

# Construction d'un indice mesurant la position sociale des élèves à partir des PCS de leurs parents

Thierry Rocher

DEPP

Journées de méthodologie statistique, 31 mars 2015

# Motivations

Point de départ : étude des relations entre la performance scolaire et l'origine sociale

Deux besoins :

- Une variable qui synthétise plusieurs dimensions (sociale, économique, scolaire)
- Une variable quantitative pour passer facilement du niveau de l'élève à celui de l'établissement, de la classe

Variable centrale en éducation : la profession et catégorie socioprofessionnelle (PCS)

→ construction d'un indice dérivé de la PCS

# PCS utilisée au MEN

10	Agriculteurs exploitants
21	Artisans
22	Commerçants et assimilés
23	Chefs d'entreprise de 10 salariés ou plus
31	Professions libérales et assimilés
33	Cadres de la fonction publique
34	Professeurs, professions scientifiques
35	Professions de l'information, des arts et des spectacles
37	Cadres administratifs et commerciaux d'entreprise
38	Ingénieurs et cadres techniques d'entreprise
42	Professeurs des écoles, instituteurs et assimilés
43	Professions intermédiaires de la santé et du travail social
44	Clergé, religieux
45	Professions intermédiaires administratives de la fonction publique
46	Professions intermédiaires administratives et commerciales des entreprises
47	Techniciens
48	Contremaîtres, agents de maîtrise
52	Employés civils et agents de service de la fonction publique
53	Policiers et militaires
54	Employés administratifs d'entreprise
55	Employés de commerce
56	Personnels des services directs aux particuliers
61	Ouvriers qualifiés
66	Ouvriers non qualifiés
69	Ouvriers agricoles
71	Anciens prof. Intermédiaires et supérieures
72	Anciens artisans, commerçants, chefs d'entreprise
73	Anciens cadres et professions intermédiaires
76	Anciens employés et ouvriers
81	Chômeurs n'ayant jamais travaillé
82	Inactifs divers (autres que retraités)

32 postes :  
 Nomenclature  
 24 postes INSEE  
 + détails 42 postes

# PCS utilisée au MEN

10	Agriculteurs exploitants
21	Artisans
22	Commerçants et assimilés
23	Chefs d'entreprise de 10 salariés ou plus
31	Professions libérales et assimilés
33	Cadres de la fonction publique
34	Professeurs, professions scientifiques
35	Professions de l'information, des arts et des spectacles
37	Cadres administratifs et commerciaux d'entreprise
38	Ingénieurs et cadres techniques d'entreprise
42	Professeurs des écoles, instituteurs et assimilés
43	Professions intermédiaires de la santé et du travail social
44	Clergé, religieux
45	Professions intermédiaires administratives de la fonction publique
46	Professions intermédiaires administratives et commerciales des entreprises
47	Techniciens
48	Contremaîtres, agents de maîtrise
52	Employés civils et agents de service de la fonction publique
53	Policiers et militaires
54	Employés administratifs d'entreprise
55	Employés de commerce
56	Personnels des services directs aux particuliers
61	Ouvriers qualifiés
66	Ouvriers non qualifiés
69	Ouvriers agricoles
71	Anciens prof. Intermédiaires et supérieures
72	Anciens artisans, commerçants, chefs d'entreprise
73	Anciens cadres et professions intermédiaires
76	Anciens employés et ouvriers
81	Chômeurs n'ayant jamais travaillé
82	Inactifs divers (autres que retraités)

Classes DEPP :

Favorisées A

Favorisées B

Moyennes

Défavorisées

Hierarchisation :

« distance scolaire »

# Limite : décrire des groupes

Limite des typologies actuelles pour décrire la tonalité sociale de groupes d'élèves (classes, établissements, académies) :

- tendance : niveau social
- dispersion : mixité sociale

Indicateur usuel : un pourcentage d'élèves d'une certaine catégorie sociale (ex. : % d'élèves défavorisés)

→ un indice quantitatif

Une mesure unidimensionnelle de la stratification sociale :

- En théorie, une vision discutable de l'univers social ( $\neq$  tradition sociologique française)
- En pratique, un outil très utile

# Principe de construction

Idée : attribuer une valeur numérique à chaque PCS

- Moyenne pondérée (combinaison linéaire) de caractéristiques par PCS
- Techniquement : score factoriel (1er axe d'une ACM) moyen par PCS

⇒ modèle sociologique

Exemples simples : revenu moyen par PCS, nombre d'années d'études par PCS, score moyen par PCS

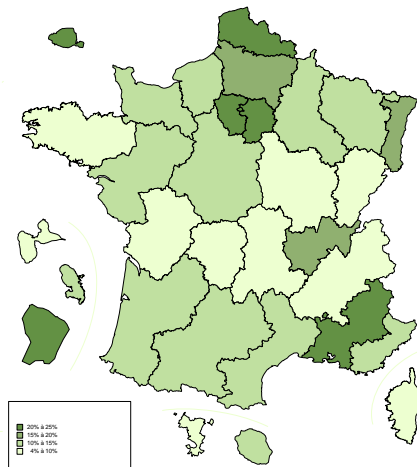
Travaux similaires :

- Indice socio-économique (Ganzeboom, 1992)
- Indice de statut social (Chambaz et al., 1998)

# Exemple d'application

Version 0 : indice de synergie socio-scolaire (Le Donné & Rocher, 2010), refonte en cours

Part de variance expliquée par le collège - Public et Privé sous contrat



France métropolitaine + DOM : 19%

# Indice rénové

Même principe : la valeur de l'indice correspond à la moyenne pondérée de caractéristiques sociales par PCS

## Aspects revisités

- La méthode retenue pour calculer l'indice
- Le choix des variables entrant en compte dans sa construction

## Les données utilisées :

- Panel DEPP : 35 000 élèves entrant en 6e recrutés à la rentrée 2007
- Une description très fine du contexte familial (questionnaire famille 2008)
- Des tests standardisés mesurant les acquis cognitifs et conatifs des élèves (épreuves 2008)



# Choix des variables

Indice de position socio-scolaire : l'indice doit synthétiser un ensemble de variables susceptibles d'influer sur la réussite scolaire

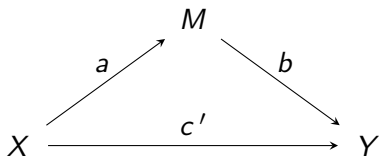
Concept de **médiation** :

- Les PCS n'agissent pas directement sur les résultats scolaires mais *via* un certain nombre de caractéristiques
- La catégorie sociale est le reflet de capitaux (économiques, éducatifs, culturels), de pratiques et d'attitudes qui tendent à plus ou moins favoriser les conditions d'apprentissage, la réussite scolaire, etc.
- L'« effet » de la PCS des parents sur la réussite scolaire des enfants transite par cet ensemble de variables dites « médiatrices »

# Médiation - Principe

$$X \xrightarrow{c} Y$$

L'effet de  $X$  sur  $Y$  est « médiatisé » par une variable  $M$  :



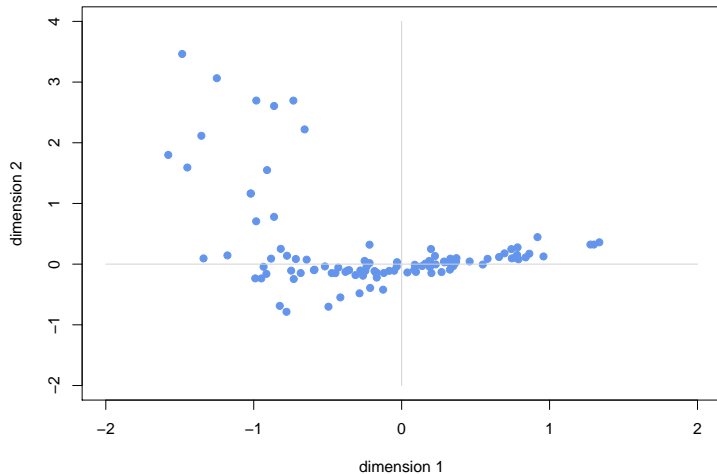
Quatre conditions :

- 1 La variable causale  $X$  est corrélée à la variable de résultat  $Y$
- 2 La variable causale  $X$  est corrélée à la variable médiatrice  $M$
- 3 La variable médiatrice  $M$  affecte la variable de résultat  $Y$ , lorsque la variable causale  $X$  est contrôlée
- 4 L'effet de  $X$  sur  $Y$  est atténué, voire annulé, lorsque l'on contrôle par  $M$  : la différence  $c - c'$  est positive et significative.

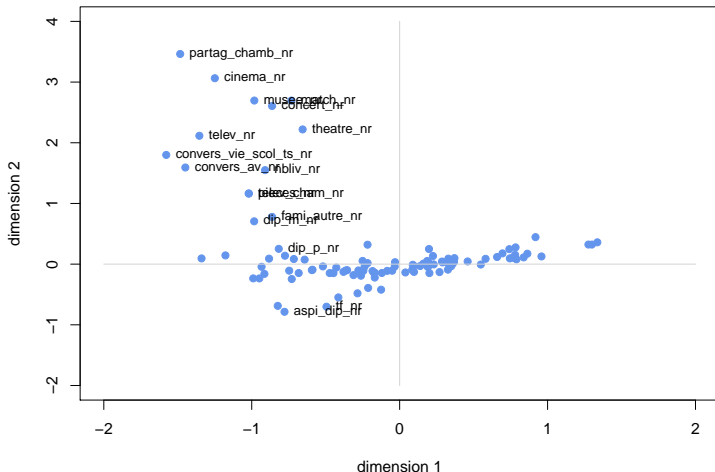
# Variables retenues

Variables	Modalités	Variables	Modalités
<b>Diplômes des parents</b>		<b>Ambition et implication</b>	
Diplôme mère	9	Aspiration	6
Diplôme père	9	Diplôme le plus utile	5
<b>Conditions matérielles</b>		Implication des parents	5
Revenus mensuels du foyer	10	Conversations (vie scolaire)	3
Nombre de pièces du logement	6	Conversations (avenir scolaire)	3
Chambre seul	3	<b>Pratiques culturelles (6 derniers mois)</b>	
Ordinateur	2	Événement sportif	4
Internet	2	Concert	4
<b>Composition familiale</b>		Théâtre	4
Structure familiale	5	Cinéma	4
Taille fratrie	6	Musée	4
<b>Capital culturel</b>		Activités extra-scolaires	3
Nb de livres	4		
TV chambre	3		
TV régulièrement	3		

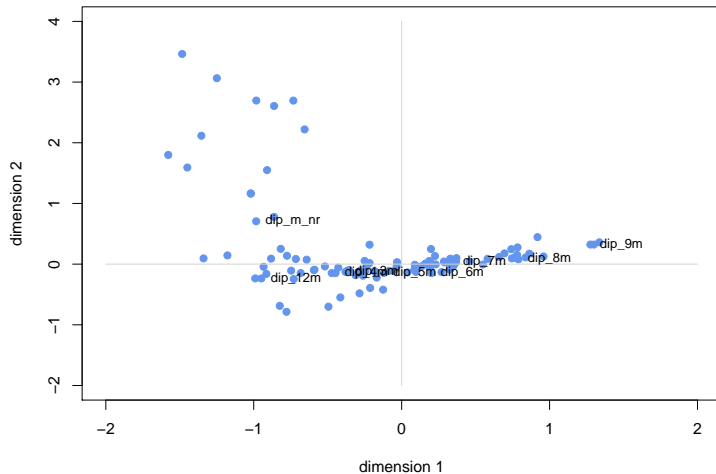
## ACM



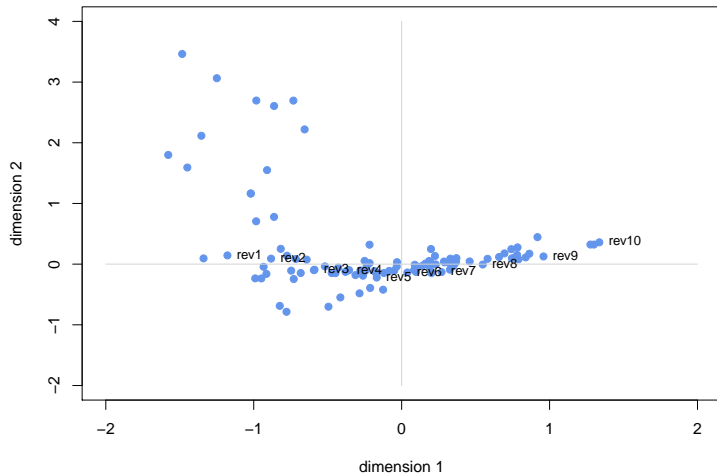
# ACM - indicatrices de non-réponses



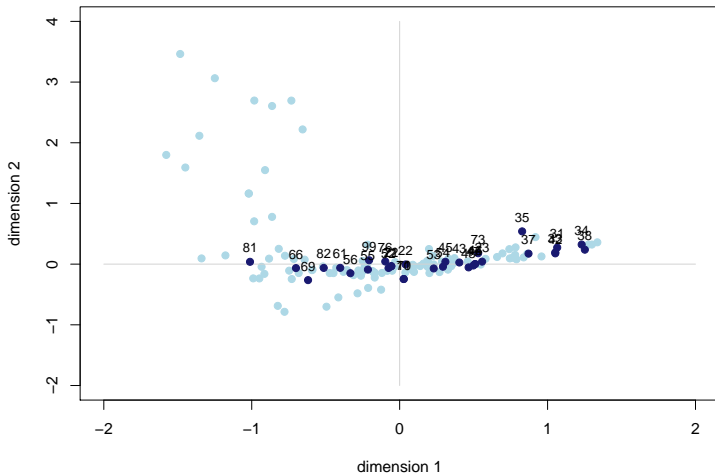
# Axe 1 ACM - diplôme père



# Axe 1 ACM - revenus

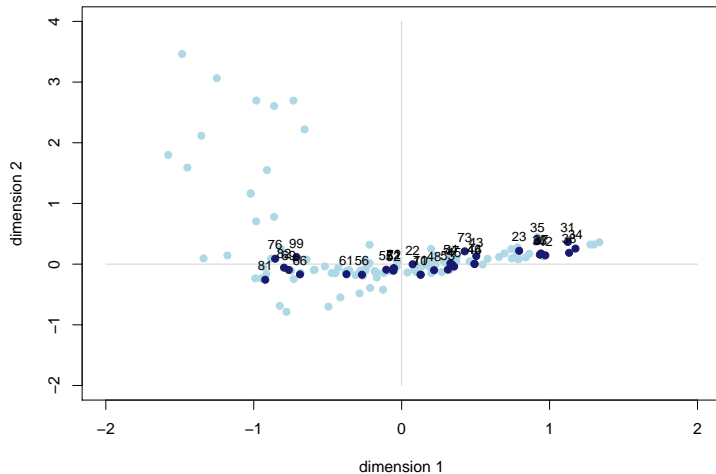


# Projection des PCS - mère





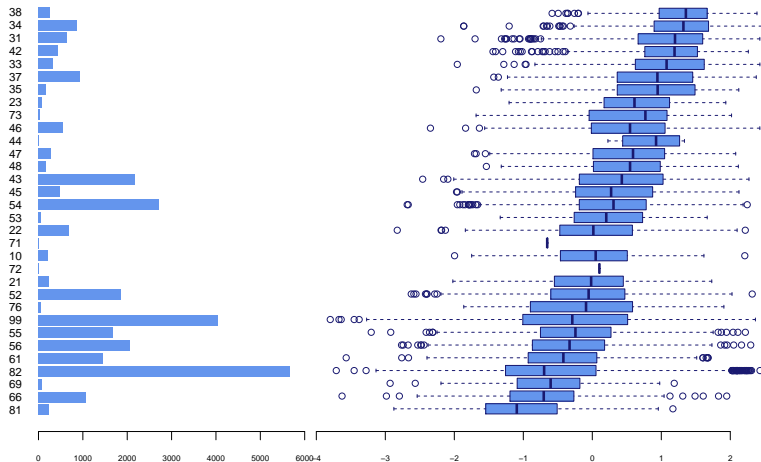
# Projection des PCS - père



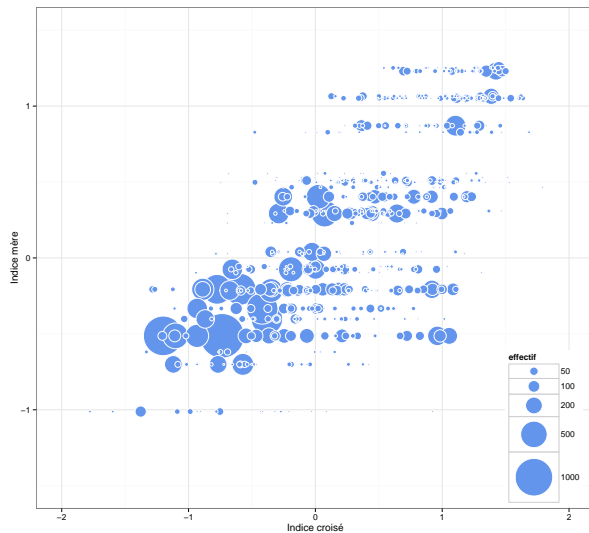
## Indices

PCS	Libellé	mère	père
10	Agriculteurs exploitants	0,03	0,13
21	Artisans	-0,06	-0,05
22	Commerçants et assimilés	0,04	0,07
23	Chefs d'entreprise de dix salariés ou plus	0,56	0,79
31	Professions libérales	1,06	1,12
33	Cadres de la fonction publique	1,05	0,93
34	Professeurs, professions scientifiques	1,23	1,18
35	Professions de l'information, des arts et des spectacles	0,83	0,92
37	Cadres administratifs et commerciaux d'entreprises	0,87	0,94
38	Ingénieurs et cadres techniques d'entreprises	1,25	1,13
42	Instituteurs et assimilés	1,05	0,97
43	Professions intermédiaires de la santé et du travail social	0,40	0,50
44	Clergé, religieux	0,51	0,49
45	Professions intermédiaires administratives de la fonction publique	0,31	0,35
46	Professions intermédiaires administratives commerciales en entreprise	0,51	0,49
47	Techniciens	0,50	0,34
48	Contremaîtres, agents de maîtrise	0,47	0,22
52	Employés civils et agents de service de la fonction publique	-0,07	-0,06
53	Policiers et militaires	0,23	0,31
54	Employés administratifs d'entreprises	0,29	0,33
55	Employés de commerce	-0,21	-0,10
56	Personnels des services directs aux particuliers	-0,33	-0,27
61	Ouvriers qualifiés	-0,40	-0,37
66	Ouvriers non qualifiés	-0,70	-0,69
69	Ouvriers agricoles	-0,62	-0,76
71	Retraités agriculteurs exploitants	0,03	0,13
72	Retraités artisans, commerçants, chefs entreprise	-0,06	-0,05
73	Retraités cadres et professions intermédiaires	0,53	0,43
76	Retraités employés et ouvriers	-0,10	-0,86
81	Chômeurs n'ayant jamais travaillé	-1,01	-0,92
82	Personnes diverses sans activité professionnelle	-0,51	-0,79
99	Non renseignée (inconnue ou sans objet)	-0,21	-0,71

# Dispersion (mère)



# Lien indice croisé $\times$ indice mère



# Autres méthodes

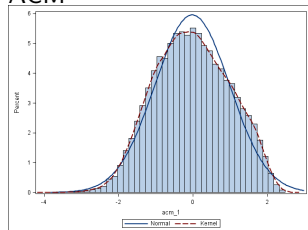
Le principe retenu consiste à synthétiser dans les PCS un ensemble de caractéristiques familiales ( $\mathbf{X}$ ), en lien avec les résultats scolaires ( $\mathbf{Y}$ ).

Selon ce principe, au-delà de l'ACM des variables  $\mathbf{X}$ , deux autres possibilités envisagées :

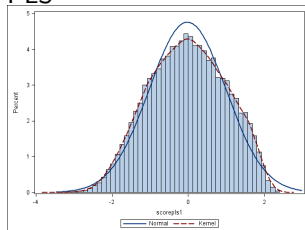
- L'analyse discriminante : déterminer la composante des  $\mathbf{X}$  qui discrimine le plus les PCS
- La régression PLS (*Partial Least Square*) : déterminer la composante des  $\mathbf{X}$  qui rend le mieux compte des covariances entre  $\mathbf{X}$  et  $\mathbf{Y}$

# Distributions

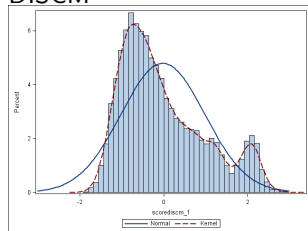
## ACM



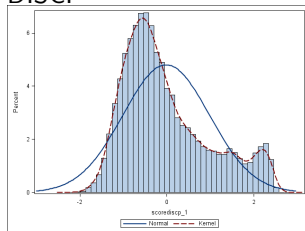
## PLS



## DISCM



## DISCP



# Travaux complémentaires

## Points de discussion

- Rapport construction-utilisation
- Influence de la méthode de calcul
- Influence du choix des variables
- Choix des PCS (nomenclatures, père/mère, responsable, etc.)
- Base de données utilisée
- Standardisation de l'indice
- ...

# Travaux complémentaires

## Points de discussion

- Rapport construction-utilisation
- Influence de la méthode de calcul
- Influence du choix des variables
- Choix des PCS (nomenclatures, père/mère, responsable, etc.)
- Base de données utilisée
- Standardisation de l'indice
- ...

Merci !