

# Effets de Pairs en Éducation : Applications sur Données Réunionnaises

Présentation au séminaire du CEMOI

Daniel Rakotomalala

Centre d'Économie et de Management de l'Océan Indien

17 Juin 2021

Introduction

Background théorique

Revue de littérature

Éléments institutionnels et données

Modélisation

Résultats

Conclusion



# Introduction

# Contexte

- ▶ Un élève de 3<sup>ème</sup> du système français passe 26 heures par semaines avec ses camarades de classe (ses **pairs**).<sup>1</sup>
- ▶ À part les interactions qui passent par l'enseignant, les élèves d'une même classe **interagissent directement** : entraide ou compétition, entre autres (**Davezies 2004**).
- ▶ Ces interactions directes peuvent affecter les performances scolaires : ce sont les **effets de pairs**.

---

<sup>1</sup> <<https://www.education.gouv.fr/les-horaires-par-cycle-au-college-9884>>.

# Le papier

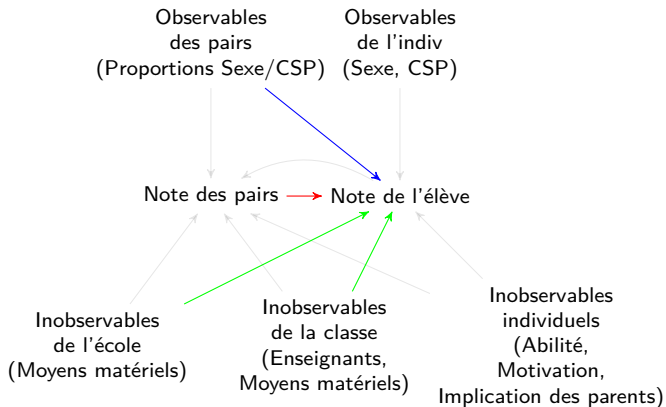
- ▶ Objectif : Mesurer l'influence du niveau initial des pairs sur les résultats en fin de collège à La Réunion.
- ▶ Intérêts :
  - ▶ Complète la connaissance relativement limitée sur cette question, pour le cas de la France (Monso et al. 2019).
  - ▶ Territoire étudié : La Réunion, relativement défavorisée sur les côtés social<sup>2</sup> et scolaire par rapport à la métropole (Arneton, Bocéran, and Flieller 2013).
  - ▶ Utilisation d'un modèle tenant compte de la présence non négligeable et non aléatoire des données manquantes sur le niveau des pairs : Sojourner (2013).
  - ▶ Contribue au débat sur les politiques de mixité scolaire : faut-il former des classes de niveau hétérogènes ou homogènes ?

---

<sup>2</sup><https://www.insee.fr/fr/statistiques/4482473#en-six-questions>

## Background théorique

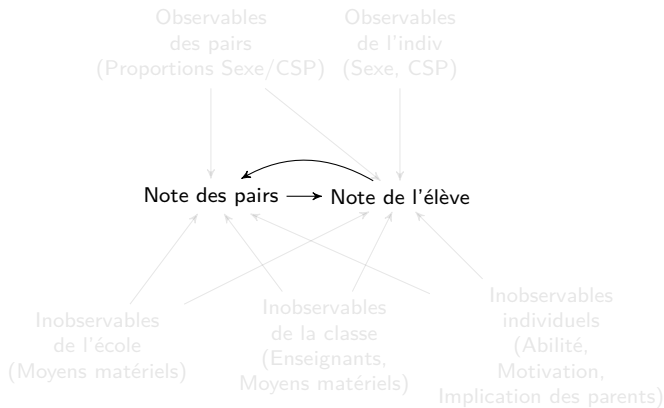
## Définitions (Manski 1993) : illustration via les déterminants de la note d'un élève



- ▶ Deux types d'effets de pairs : *effets de pairs "endogènes"* (→) et *effets de pairs "exogènes"* (→).
- ▶ Un troisième type d'effet à distinguer des effets de pairs : les *effets corrélés* (→).

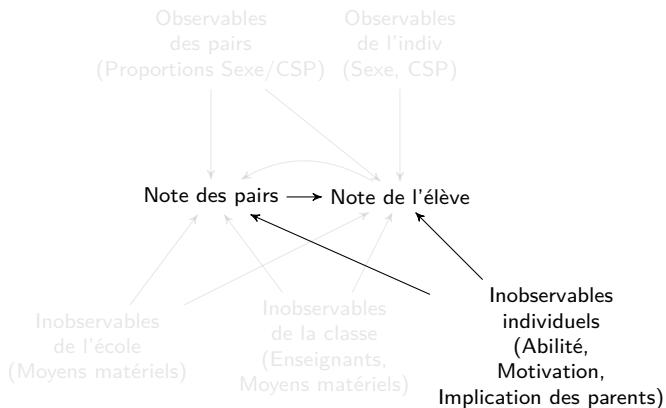


# Problèmes d'identification : le problème de la réflexion



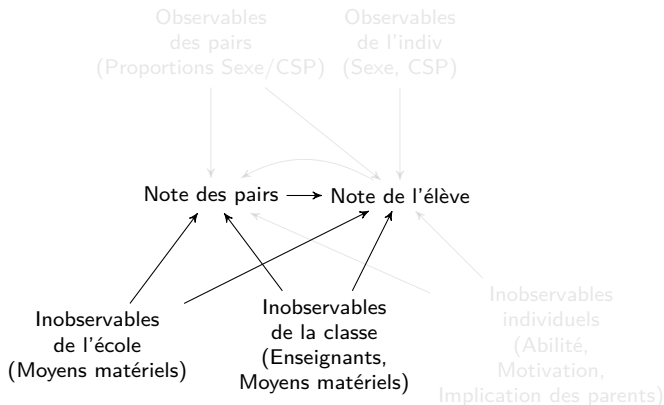
- ▶ L'individu est influencé par les pairs mais l'inverse est aussi vrai (exemple : élève modèle ou perturbateur).
- ▶  $\Rightarrow$  Si on observe une corrélation entre la note des pairs et la note de l'élève, il est impossible de connaître qui a influencé qui ([Angrist 2014](#)).

# Problèmes d'identification : la formation endogène des pairs (sélection)



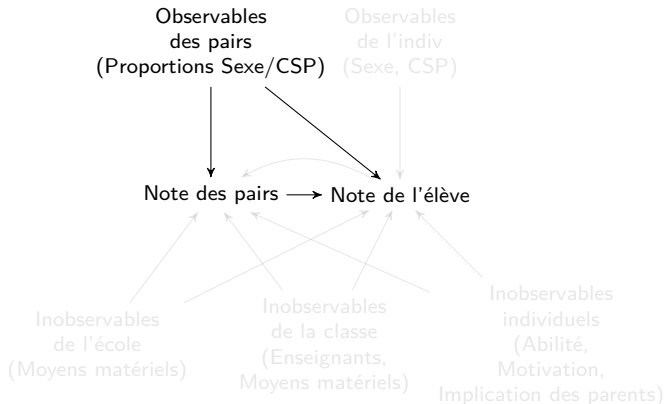
- ▶ Exemple : ceux déjà bons *a priori* ont tendance à se regrouper entre eux.
- ▶ ⇒ Corrélation positive entre note des pairs et note de l'élève : difficile de séparer les effets de pairs de la simple sélection.

# Problèmes d'identification : les effets corrélés



- ▶ Exemple : l'effet de la qualité de l'enseignant (effet corrélé de classe) peut se confondre avec l'effet exclusif des pairs.
- ▶ Exemple : l'effet des moyens matériels de l'établissement (effet corrélé d'établissement) peut se confondre avec l'effet exclusif des pairs.

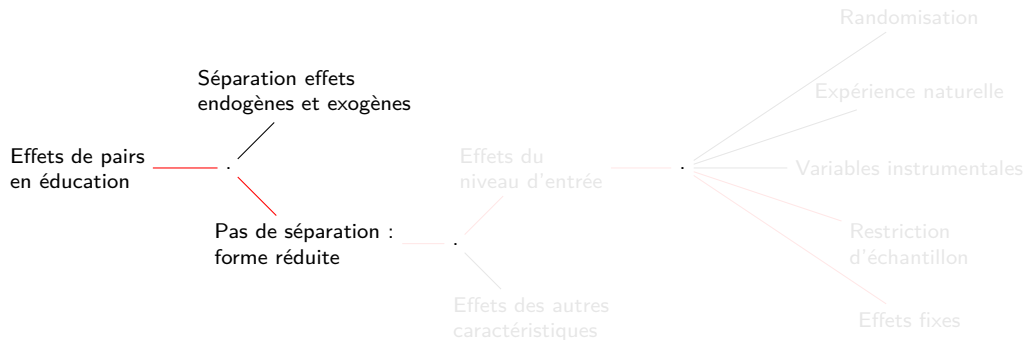
## Problèmes d'identification : forte corrélation entre les effets de pairs endogènes et exogènes (Sacerdote 2011)



- ▶ Exemple : des pairs ayant de bonnes notes ont tendance à être de catégorie sociale aisée.
- ▶ ⇒ Difficile de séparer l'effet de la note des pairs de l'effet de la catégorie sociale.

## Revue de littérature

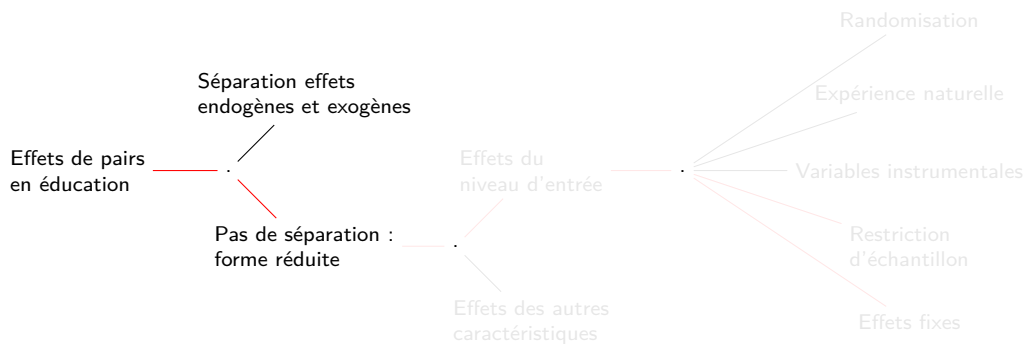
# Deux groupes de papiers, deux objectifs (Yeung and Nguyen-Hoang 2016)



— : Là où se situe le papier présent.

- ▶ Séparation des effets endogènes et exogènes (Boucher et al. 2014, par exemple).
  - ▶ Méthodologie en plein essor (Lee 2007, entre autres).
  - ▶ Intérêts en politiques publiques : politiques ciblées sur certains types d'individus pour que les effets endogènes se manifestent (Monso et al. 2019).

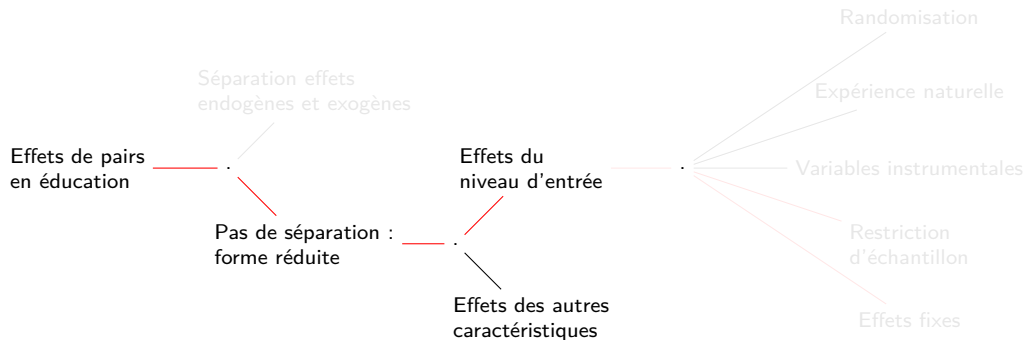
# Deux groupes de papiers, deux objectifs (Yeung and Nguyen-Hoang 2016)



— : Là où se situe le papier présent.

- ▶ Forme réduite ([Boutchenik and Maillard 2018](#), pour la France, par exemple).
  - ▶ Ne mettent pas la note contemporaine des pairs dans l'équation.
  - ▶ Intérêts en politiques publiques : politiques de mixité scolaire ou sociale.

# Deux types d'effets de pairs d'intérêt

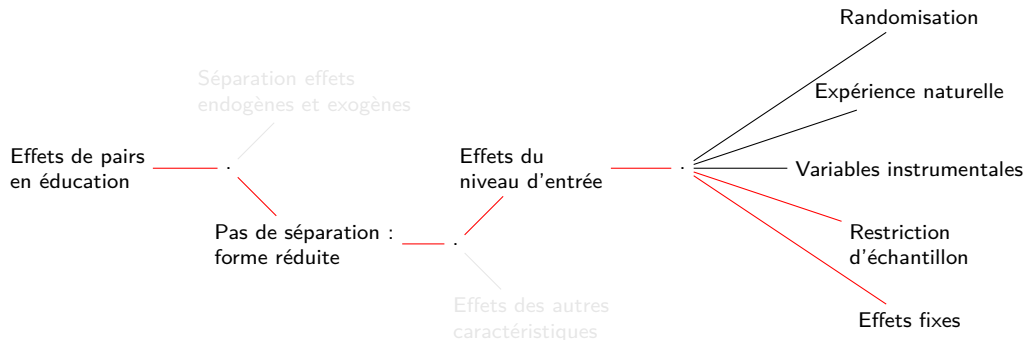


— : Là où se situe le papier présent.

- ▶ Effets du niveau d'entrée ([Lavy, Silva, and Weinhardt 2012](#) ; [Boutchenik and Maillard 2018](#) ; [Sojourner 2013](#), par exemple.) : intérêt = politiques de mixité scolaire.
- ▶ Effets des autres caractéristiques ([Hoxby and Weingarth 2005](#), en partie) : intérêt = politiques de mixité sociale.



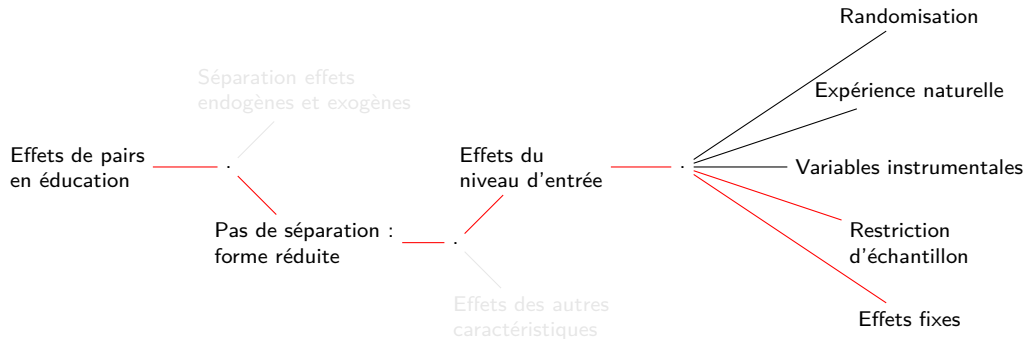
# Plusieurs stratégies d'identification



— : Là où se situe le papier présent.

- ▶ Randomisation ([Duflo, Dupas, and Kremer 2011](#), par exemple) : on contrôle explicitement la formation des pairs.
- ▶ Expérience naturelle ([Hoxby and Weingarth 2005](#), par exemple) : on s'appuie sur un événement exogène faisant varier la composition des pairs.
- ▶ Variables instrumentales ([Davezies 2004](#), par exemple) : on trouve une variable exogène corrélée avec le niveau des pairs sans relation directe avec la note de l'élève.

# Plusieurs stratégies d'identification



— : Là où se situe le papier présent.

- ▶ Restriction d'échantillon ([Lavy, Silva, and Weinhardt 2012](#), par exemple) : on se limite à une partie de l'échantillon où on peut considérer la formation des pairs comme aléatoire.
- ▶ Effets fixes ([Sojourner 2013](#), par exemple) : on compare les pairs au sein d'une même structure (ex : établissement) dans laquelle on peut considérer la formation des pairs comme aléatoire.

# Tendances dans les résultats (forme réduite, effet du niveau initial des pairs)

Papiers	Pays	Niveau	Stratégie d'identification	Résultats : effets LEM	Autres résultats
Duflo et al. (2011)	Kenya	Primaire	Randomisation	1 UET → 0.5 UET	Effet très fort pour les plus forts, moins fort pour les plus faibles et inexistant pour les moyens.
Hoxby et Weingarth (2005)	États-Unis	Primaire et collège	Expérience naturelle	1 point → 0.25 point	Les plus faibles bénéficient des pairs légèrement plus forts.
Davezies (2004)	France	Primaire	Variable instrumentale	1 point → 0.5 point	Effet moyen plus fort pour les plus faibles.
Lavy et al. (2012)	Angleterre	Collège	Effets fixes (établissement-niveau) et restriction d'échantillon	Non significatif	Effet négatif des pairs les plus faibles.
Boutchenik et Maillard (2018)	France	Lycée	Effets fixes (établissement-filière) et restriction d'échantillon	1 RP → 0.08 RP	Effet moyen plus fort pour les plus faibles.
Sojourner (2013)	États-Unis	Primaire	Effets fixes (établissement)	1 RP → 0.4 RP	Effet moyen plus forts pour les plus faibles.

Sources : Papiers cités, auteur.

Notes : LEM : linéaire en moyennes. UET : unité d'écart-type. RP : rang de percentiles.

- Tendance générale : existence d'un effet positif.

# Tendances dans les résultats (forme réduite, effet du niveau initial des pairs)

Papiers	Pays	Niveau	Stratégie d'identification	Résultats : effets LEM	Autres résultats
Duflo et al. (2011)	Kenya	Primaire	Randomisation	1 UET → 0.5 UET	Effet très fort pour les plus forts, moins fort pour les plus faibles et inexistant pour les moyens.
Hoxby et Weingarth (2005)	États-Unis	Primaire et collège	Expérience naturelle	1 point → 0.25 point	Les plus faibles bénéficient des pairs légèrement plus forts.
Davezies (2004)	France	Primaire	Variable instrumentale	1 point → 0.5 point	Effet moyen plus fort pour les plus faibles.
Lavy et al. (2012)	Angleterre	Collège	Effets fixes (établissement-niveau) et restriction d'échantillon	Non significatif	Effet négatif des pairs les plus faibles.
Boutchenik et Maillard (2018)	France	Lycée	Effets fixes (établissement-filière) et restriction d'échantillon	1 RP → 0.08 RP	Effet moyen plus fort pour les plus faibles.
Sojourner (2013)	États-Unis	Primaire	Effets fixes (établissement)	1 RP → 0.4 RP	Effet moyen plus forts pour les plus faibles.

Sources : Papiers cités, auteur.

Notes : LEM : linéaire en moyennes. UET : unité d'écart-type. RP : rang de percentiles.

- Cet effet semble dépendre du niveau d'éducation : plus fort en bas niveau.

# Tendances dans les résultats (forme réduite, effet du niveau initial des pairs)

Papiers	Pays	Niveau	Stratégie d'identification	Résultats : effets LEM	Autres résultats
Duflo et al. (2011)	Kenya	Primaire	Randomisation	1 UET $\rightarrow$ 0.5 UET	<b>Effet très fort pour les plus forts, moins fort pour les plus faibles et inexistant pour les moyens.</b>
Hoxby et Weingarth (2005)	États-Unis	Primaire et collège	Expérience naturelle	1 point $\rightarrow$ 0.25 point	Les plus faibles bénéficient des pairs légèrement plus forts.
Davezies (2004)	France	Primaire	Variable instrumentale	1 point $\rightarrow$ 0.5 point	Effet moyen plus fort pour les plus faibles.
Lavy et al. (2012)	Angleterre	Collège	Effets fixes (établissement-niveau) et restriction d'échantillon	Non significatif	Effet négatif des pairs les plus faibles.
Boutchenik et Maillard (2018)	France	Lycée	Effets fixes (établissement-filière) et restriction d'échantillon	1 RP $\rightarrow$ 0.08 RP	Effet moyen plus fort pour les plus faibles.
Sojourner (2013)	États-Unis	Primaire	Effets fixes (établissement)	1 RP $\rightarrow$ 0.4 RP	<b>Effet moyen plus forts pour les plus faibles.</b>

Sources : Papiers cités, auteur.

Notes : LEM : linéaire en moyennes. UET : unité d'écart-type. RP : rang de percentiles.

- Les autres résultats semblent dépendre du système éducatif : voir [Duflo, Dupas, and Kremer \(2011\)](#) vs. [Sojourner \(2013\)](#).

## Éléments institutionnels et données

## Le Cours Moyen 2<sup>ème</sup> année (CM2)

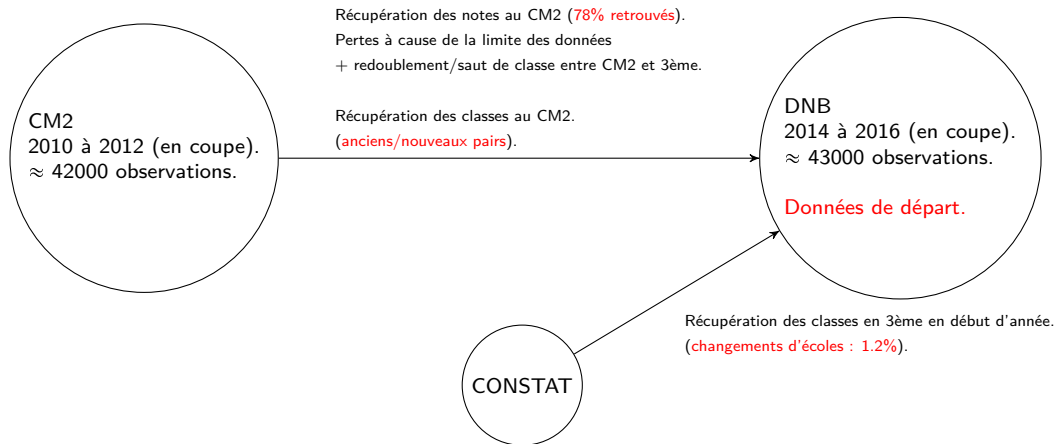
- ▶ Fin (5<sup>ème</sup> année) du primaire.
- ▶ Année scolaire : de mi-Août environ à début Juillet.
- ▶ En moyenne 2 classes par école.
- ▶ Classes multiniveaux (CM1 et CM2 dans la même salle de classe).
- ▶ Vers fin Janvier sauf exception : **évaluation nationale des acquis de CM2.**

## La 3ème

- ▶ Fin (4<sup>ème</sup> année) du collège.
- ▶ Année scolaire : de mi-Août environ à début Juillet.
- ▶ En moyenne 7 classes par établissement.
- ▶ Pas de classes multiniveaux.
- ▶ Sections spécifiques pour les élèves en grande difficulté (ex : *Section d'Enseignement Général et Professionnel Adapté - SEGPA*) ou exceptionnellement bons (ex : options internationales). Certaines classes accueillent exclusivement l'une ou l'autre (**classes atypiques**).
- ▶ Vers fin Juin : **épreuves pour obtenir le Diplôme National de Brevet (DNB)**.



# Construction des données



- **Fichiers CONSTAT :** Fichiers établis en début d'année, recouvrant toutes les écoles, collèges et lycées ainsi que tous les élèves qui y sont inscrits. Contient les informations au niveau individuel, identifiants de classes et informations au niveau de l'établissement. Ne contient pas de données de résultats scolaires.

### 3 échantillons

- ▶ Données complètes ( $\approx 43000$  observations) :
  - ▶ 14% des classes sont des classes atypiques inférieures.
  - ▶ 3% des classes sont des classes atypiques supérieures.
- ▶ Données sans classes atypiques ( $\approx 36000$  observations,  $\approx 84\%$  des données complètes).
- ▶ Données sans classes atypiques après restrictions statistiques ([Brodaty and Gurgand 2016](#) ; [Boutchenik and Maillard 2018](#)) : tests KS<sup>3</sup> ( $\approx 34000$  observations,  $\approx 95\%$  des données sans classes atypiques).
  - ▶ Les tests permettent *a priori* de détecter d'éventuels regroupements supplémentaires des élèves selon le niveau initial.
  - ▶ Échantillon pour les estimations.

---

<sup>3</sup>Kolmogorov-Smirnov.

# Statistiques descriptives

	Données complètes (1)	(1) sans classes atypiques (2)	(2) après tests KS (3)
Nombre d'établissements	84	82	79
Nombre de classes par établissement	7.075 (2.226)	6.016 (1.895)	5.97 (1.859)
Taille de classe	24.528 (3.196)	24.803 (3.002)	24.753 (2.989)
Note au DNB	49.633 (29.077)	49.31 (29.853)	49.1 (29.841)
Proportion de retrouvés	0.775	0.776	0.778
Note au CM2	51.95 (27.613)	53.935 (27.067)	53.656 (26.96)
Proportion de nouveaux pairs	0.919	0.92	0.921
Note au CM2 des nouveaux pairs	51.843 (15.705)	53.802 (13.529)	53.517 (13.128)
Note au CM2 des anciens pairs	32.404 (32.302)	33.493 (32.672)	33.136 (32.496)
Proportion de changements d'établissement	0.012	0.012	0.012
Note au CM2, changement d'établissement	40.711 (25.063)	41.3 (25.086)	40.706 (24.783)
Note au CM2, pas de changement d'établissement	52.064 (27.614)	54.063 (27.057)	53.789 (26.95)
Observations	42686	35978	34179

*Sources* : Fichiers DNB (2014 à 2016), fichiers CM2 (2010 à 2012), fichiers CONSTAT, calculs de l'auteur.

*Notes* : Toutes les notes sont exprimées en rang de percentiles. Les moyennes de notes au CM2 sont calculées sur les retrouvés uniquement.

Colonne (3) : échantillon principal des estimations. Écart-types entre parenthèses. KS : Kolmogorov-Smirnov.

- ▶ Entre (1) et (2) : on enlève des classes (17%).
- ▶ Augmentation de la note au CM2 entre (1) et (2) : prévalence des classes atypiques inférieures dans (1).

# Statistiques descriptives

	Données complètes (1)	(1) sans classes atypiques (2)	(2) après tests KS (3)
Nombre d'établissements	84	82	79
Nombre de classes par établissement	7.075 (2.226)	6.016 (1.895)	5.97 (1.859)
Taille de classe	24.528 (3.196)	24.803 (3.002)	24.753 (2.989)
Note au DNB	49.633 (29.077)	49.31 (29.853)	49.1 (29.841)
Proportion de retrouvés	0.775	0.776	0.778
Note au CM2	51.95 (27.613)	53.935 (27.067)	53.656 (26.96)
Proportion de nouveaux pairs	0.919	0.92	0.921
Note au CM2 des nouveaux pairs	51.843 (15.705)	53.802 (13.529)	53.517 (13.128)
Note au CM2 des anciens pairs	32.404 (32.302)	33.493 (32.672)	33.136 (32.496)
Proportion de changements d'établissement	0.012	0.012	0.012
Note au CM2, changement d'établissement	40.711 (25.063)	41.3 (25.086)	40.706 (24.783)
Note au CM2, pas de changement d'établissement	52.064 (27.614)	54.063 (27.057)	53.789 (26.95)
Observations	42686	35978	34179

*Sources* : Fichiers DNB (2014 à 2016), fichiers CM2 (2010 à 2012), fichiers CONSTAT, calculs de l'auteur.

*Notes* : Toutes les notes sont exprimées en rang de percentiles. Les moyennes de notes au CM2 sont calculées sur les retrouvés uniquement.

Colonne (3) : échantillon principal des estimations. Écart-types entre parenthèses. KS : Kolmogorov-Smirnov.

- ▶ Entre (2) et (3) : on enlève des établissements (trois).
- ▶ Pas de changement considérable des chiffres entre (2) et (3).

# Statistiques descriptives

	Données complètes (1)	(1) sans classes atypiques (2)	(2) après tests KS (3)
Nombre d'établissements	84	82	79
Nombre de classes par établissement	7.075 (2.226)	6.016 (1.895)	5.97 (1.859)
Taille de classe	24.528 (3.196)	24.803 (3.002)	24.753 (2.989)
Note au DNB	49.633 (29.077)	49.31 (29.853)	49.1 (29.841)
Proportion de retrouvés	0.775	0.776	0.778
Note au CM2	51.95 (27.613)	53.935 (27.067)	53.656 (26.96)
Proportion de nouveaux pairs	0.919	0.92	0.921
Note au CM2 des nouveaux pairs	51.843 (15.705)	53.802 (13.529)	53.517 (13.128)
Note au CM2 des anciens pairs	32.404 (32.302)	33.493 (32.672)	33.136 (32.496)
Proportion de changements d'établissement	0.012	0.012	0.012
Note au CM2, changement d'établissement	40.711 (25.063)	41.3 (25.086)	40.706 (24.783)
Note au CM2, pas de changement d'établissement	52.064 (27.614)	54.063 (27.057)	53.789 (26.95)
Observations	42686	35978	34179

*Sources* : Fichiers DNB (2014 à 2016), fichiers CM2 (2010 à 2012), fichiers CONSTAT, calculs de l'auteur.

*Notes* : Toutes les notes sont exprimées en rang de percentiles. Les moyennes de notes au CM2 sont calculées sur les retrouvés uniquement.

Colonne (3) : échantillon principal des estimations. Écart-types entre parenthèses. KS : Kolmogorov-Smirnov.

- La part des retrouvés, la proportion de nouveaux pairs et la proportion d'élèves qui changent d'école ne dépendent pas du type de classes.

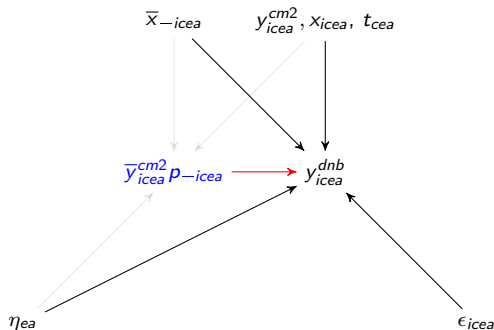
# Modélisation

# Modèle économétrique : notations

- ▶  $i, c, e, a$  : indices des individu, classe, école et année.
- ▶  $y_{icea}^{dnb}$  : note au DNB en rang de percentiles, variable dépendante. Les rangs de percentiles sont calculés par année de DNB.
- ▶  $y_{icea}^{cm2}$  : note au CM2 en rang de percentiles. Les rangs de percentiles sont calculés par année de CM2.
- ▶  $\bar{y}_{icea}^{cm2}$  : note au CM2 des pairs (de 3<sup>ème</sup>), calculée sur les retrouvés uniquement.
- ▶  $Qn_{icea}$  : indicateur d'appartenance au quintile  $n$  de la note au CM2, calculée sur les retrouvés uniquement.
- ▶  $\bar{Q}m_{icea}$  : proportion de pairs au quintile  $m$  de la note au CM2, calculée sur les retrouvés uniquement.
- ▶  $p_{icea}$  : proportion de retrouvés chez les pairs.
- ▶  $x_{icea}$  : variables de contrôle au niveau individuel.
- ▶  $\bar{x}_{icea}$  : variables de contrôle chez (tous) les pairs.
- ▶  $t_{cea}$  : taille de classe.
- ▶  $\eta_{ea}$  : effet établissement-année.
- ▶  $\epsilon_{icea}$  : terme d'erreur.

## Sojourner (2013), effets de pairs linéaires en moyennes

$$y_{icea}^{dnb} = \beta \bar{y}_{icea}^{cm2} p_{-icea} + \iota y_{icea}^{cm2} + \pi p_{-icea} + x_{icea} \gamma + \delta t_{cea} + \bar{x}_{-icea} \theta + \eta_{ea} + \epsilon_{icea} \quad (1)$$

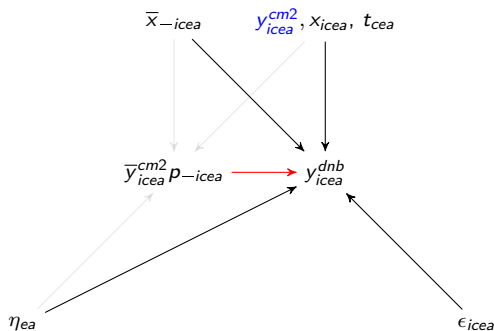


- Prise en compte du niveau **passé (CM2)** des pairs (en 3ème) : permet de contourner le problème de réflexion (Yeung and Nguyen-Hoang 2016).
- Spécificité du modèle : prise en compte des données manquantes via  $p_{-icea}$ .



## Sojourner (2013), effets de pairs linéaires en moyennes

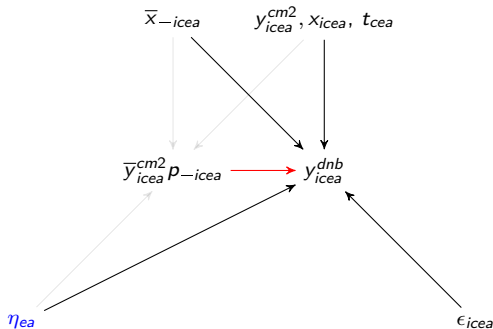
$$y_{icea}^{dnb} = \beta \bar{y}_{icea}^{cm2} p_{-icea} + \iota y_{icea}^{cm2} + \pi p_{-icea} + x_{icea} \gamma + \delta t_{cea} + \bar{x}_{icea} \theta + \eta_{ea} + \epsilon_{icea} \quad (2)$$



- $y_{icea}^{cm2}$  : prise en compte partielle de la formation endogène des pairs (Davezies 2004 ; Sojourner 2013). Biais de surestimation si non.

## Sojourner (2013), effets de pairs linéaires en moyennes

$$y_{icea}^{dnb} = \beta \bar{y}_{icea}^{cm2} p_{icea} + \iota y_{icea}^{cm2} + \pi p_{icea} + x_{icea} \gamma + \delta t_{cea} + \bar{x}_{icea} \theta + \eta_{ea} + \epsilon_{icea} \quad (3)$$



- $\eta_{ea}$  : prise en compte de la partie inobservée spécifique à l'établissement-année de laquelle peut résulter une sélection endogène (moyens, stratégie parentale, sélectivité de l'établissement).
- S'appuie sur la variation intra établissement, inter classes : nécessite l'hypothèse d'assignation aléatoire des classes au sein d'un établissement-année : d'où les restrictions d'échantillon.

## Cependant, le modèle linéaire en moyennes est très limité

- ▶ L'effet estimé ne sert pas en termes de politiques publiques. Si ce modèle est vrai :
  - ▶ Le gain de l'arrivée d'un bon camarade dans une classe est exactement compensée par la perte due à son départ dans sa classe d'origine ([Brodaty 2010](#)).
  - ▶ La majorité des papiers rejettent empiriquement ce modèle ([Hoxby and Weingarth 2005](#), par exemple).
- ▶ D'où les modèles dits
  - ▶ **hétérogènes** : l'effet de pair dépend du niveau de l'individu.
  - ▶ **non linéaires** : l'effet de pair dépend du niveau des pairs.
- ▶ Les fondations ( $p_{-icea}, \eta_{ea}$ ) sont maintenues dans ces modèles plus flexibles ([Sojourner 2013](#)).

Sojourner (2013), effets de pairs hétérogènes (dépendent du niveau de l'individu)

$$y_{icea}^{dnb} = \sum_{n=1}^5 \beta_{Qn} (\bar{y}_{-icea}^{cm2} p_{-icea}) Qn_{icea} + \sum_{n=2}^5 \iota Qn_{icea} + \pi p_{-icea} + x_{icea} \gamma + \delta t_{cea} + \bar{x}_{-icea} \theta + \eta_{ea} + \epsilon_{icea} \quad (4)$$

- Exemple :  $\beta_{Q1}$  désigne l'effet de l'augmentation d'une unité de la note au CM2 des pairs pour un élève au Q1 (très faible).

Sojourner (2013), effets de pairs hétérogènes et non linéaires (dépendent du niveau de l'individu et des pairs)

$$y_{icea}^{dnb} = \sum_{m \in \{1,2,4,5\}} \left[ \sum_{n=1}^5 \beta_{\bar{Q}m, Qn} (\bar{Q}m_{-icea} p_{-icea}) Qn_{icea} \right] + \sum_{n=2}^5 \iota_{Qn} Qn_{icea} + \pi p_{-icea} \quad (5)$$

$$+ x_{icea} \gamma + \delta t_{cea} + \bar{x}_{-icea} \theta + \eta_{ea} + \epsilon_{icea}$$

- Exemple :  $\beta_{\bar{Q}5, Q1}$  désigne l'effet, l'effet du passage de 0% à 100% de la part de pairs au Q5 (très forts) contre des pairs au Q3 (moyen) pour un individu au Q1 (très faible).

## Résultats

# Résultats : effets de pairs linéaires en moyenne

	Variable dépendante : note au DNB	
	(1)	(2)
Note CM2	0.623*** (0.008)	0.641*** (0.007)
Note CM2 pairs	0.224*** (0.035)	0.346*** (0.032)
Indicateurs de quintiles individuels	✓	✓
Effets établissement-année	✓	✓
Contrôles individuels	✓	✓
Taille de classe	✓	✓
Contrôles chez les pairs	✓	✓
Changement d'école	✓	✓
Nouveaux pairs/anciens pairs	x	✓
N	26016	26016
R <sup>2</sup> ajusté	0.513	0.52

Sources : Fichiers DNB(2014 à 2016), fichiers CM2 (2010 à 2012) et fichiers CONSTAT, calculs de l'auteur.

Notes : Les notes sont exprimées en rang de percentiles. La variable Note CM2 pairs n'est pas la même dans (1) et (2) mais est présentée ainsi par souci de visualisation. Dans (2), il s'agit de la note au CM2 des nouveaux pairs.

Significativité : 10% \* 5% \*\* 1% \*\*\*. Écart-types robustes entre parenthèses.

- ▶ 1 RP note CM2 pairs → 0.2 RP ( $\approx 1/3$  coefficient de la note au CM2) note DNB.
- ▶ L'effet passe à 0.35 RP ( $\approx 1/2$  coefficient de la note au CM2) lorsque l'on prend en compte les nouveaux/anciens pairs.

# Résultats : effets de pairs linéaires en moyenne

	Variable dépendante : note au DNB	
	(1)	(2)
Note CM2	0.623*** (0.008)	0.641*** (0.007)
Note CM2 pairs	0.224*** (0.035)	0.346*** (0.032)
Indicateurs de quintiles individuels	✓	✓
Effets établissement-année	✓	✓
Contrôles individuels	✓	✓
Taille de classe	✓	✓
Contrôles chez les pairs	✓	✓
Changement d'école	✓	✓
Nouveaux pairs/anciens pairs	✗	✓
N	26016	26016
R <sup>2</sup> ajusté	0.513	0.52

Sources : Fichiers DNB(2014 à 2016), fichiers CM2 (2010 à 2012) et fichiers CONSTAT, calculs de l'auteur.

Notes : Les notes sont exprimées en rang de percentiles. La variable Note CM2 pairs n'est pas la même dans (1) et (2) mais est présentée ainsi par souci de visualisation. Dans (2), il s'agit de la note au CM2 des nouveaux pairs.

Significativité : 10% \* 5% \*\* 1% \*\*\*. Écart-types robustes entre parenthèses.

- (2) : De même ordre de grandeur que les coefficients de [Sojourner \(2013\)](#). Mais pas [Boutchenik and Maillard \(2018\)](#) : niveau d'éducation ? territoire ?.



# Résultats : effets de pairs hétérogènes

Variable dépendante : note au DNB

	(1)	(2)
Note CM2 pairs x Q1	0.04 (0.04)	0.172*** (0.039)
Note CM2 pairs x Q2	0.182*** (0.042)	0.305*** (0.039)
Note CM2 pairs x Q3	0.315*** (0.041)	0.442*** (0.04)
Note CM2 pairs x Q4	0.337*** (0.042)	0.459*** (0.041)
Note CM2 pairs x Q5	0.256*** (0.048)	0.342*** (0.044)
Effets établissement-année	✓	✓
Contrôles individuels	✓	✓
Taille de classe	✓	✓
Contrôles chez les pairs	✓	✓
Changement d'école	✓	✓
Nouveaux pairs/anciens pairs	x	✓
N	26016	26016
R <sup>2</sup> ajusté	0.497	0.503

Sources : Fichiers DNB(2014 à 2016), fichiers CM2 (2010 à 2012) et fichiers CONSTAT, calculs de l'auteur.

Notes : Les notes sont exprimées en rang de percentiles. Les quintiles sont calculés sur les candidats au DNB, par année. La variable Note CM2 pairs n'est pas la même dans (1) et (2) mais est présentée ainsi par souci de visualisation. Dans (2), il s'agit de la note au CM2 des nouveaux pairs. Significativité : 10% \* 5% \*\* 1% \*\*\*. Écart-types robustes entre parenthèses.

- ▶ Effet croissant de manière monotone avec le niveau au CM2 jusqu'à Q4.
- ▶ Effets dans (2) > Effets dans (1).

# Résultats : effets de pairs hétérogènes

	Variable dépendante : note au DNB	
	(1)	(2)
Note CM2 pairs x Q1	0.04 (0.04)	0.172*** (0.039)
Note CM2 pairs x Q2	0.182*** (0.042)	0.305*** (0.039)
Note CM2 pairs x Q3	0.315*** (0.041)	0.442*** (0.04)
Note CM2 pairs x Q4	0.337*** (0.042)	0.459*** (0.041)
Note CM2 pairs x Q5	0.256*** (0.048)	0.342*** (0.044)
Effets établissement-année	✓	✓
Contrôles individuels	✓	✓
Taille de classe	✓	✓
Contrôles chez les pairs	✓	✓
Changement d'école	✓	✓
Nouveaux pairs/anciens pairs	x	✓
N	26016	26016
R <sup>2</sup> ajusté	0.497	0.503

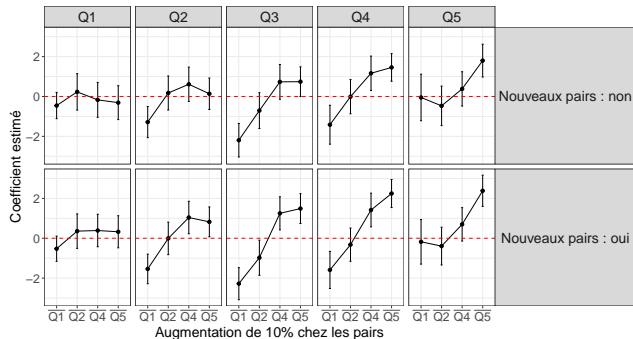
Sources : Fichiers DNB(2014 à 2016), fichiers CM2 (2010 à 2012) et fichiers CONSTAT, calculs de l'auteur.

Notes : Les notes sont exprimées en rang de percentiles. Les quintiles sont calculés sur les candidats au DNB, par année. La variable Note CM2 pairs n'est pas la même dans (1) et (2) mais est présentée ainsi par souci de visualisation. Dans (2), il s'agit de la note au CM2 des nouveaux pairs. Significativité : 10% \* 5% \*\* 1% \*\*\*. Écart-types robustes entre parenthèses.

- Pas en accord avec la majorité des papiers, sauf [Duflo, Dupas, and Kremer \(2011\)](#) (Kenya) : contexte socio-économique similaire ?.

# Résultats : effets de pairs hétérogènes et non linéaires

Variable dépendante : note au DNB.



Sources : Fichiers DNB (2014 à 2016), fichiers CM2 (2010 à 2012), fichiers, calculs de l'auteur.

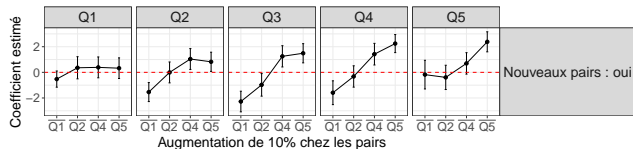
Notes : Les quintiles sont calculées par année (DNB) sur les données complètes. N = 26015 dans les deux lignes de panels.

Intervalle de confiance à 95% à partir d'écart-types robustes.

- ▶ Pas d'effet considérable (ampleur, significativité) pour les individus les plus faibles (colonne de panels Q1).
- ▶ Les pairs les plus faibles ont un effet négatif sur tous les individus ayant un niveau qui leur sont supérieurs, sauf les plus forts (Q5). Voir  $\bar{Q}1$ , colonnes de panels Q2 à Q4.
- ▶ Les pairs les plus forts ont un effet positif sur les individus au moins de niveau moyen.

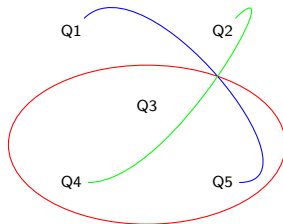
# Implications (spéculatif)

Variable dépendante : note au DNB.



Sources : Fichiers DNB (2014 à 2016), fichiers CM2 (2010 à 2012), fichiers CONSTAT, calculs de l'auteur.

Notes : Les quintiles sont calculées par année (DNB) sur les données complètes. N = 26015 dans les deux lignes de panels. Intervalles de confiance à 95% à partir d'écart-types robustes.



- ▶ Regrouper **Q3**, **Q4**, **Q5S** équitablement entre eux.
- ▶ Favoriser les combinaisons (**Q1**, **Q5**) et (**Q2**, **Q4**).

## Conclusion

## En résumé

- ▶ Objectif : Mesurer l'influence du niveau initial des pairs sur les résultats en fin de collège à La Réunion.
- ▶ Méthodologie :
  - ▶ Forme réduite (niveau initial des pairs).
  - ▶ Effets établissement (variation intra établissement inter classes).
  - ▶ Restriction d'échantillon : suppression des classes atypiques et tests statistiques.
  - ▶ Prise en compte des valeurs manquantes sur le niveau initial des pairs : [Sojourner \(2013\)](#).
- ▶ Résultats :
  - ▶ Un effet linéaire en moyennes positif et significatif, en cohérence avec la littérature.
  - ▶ Un effet croissant avec le niveau de l'élève jusqu'à un certain niveau, en contradiction avec la littérature, sauf [Duflo, Dupas, and Kremer \(2011\)](#).
  - ▶ Des effets hétérogènes et non linéaires incitant à regrouper les plus forts entre eux puis à départager les plus faibles dans ces groupes en respectant quelques règles.

# Discussions

- ▶ Limites (liste non exhaustive) :
  - ▶ Absence de modèle général englobant les 3 spécifications (linéaires en moyenne, hétérogène, hétérogène et non linéaire).
  - ▶ Pas de données sur les enseignants (biais de surestimation si les meilleurs enseignants affectés aux classes de bon niveau).
  - ▶ Manque d'information explicite sur la pratique de regroupement des élèves selon leur niveau de la part des établissements.
- ▶ Perspectives (liste non exhaustive) :
  - ▶ Tests de robustesses.
  - ▶ Confrontation avec les résultats des études comparables.

Merci de votre attention.



## Bibliographie I

- Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30: 98–108.
- Arneton, Mélissa, Christine Bocéran, and André Flieller. 2013. "Les Performances En Mathématiques Des élèves Des départements d'outre-Mer." *L'orientation Scolaire Et Professionnelle*, no. 42/1.
- Boucher, Vincent, Yann Bramoullé, Habiba Djebbari, and Bernard Fortin. 2014. "Do Peers Affect Student Achievement? Evidence from Canada Using Group Size Variation." *Journal of Applied Econometrics* 29 (1): 91–109.
- Boutchenik, Béatrice, and Sophie Maillard. 2018. "Peer Effects with Peer and Student Heterogeneity: An Assessment for French Baccalaureat." *Contribution Aux 13es Journées de Méthodologie Statistique de l'Insee*, 12–14.
- Brodaty, Thibault. 2010. "Les Effets de Pairs Dans l'Éducation: Une Revue de Littérature." *Revue d'économie Politique* 120 (5): 739–57.

## Bibliographie II

- Brodaty, Thibault, and Marc Gurgand. 2016. "Good Peers or Good Teachers? Evidence from a French University." *Economics of Education Review* 54: 62–78.
- Davezies, Laurent. 2004. "Influence Des Caractéristiques Du Groupe Des Pairs Sur La Scolarité élémentaire." *Éducation Et Formations* 72: 171.
- Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101 (5): 1739–74.
- Hoxby, Caroline M, and Gretchen Weingarth. 2005. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Working paper.
- Lavy, Victor, Olmo Silva, and Felix Weinhardt. 2012. "The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools." *Journal of Labor Economics* 30 (2): 367–414.

## Bibliographie III

- Lee, Lung-fei. 2007. "Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects." *Journal of Econometrics* 140 (2): 333–74.
- Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–42.
- Monso, Olivier, Denis Fougere, Pauline Givord, and Claudine Pirus. 2019. "Les Camarades Influent-ils La réussite Et Le Parcours Des élèves?"
- Sacerdote, Bruce. 2011. "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In *Handbook of the Economics of Education*, 3:249–77. Elsevier.
- Sojourner, Aaron. 2013. "Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR." *The Economic Journal* 123 (569): 574–605.

## Bibliographie IV

Yeung, Ryan, and Phuong Nguyen-Hoang. 2016. "Endogenous Peer Effects: Fact or Fiction?" *The Journal of Educational Research* 109 (1): 37–49.