à  $\lambda + \gamma$  car le terme  $\frac{1}{1-\beta}$  est supérieur à 1. Les effets de l'avantage initial (avoir des parents diplômés) et de la ségrégation entre les classes (être entouré d'élèves similaires), sont donc amplifiés du fait de l'existence d'effets de pairs endogènes, c'est-à-dire du fait que la réussite de chaque élève profite à tous les autres par le biais du « multiplicateur social »<sup>26</sup>.

### 1.3. Le problème de la réflexion et ses conséquences pour l'estimation des effets endogènes

Compte tenu du problème de la réflexion exposé par Manski (1993), si l'on estime les effets de pairs endogènes en régressant le résultat des élèves sur le résultat moyen de leur classe, on obtient un résultat trivial, c'est-à-dire un coefficient de 1. Cette corrélation n'est pas la preuve de l'influence du groupe sur l'élève, mais une tautologie.

Pour pallier cette difficulté, on peut s'appuyer sur l'équation (4) en forme réduite, qui fait apparaître le paramètre d'intérêt  $\beta$ . Toutefois, les paramètres  $\gamma$  et  $\beta$ , qui interviennent de façon combinée, ne peuvent plus être alors identifiés séparément. Les effets endogènes et exogènes ne sont pas identifiables, ce qui est une autre conséquence du problème de réflexion.

Trois solutions à ce problème peuvent être envisagées. Tout d'abord, des méthodologies existent pour permettre une telle identification, par exemple lorsque l'élève observé est retiré du calcul de la moyenne de la classe, et qu'au moins trois classes de tailles différentes sont présentes dans l'échantillon (Lee, 2007 ; Bramoullé *et alii*, 2009 ; Davezies, d'Haultfoeuille et Fougère, 2009). Ensuite, il est possible d'ajouter une contrainte sur les paramètres ; les analystes uniquement intéressés par l'identification des effets endogènes supposent en général que le paramètre  $\gamma$  est égal à 0, soit l'absence d'effets exogènes. Enfin, de nombreux travaux se limitent à l'estimation de l'équation (4), c'est-à-dire qu'ils estiment un effet exogène en forme réduite correspondant au terme  $\frac{\gamma+\beta\lambda}{1-\beta}$ . Cet effet exogène inclut les effets purement exogènes  $\gamma$  sans pouvoir les distinguer spécifiquement.

---

<sup>26</sup> On notera la parenté de ce mécanisme avec la littérature économique sur les « multiplicateurs », dont le plus connu est le multiplicateur keynésien. Le principe général est qu'une politique peut avoir des effets de long terme supérieurs aux effets immédiats, car la dynamique introduite par cette politique est auto-entretenu.

## Annexe 2. Illustration des principales méthodes d'estimation des effets de pairs exogènes

### 2.1 Le biais d'endogénéité

On s'intéresse ici uniquement à l'estimation d'effets exogènes, c'est-à-dire à des effets traduisant l'influence des caractéristiques des pairs. L'équation estimée dans la plupart des études est de la forme<sup>27</sup> :

$$Y_i = b_0 + X_i b_1 + \bar{X}_{-i} b_2 + \varepsilon_i \quad (5)$$

Dans cette équation :

- $Y_i$  est la variable de résultat observée pour l'élève  $i$  ;
- les variables  $X_i$  sont les caractéristiques de cet élève (milieu social...) ;
- les variables  $\bar{X}_{-i}$  sont les caractéristiques « moyennes » du groupe de pairs, qu'on considèrera être la classe. Dans le calcul de l'expression  $\bar{X}_{-i}$ , on retire<sup>28</sup> l'élève  $i$  ;
- $\varepsilon_i$  correspond à un terme aléatoire.

Contrairement à l'annexe 1, on considère désormais que le terme aléatoire peut être corrélé aux caractéristiques moyennes des camarades, représentées par le vecteur  $\bar{X}_{-i}$ , cette corrélation pouvant biaiser l'estimation du coefficient  $b_2$ .

Malgré ce risque de biais, l'estimation par moindres carrés ordinaires de l'équation (5) peut être envisagée dans deux cas. Le premier est celui où l'on dispose d'une liste de variables  $X_i$  suffisamment riche de façon à prendre en compte de nombreuses dimensions liées à la composition des classes et aux niveaux scolaires des élèves. Le second cas est celui où les données ont été collectées, ou sélectionnées, de manière à ce que l'on puisse considérer que l'affectation des élèves aux établissements ou aux classes est aléatoire. C'est le cas des expérimentations aléatoires, ainsi que des expériences naturelles ou quasi-expériences.

<sup>27</sup> Cette équation constitue la contrepartie empirique de l'équation (4) de l'annexe 1 ; les espérances mathématiques apparaissant dans l'équation (4) sont ici remplacées par les moyennes calculées avec les données observées.

<sup>28</sup> Ce choix est fréquemment fait dans les travaux empiriques, notamment parce que l'on suppose un mécanisme particulier d'interactions entre pairs (l'élève est influencé par ses camarades, mais ne s'influence pas lui-même). On pourrait également laisser l'élève  $i$  dans le calcul de la moyenne, par exemple si l'on suppose que la composition de la classe dans son ensemble crée un climat plus ou moins favorable à l'enseignement.

## 2.2 Application de la méthode des variables instrumentales

Le cas d'un modèle avec variables instrumentales s'illustre par un système à deux équations :

$$Y_i = b_0 + X_i b_1 + \bar{X}_{-i} b_2 + \varepsilon_i \quad (6)$$

$$\bar{X}_{-i} = c_0 + X_i c_1 + Z_i c_2 + \mu_i \quad (7)$$

La seconde équation fait apparaître les variables instrumentales  $Z_i$ . Celles-ci sont supposées être corrélées à la composition de la classe, représentée par le vecteur  $\bar{X}_{-i}$ , mais non corrélées au terme d'erreur  $\varepsilon_i$  (voir la section « Les méthodologies fondées sur des sources de variation exogène »).

## 2.3 Exemple d'application de modèles avec effets fixes

Estimer des effets fixes suppose que l'analyste ait accès à des données appropriées : pour un même groupe, on peut disposer soit d'observations longitudinales, soit d'observations multiples. Par exemple, on peut observer plusieurs classes d'un même établissement, ou bien le même établissement au cours du temps.

Dans l'équation d'intérêt, on intègre alors un effet fixe spécifique à l'établissement  $e_j$  :

$$Y_i = b_0 + X_i b_1 + \bar{X}_{-i} b_2 + e_j + \varepsilon_i \quad (8)$$

Compte tenu des effets de sélection, les élèves ne sont pas répartis aléatoirement entre les établissements. Cela implique que les caractéristiques de l'élève  $i$  représentées par le vecteur  $X_i$  et les caractéristiques moyennes de ses camarades, représentées par le vecteur  $\bar{X}_{-i}$ , sont corrélées à l'effet établissement  $e_j$ , cette corrélation pouvant biaiser l'estimation du coefficient d'intérêt  $b_2$ .

L'objectif de la méthode est donc d'estimer un coefficient  $b_2$  en tenant compte des effets fixes d'établissement (Givord et Guillerm, 2016). Dans cet exemple, on agrège les données des élèves qui sont scolarisés dans l'établissement  $j$  : la moyenne des résultats obtenus par les élèves de l'établissement  $j$  est notée  $\bar{Y}_j$ .

$$\bar{Y}_j = b_0 + \bar{X}_j b_1 + \bar{X}_j b_2 + e_j + \bar{\varepsilon}_j \quad (9)$$

On fait ensuite la différence terme à terme des équations (8) et (9) :

$$Y_i - \bar{Y}_j = (X_i - \bar{X}_j) b_1 + (\bar{X}_{-i} - \bar{X}_j) b_2 + \varepsilon_i - \bar{\varepsilon}_j \quad (10)$$

Dans cette équation, les régresseurs sont par hypothèse non corrélés avec l'erreur  $\varepsilon_i - \bar{\varepsilon}_j$ . Cette différenciation a permis d'éliminer l'effet fixe associé à l'établissement  $e_j$  et le biais.

Cette logique s'applique de façon comparable à d'autres situations. Par exemple, l'identification d'un effet fixe spécifique à l'enseignant suppose que l'on observe un même enseignant devant des classes de compositions différentes (à différents niveaux de la scolarité, ou à différentes périodes), sous l'hypothèse que l'effet fixe associé à cet enseignant est le même dans toutes ces situations. L'identification d'un effet fixe spécifique à l'élève suppose de façon analogue que l'on observe un même élève à différentes périodes, ou dans différentes matières évaluées, là encore sous l'hypothèse que cet effet fixe est le même à toutes les périodes, ou pour toutes les matières considérées. Les données permettant de prendre en compte ces différentes formes d'effets fixes doivent donc être suffisamment riches. Par ailleurs, ces méthodes nécessitent des hypothèses souvent assez fortes.

## Bibliographie

- ABDULKADIROĞLU A., ANGRIST J., PATHAK P., 2014, “The Elite Illusion: Achievement Effects at Boston and New York Exam Schools”, *Econometrica*, vol. 82, n°1, p. 137-196.
- ANGRIST J., LANG K., 2004, “Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program”, *The American Economic Review*, vol. 94, n°5, p. 1613-1634.
- AMMERMUELLER A., PISCHKE J. S., 2009, “Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study”, *Journal of Labor Economics*, vol. 27, n°3, p. 315-348.
- BELFI B., GOOS M., DE FRAINE B., VAN DAMME J., 2012, “The effect of class composition by gender and ability on secondary school students' school well-being and academic self-concept: A literature review”, *Educational Research Review*, vol. 7, n°1, p. 62-74.
- BIFULCO R., FLETCHER J., ROSS S., 2011, “The Effect of Classmate Characteristics on Post-Secondary Outcomes: Evidence from the Add Health”, *American Economic Journal: Economic Policy*, vol. 3, n°1, p. 25-53.
- BOUTCHENIK B., MAILLARD S., 2018, « Peer Effects with Peer and Student Heterogeneity: An Assessment for French Baccalauréat », Actes des XIIIes Journées de Méthodologie Statistique de l'Insee, 12-14 juin 2018.
- BRAMOULLÉ Y., DJEBBARI H., FORTIN B. (2009), “Identification of peer effects through social networks”, *Journal of Econometrics*, vol. 150, issue 1, p. 41-55.
- BRESSOUX P., 2008, « Effet-établissement », in VAN ZANTEN A. (dir), *Dictionnaire de l'éducation*, Paris : Presses Universitaires de France, p. 212-216.
- BRODATY T., 2010, « Les effets de pairs dans l'éducation : une revue de littérature », *Revue d'économie politique*, vol. 120, n°5, p. 739-757.
- BURKE M., SASS T., 2013, “Classroom Peer Effects and Student Achievement”, *Journal of Labor Economics*, vol. 31, n° 1, p. 51-82.
- CAILLE J.-P., LEMAIRE S., 2009, « Les bacheliers « de première génération » : des trajectoires scolaires et des parcours dans l'enseignement supérieur « bridés » par de moindres ambitions ? », *France Portrait Social*, p. 171-193.
- CEBOLLA-BOADO H., 2007, “Immigrant Concentration in Schools: Peer Pressures in Place?”, *European Sociological Review*, vol. 23, n°3, p. 341-356.
- CROSNOE R., 2009, “Low-Income Students and the Socioeconomic Composition of Public High Schools”, *American Sociological Review*, vol. 74, n°5, p. 709-730.
- DAVEZIES L., 2005, « Influence des caractéristiques des pairs sur la scolarité », *Educations & formations*, n°72, MENESR-DEPP, p. 171-199.
- DAVEZIES L., D'HAULTFOEUILLE X., FOUGÈRE D., 2009, “Identification of Peer Effects Using Group Size Variation”, *The Econometrics Journal*, vol. 12, n°3, p. 397-413.
- DOBBIE W., FRYER R. G., 2014, “The Impact of Attending a School with High-Achieving Peers: Evidence from New York City Exam Schools”, *American Economic Journal: Applied Economics*, vol. 6, n°3, p. 58-75.

- DUFLO E., DUPAS P., KREMER M., 2011, “Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya”, *American Economic Review*, vol. 101, n°5, p. 1739-1774.
- DUPRIEZ V., MONSEUR C., VAN CAMPENHOUDT M., 2012, « Le poids de l’origine socioculturelle des élèves et de leur environnement scolaire sur leurs aspirations d’études supérieures : les bases d’une comparaison internationale », *Revue de l’orientation scolaire et professionnelle*, vol. 41, n°1, p. 29-56.
- DURU-BELLAT M., MINGAT A., 1988, « Le déroulement de la scolarité au collège : le contexte « fait des différences » », *Revue française de sociologie*, vol. 29, n°4, p. 649-666.
- EPPLE D., ROMANO R., 2011, “Peer Effects in Education: A Survey of the Theory and Evidence”, in *Handbook of Social Economics*, vol. 1, p. 1053-1163.
- FACK G., GRENET J., 2016, « Mixité sociale et scolaire dans les lycées parisiens », *Éducation & formations*, n° 91, MENESR-DEPP, p. 77-100.
- FELOUZIS G., 2003, « La ségrégation ethnique au collège et ses conséquences », *Revue française de sociologie*, vol. 44, n°3, p. 413-447.
- FOUGERE D., 2010, « Les méthodes d’expérimentation en question », *Éducation & formations*, n°81, MEN-DEPP, p. 41-47.
- FOUGERE D., KIEFER N., MONSO O., PIRUS C., 2017, « La concentration des enfants étrangers dans les classes de collèges », *Éducation & formations*, n°95, MEN-DEPP, p. 139-172.
- FOUGERE D., GIVORD P., MONSO O., PIRUS C., 2018, « Les effets de pairs en éducation : comment sont-ils mesurés et pour quelles conclusions ? », Actes des XIIIes Journées de Méthodologie Statistique de l’Insee, 12-14 juin 2018.
- FOUQUET-CHAUPRADE B., 2013, « L’ethnicité au collège : bien-être et effet de contexte », *Sociologie*, vol. 4, n°4, p. 431-449.
- GIBBONS S., TELHAJ S., 2016, “Peer Effects: Evidence from Secondary School Transition in England”, *Oxford Bulletin of Economics and Statistics*, vol. 78, n°4, p. 548-575.
- GIVORD P., GUILLERM M., 2016, « Les modèles multiniveaux », document de travail de l’INSEE, n°M2016/05.
- GOUX D., MAURIN É., 2007, “Close neighbours matter: neighbourhood effects on early performance at school”, *The Economic Journal*, vol. 117, n° 523, p. 1193-1215.
- GRAHAM B. S., 2008, “Identifying Social Interactions Through Conditional Variance Restrictions”, *Econometrica*, vol. 76, n°3, p. 643-660.
- GUYON N., MAURIN É., McNALLY S., 2012, “The Effect of Tracking Students by Ability into Different Schools: A Natural Experiment”, *The Journal of Human Resources*, vol. 47, n°3, p. 684-721.
- HANUSHEK E. A., KAIN J. F., RIVKIN S. G., 2009, “New Evidence about *Brown v. Board Education*: The Complex Effects of School Racial Composition on Achievement”, *Journal of Labor Economics*, vol. 27, p. 349-383.
- HATTIE J. A. C., 2002, “Classroom composition and peer effects”, *International Journal of Educational Research*, n° 37, p. 449-481.
- HOEKSTRA M., MOUGANIE P., WANG Y., 2018, “Peer Quality and the Academic Benefits to Attending Better Schools”, *Journal of Labor Economics*, vol. 36, n°4, p. 841-884.



- HOXBY C., WEINGARTH G., 2005, "Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects", Working paper, Harvard University.
- IMBERMAN S., KUGLER A., SACERDOTE B., 2012, "Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees", *American Economic Review*, vol. 102, n°5, p. 2048-2082.
- JACKSON K., 2010, "Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago", *The Economic Journal*, vol. 120, issue 549, p. 1399-1429.
- JONSSON J., MOOD C., 2008, "Choice by Contrast in Swedish Schools: How Peers' Achievement Affects Educational Choice", *Social Forces*, vol. 87, n° 2, p. 741-765.
- LANDAUD F., LY S.-T., MAURIN É., 2018, "Competitive Schools and the Gender Gap in the Choice of Field of Study", *Journal of Human Resources*, 10.3368/jhr.55.2.0617.8864R.
- LARSEN B. Ø., KRISTENSEN N., 2017, "Building Human or Criminal Capital? School Peer Effects on Future Offending", IZA Discussion Paper, n°11124, Bonn.
- LAVY V., SILVA O., WEINHARDT F., 2012, "The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools", *Journal of Labor Economics*, vol. 30, n°2, p. 367-414.
- LEE L., 2007, "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects", *Journal of Econometrics*, vol. 140, n°2, p. 333-374.
- LUCAS A., MBITI I., 2014, "Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya", *American Economic Journal: Applied Economics*, vol. 6, n°3, p. 234-263.
- LY S.-T., RIEGERT A., 2015, *Mixité sociale et scolaire, ségrégation inter et intra-établissement dans les collèges et lycées français*, Rapport pour le Conseil national d'évaluation du système scolaire (CNESEO).
- LY S.-T., RIEGERT A., 2016, « Retrouver ses camarades de classe en seconde. Un atout pour la scolarité au lycée », *Éducation & formations*, n°91, MENESR-DEPP, p.101-114.
- MANSKI C., 1993, "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, vol. 60, n°3, p. 531-542.
- MURAT F., 1998, « Les différentes façons d'évaluer le niveau des élèves en fin de collège », *Éducation & formations*, n°53, MENERT-DEP, p. 35-49.
- NAKHILI N., 2005, « Impact du contexte scolaire dans l'élaboration des choix d'études supérieures des élèves de terminale », *Éducation & formations*, n°72, p. 155-167.
- PIKETTY T., 2004, « L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises. Une estimation à partir du panel primaire 1997 », Document de travail Paris-Jourdan.
- 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 », *Les dossiers évaluations et statistiques*, MEN-DEP, n°173.
- POP-ELECHES C., URQUIOLA M., 2013, "Going to a Better School: Effects and Behavioral Responses", *American Economic Review*, vol. 103, n°4, p. 1289-1324.
- RANGVID B. S., 2007, "School Composition Effects in Denmark: Quantile Regression Evidence from PISA 2000", *Empirical Economics*, n° 33, p. 359-388.

SACERDOTE B., 2011, "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?", *Handbook of the Economics of Education*, vol. 3, p. 249-277.

SCHNEEWEIS N., WINTER-EBMER R., 2007, "Peer Effects in Austrian schools", *Empirical Economics*, vol. 32, p. 387-409.

THRUPP M., LAUDER H., ROBINSON T., 2002, "School Composition and Peer Effects", *International Journal of Educational Research*, vol. 37, n°5, p. 483-504.

VALLET L.-A., 2005, « La mesure des effets de quartier/voisinage : un objet important et difficile à la croisée des sciences sociales », commentaire sur l'article de D. GOUX et É. MAURIN, « Composition sociale du voisinage et échec scolaire », *Revue économique*, vol. 56, n°2, p. 349-361.

VAN ZANTEN A., 2015, « Les inégalités d'accès à l'enseignement supérieur : quel rôle joue le lycée d'origine des futurs étudiants ? », *Regards Croisés sur l'Économie*, n° 16, p. 80-92.



**U-PC**  
Université Sorbonne  
Paris Cité

**Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex).  
Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR).  
Il est financé dans le cadre des investissements d'avenir.**

(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)

**[www.sciencespo.fr/liepp](http://www.sciencespo.fr/liepp)**

## **A propos de la publication**

### **Procédure de soumission :**

Rédigé par un ou plusieurs chercheurs sur un projet en cours, le *Working paper* vise à susciter la discussion scientifique et à faire progresser la connaissance sur le sujet étudié. Il est destiné à être publié dans des revues à comité de lecture (peer review) et à ce titre répond aux exigences académiques. Les textes proposés peuvent être en français ou en anglais. En début de texte doivent figurer : les auteurs et leur affiliation institutionnelle, un résumé et des mots clefs.

Le manuscrit sera adressé à : [liepp@sciencespo.fr](mailto:liepp@sciencespo.fr)

Les opinions exprimées dans les articles ou reproduites dans les analyses n'engagent que leurs auteurs.

### **Directeur de publication :**

Bruno Palier

### **Comité de rédaction :**

Andreana Khristova, Samira Jebli

Sciences Po - LIEPP  
27 rue Saint Guillaume  
75007 Paris - France  
+33(0)1.45.49.83.61  
[liepp@sciencespo.fr](mailto:liepp@sciencespo.fr)

