

Statistiques et informatique

appliquées aux sciences sociales

Régressions logistiques : étude d'un exemple

Mardi 28/11/2023 15h30-17h Censier D2

Licence de sciences sociales 3^e année
Université Paris 1 Panthéon Sorbonne

Exemple étudié ici :

Claude Dargent, 2004, « Chap. 7. La religion, encore et toujours »
in Bruno Cautrès (éd.), *Le nouveau désordre électoral: Les leçons du 21 avril 2002*, Paris: Presses de Sciences Po, 161-183.

« Les élections de 2002 sont l'occasion d'un bilan : dans ce contexte nouveau, quel rôle joue désormais l'appartenance religieuse dans l'explication du vote ? [...] Un premier bilan de l'impact de la variable religieuse en matière électorale a déjà été dressé sur la séquence des enquêtes CEVIPOF 1978-1988-1995 [Boy, Mayer, 1997a]. Les résultats de l'enquête 1997 ont été également mis en rapport avec des données plus anciennes provenant d'une autre source [Michelat, 2000] ; ces deux recherches établissent la persistance en France de la relation classique entre le vote et l'appartenance religieuse. Il est donc temps de confronter ces conclusions à la séquence d'enquêtes 1995-1997-2002, et d'observer dans quelle mesure elle prolonge ou non la période antérieure. »

Tableau I. L'appartenance religieuse en 2002

(en %)

	Effectifs	
Catholique	2 788	68
Protestante	94	2
Juive	22	1
Musulmane	81	2
Bouddhiste	14	0
Autre religion	41	1
Sans religion	1 039	25
Sans réponse	28	1
Total	4 107	

Source : PEF 2002, vague 1, non pondérée.

« Cette diminution de l'influence de l'Église catholique en France s'observe d'ailleurs dans bien d'autres domaines : les ordinations de prêtres (11 % d'entre eux seulement ont moins de 50 ans en 1999), les mariages (à peine plus d'un couple marié civilement sur deux – 56 % en 1999 – passe désormais à l'Église) et dans les baptêmes de l'année (un enfant sur deux seulement est baptisé aujourd'hui dans le catholicisme) [Lambert, 2002]. »

Tableau 2. La pratique religieuse des catholiques en France (en %)

	1995	2002
Plusieurs fois par semaine	1	2
Une fois par semaine	9	8
Une ou deux fois par mois	6	5
De temps en temps, aux grandes fêtes	21	20
Uniquement pour les cérémonies, mariages, etc.	58	58
Jamais	5	7

Sources : Enquêtes CEVIPOF 1995 et PEF 2002, vague 1, non pondérée.

Tableau 3. Appartenance et pratique religieuses en France (en %)

	1995	2002
Catholique pratiquant régulier	12	10
Catholique pratiquant irrégulier	16	14
Catholique non pratiquant	47	44
Autre religion	5	6
Sans religion	19	25
Sans réponse	1	1

Source : PEF 2002, vague 1, non pondérée.

Graphique 1. Intérêt fort pour la politique selon la pratique religieuse (en %)

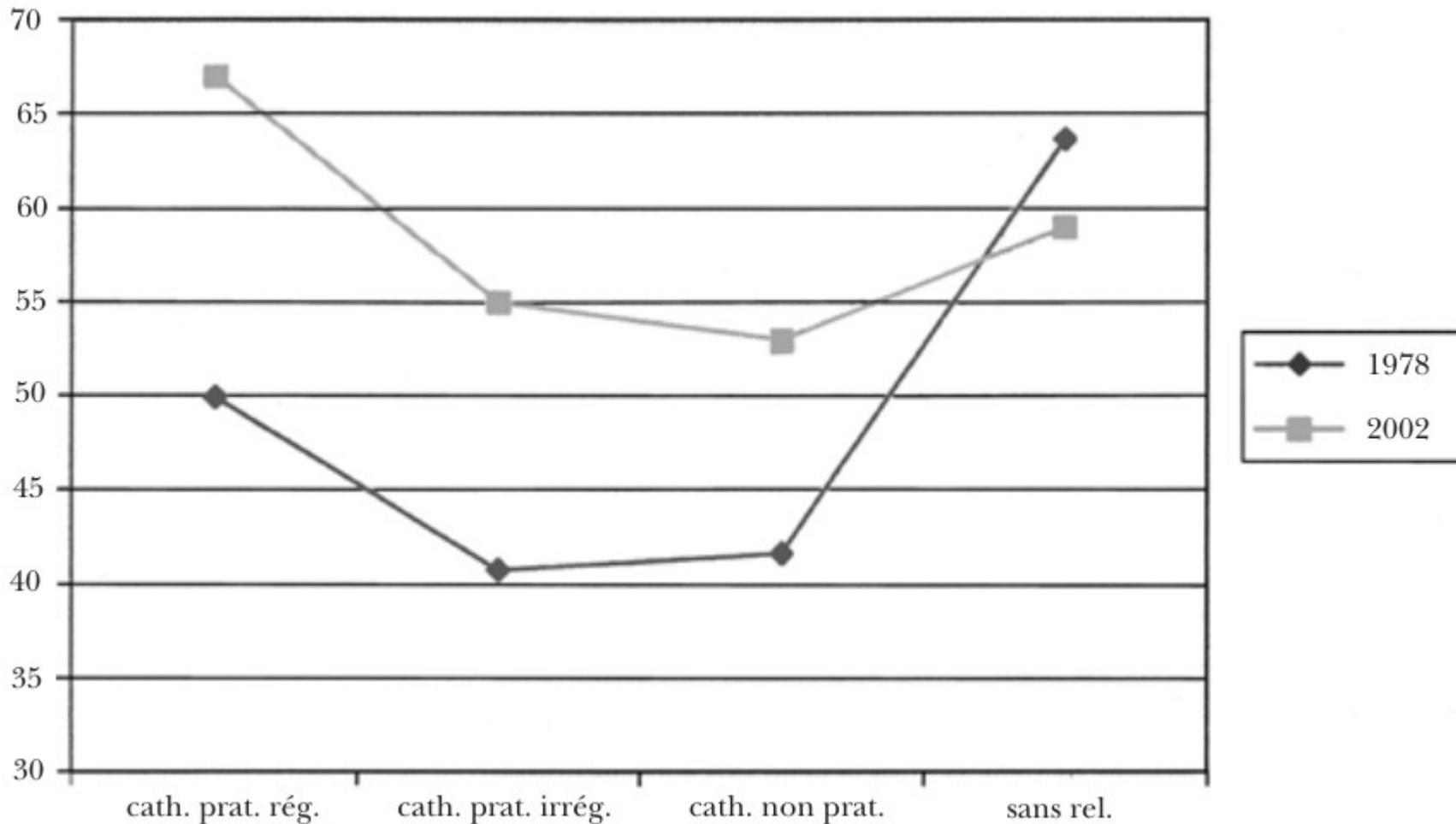


Tableau 4. Religion et vote au premier tour présidentiel 2002 (en %)

	Gauche	Droite	Exd	Effectifs
Catholique pratiquant régulier	21	66	13	(326)
Catholique pratiquant irrégulier	29	48	23	(520)
Catholique non pratiquant	42	36	22	(1 113)
Autre religion	50	35	15	(154)
Sans religion	63	20	17	(707)
Ensemble	43	38	19	(2 820)

Source : PEF 2002, vague 2 pondérée par les résultats du premier tour.

Tableau 5. Pratique religieuse des catholiques et vote au premier tour présidentiel 2002 (en %)

	Gauche	Droite	Exd	Effectifs
Plusieurs fois par semaine	12	79	9	(33)
Une fois par semaine	21	66	13	(170)
Une ou deux fois par mois	24	64	12	(123)
De temps en temps, aux grandes fêtes	29	48	23	(520)
Uniquement pour les cérémonies, mariages, etc.	41	37	22	(996)
Jamais	47	28	25	(116)
Ensemble	35	44	21	(1 958)

Source : PEF 2002, vague 2 pondérée par les résultats du premier tour de l'élection présidentielle.

Tableau 6.
Régression logistique
sur le vote de gauche

	%	Coefficient B	Exp (B)
<i>Genre</i>			
Homme	53	- 0,031	0,969
Femme	51	référence	
<i>Âge</i>			
18-24 ans	60	- 0,671**	0,511
25-39 ans	60	- 0,077	0,926
40-64 ans	51	référence	
65 ans et +	39	- 0,049	0,952
<i>Niveau d'études</i>			
Primaire	46	0,284*	1,329
BEPC, CAP, BEP	48	référence	
Baccalauréat	57	0,463***	1,588
Bac + 2	56	0,334**	1,397
2 ^e cycle et +	56	0,344**	1,411
<i>Patrimoine</i>			
0 ou 1	58	0,070	1,073
2	52	référence	
3	46	- 0,199	0,819
4 et +	39	- 0,440**	0,644
<i>Activité</i>			
Actif CDI	55	référence	
Actif précaire	56	0,210	1,234
Chômeur	57	0,091	1,095
Inactif	47	- 0,171	0,842
<i>Statut professionnel</i>			
À son compte	27	- 0,753***	0,471
Salarié du privé	49	référence	
Salarié du public	60	0,484***	1,623
<i>Religion</i>			
Cath. Prat. Rég.	26	- 1,054***	0,348
Cath. Prat. Irrég.	36	- 0,627***	0,534
Cath. Non prat.	52	référence	
Autre religion	59	0,124	1,132
Sans religion	72	0,788***	2,198
Constante		- 0,093	0,911

Source : PEF
2002, vague 2
non pondérée.

Régression logistique sur le vote de gauche. Quelques commentaires supplémentaires

-> L'auteur souligne bien évidemment le caractère significatif et très marqué de l'appartenance religieuse, qui joue dans le sens attendu : une pratique régulière (catholique) divise par trois le rapport entre vote de gauche et vote ailleurs, tandis que l'absence de religion le multiplie par deux.

-> D'autres variables jouent de manière significative et marquée : l'âge, le niveau d'étude, le patrimoine, le statut professionnel. Cependant, leur impact a tendance à être plus faible, en tout cas avec ce choix d'explicatives.

-> L'effet du patrimoine est celui attendu : les personnes les plus dotées votent moins à gauche, toutes choses égales par ailleurs. De même, les salariés du public votent plus à gauche et les indépendants moins à gauche que les salariés du privé. Moins marqué est l'effet du statut d'emploi : les précaires semblent un peu plus voter à gauche et les inactifs, moins, que les personnes en CDI.

-> En terme de niveau d'étude, les diplômés mais aussi les niveaux d'étude primaire ont plus voté à gauche toutes choses égales par ailleurs que la catégorie de référence.

-> Les jeunes ont significativement moins voté à gauche que la tranche d'âge de référence, ce qui est peut-être un point propre à cette élection mais aussi un artefact (non commenté par l'auteur) : en effet, un salarié précaire d'un niveau d'étude supérieur au bac sans patrimoine a un rapport de chance multiplié par 0,947 : il s'agit donc plutôt d'un résultat sur des jeunes inactifs et de faible niveau d'étude

Exemple étudié ici :

Joanie Cayouette-Remblière, 2015, « De l'hétérogénéité des classes populaires (et de ce que l'on peut en faire) », *Sociologie*, 6.4, <http://journals.openedition.org/sociologie/2652>

« S'il est possible de définir les classes populaires comme [...] dans une position sociale dominée et [...] à l'écart des normes sociales dominantes, [...] l'homogénéité des classes populaires est empiriquement infirmée et la catégorie ne peut plus être appréhendée qu'au pluriel. [...] cet article [...] propose d'approcher les classes sociales au moyen de configurations de positions sociales qui englobent d'un seul tenant une multitude de variables ainsi que leur modalité d'articulation. Ainsi sont construits un pôle cité [familles immigrées], un pôle cité [familles fragilisées], une « élite ouvrière », des « familles populaires entre-deux » et une « petite fonction publique », distincts des classes moyennes et supérieures. Ces configurations rendent intelligibles la diversité des modalités de socialisation familiale – et leurs contradictions – ainsi que la pluralité des trajectoires scolaires des élèves de classes populaires. »

L'enquête réalisée entre 2009 et 2012 a visé la reconstitution des trajectoires scolaires d'une population de 530 élèves, au moyen de leur dossier scolaire. Ces grandes chemises renferment de riches informations sur les élèves : outre les bulletins trimestriels de la sixième à l'éventuel accès au baccalauréat, ces dossiers rassemblent des données sur le primaire, des traces des processus d'orientation et la correspondance école/famille, entre autres, auxquelles s'ajoutent des fiches de renseignement socio-administratifs. Dans cet article, ces données seront utilisées pour ce qu'elles donnent à voir de la variété des milieux sociaux dans lesquels vivent les élèves. Il s'agit donc de partir de données renseignées par rapport à l'élève pour caractériser sa famille. Si ce tour de force est possible, c'est parce que les éléments caractérisant socialement la famille de l'élève sont nombreux et variés : on y trouve une fiche de renseignement remplie par les parents chaque année – soit pendant sept ans en moyenne – sur laquelle sont notés la profession des deux parents, la situation familiale, le statut de boursier ou non ainsi que l'adresse notamment. Puisque toutes ces déclarations impliquent des actes pratiques – la situation professionnelle des parents s'accompagne du numéro de téléphone au travail ; l'adresse implique la correspondance avec l'école... –, nous faisons l'hypothèse que les fausses déclarations y sont rares. Un autre intérêt de l'usage de données administratives est qu'il est possible d'avoir accès à une population exhaustive, contrairement aux enquêtes par questionnaires confrontées au problème de la non réponse, voire de l'attrition lorsqu'elles sont, comme ici, longitudinales. S'il présente ces avantages indéniables, le mode de recueil des données explique par ailleurs deux limites. La première a trait aux informations disponibles : dans ces dossiers, nulles traces du niveau de diplôme des parents ou encore des pays de naissance de ceux-ci, autant de déclarations que peut recueillir une enquête par questionnaire. La seconde renvoie à la délimitation de la population ainsi étudiée : ce ne sont pas toutes les classes populaires qui y sont représentées, mais les familles populaires ayant un enfant qui entre dans un collège public.

L'exploitation de ces matériaux a donné lieu à la construction d'une base de données nommée Cohortes, dont les modalités de construction à la fois statistiques et ethnographiques ont été détaillées par ailleurs (Cayouette-Remblière, 2011). Cette base de données a été complétée par une enquête ethnographique articulant entretiens, observations et travail sur archives. Seront ici principalement évoqués les entretiens avec les (anciens) élèves de la base Cohortes, réalisés lorsque ceux-ci avaient 20-21 ans.

Tableau 1. Qui se ressemble s'assemble. Profession du père et profession de la mère (en % ligne)

	Profession de la mère						
Profession du père	Enseignants (tx niveaux)	CPIS (h. enseignants)	Prof. Interm., commerçants ou artisans	Salariées d'exécution qualifiées	Sal. d'exéc. non qualifiées	Sans profession	<i>Total</i>
Enseignants (tous niveaux)	54	23	19	4			100 (n=26)
CPIS (hors enseignants)	14	24	30	14	10	8	100 (n=109)
Prof. Interm., commerçants ou artisans	7	9	31	20	22	11	100 (n=147)
Salariés d'exécution qualifiés	2	4	15	33	25	21	100 (n=112)
Salariés d'exéc. non qualifiés	3	1	13	16	43	24	100 (n=68)
Sans profession			5	26	26	42	100 (n=19)
Père absent	6	4	34	28	26	2	100 (n=47)
<i>Total</i>	9 (n=47)	10 (n=52)	24 (n=126)	21 (n=112)	22 (n=117)	14 (n=74)	100 (n=528)

Source : Base Cohortes / Significativité : Significatif au seuil de 1 sur 10 000.

Tableau 2. Stabilité du couple et profession de la mère

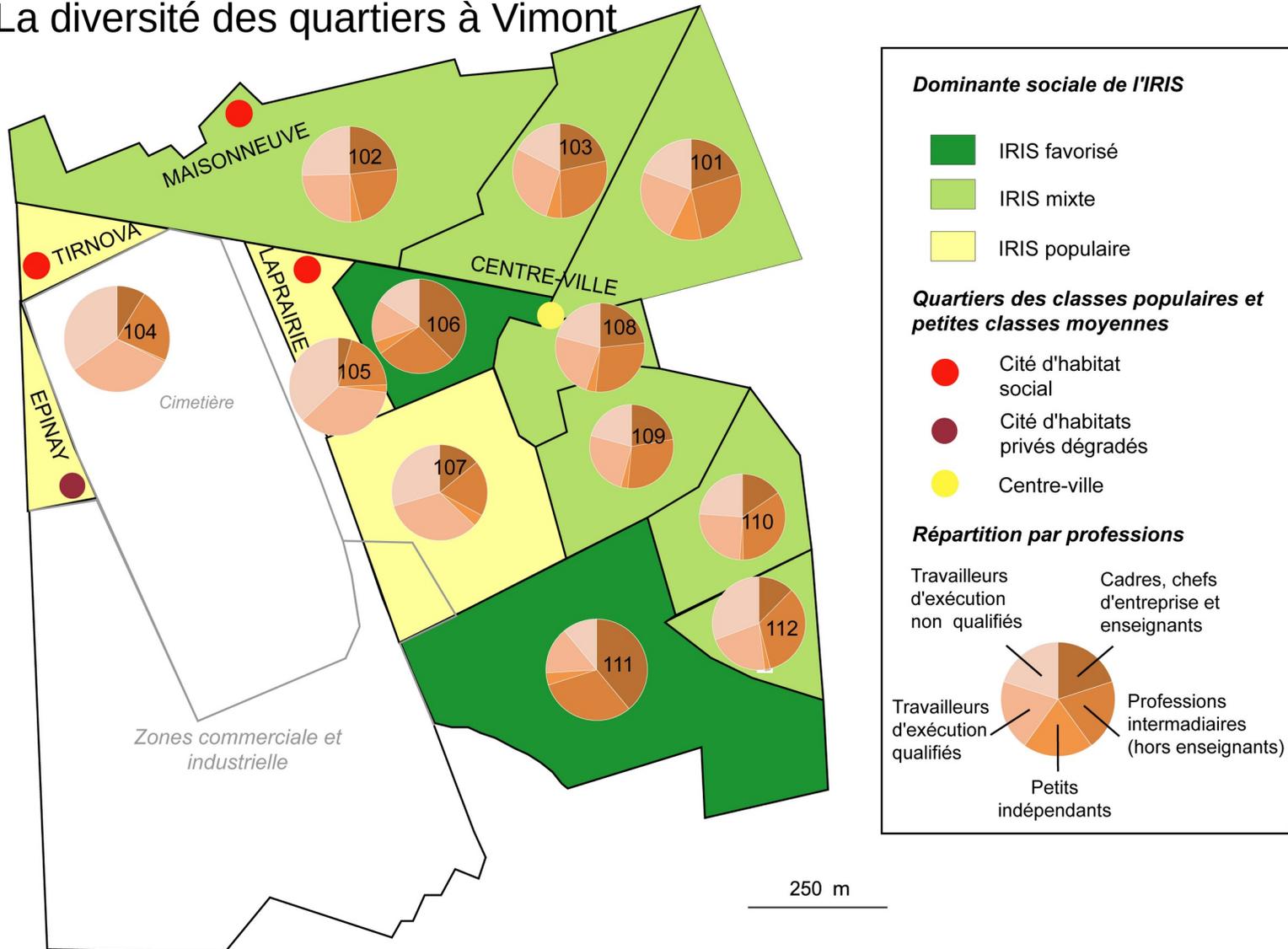
Professions de la mère	Part des mères n'étant plus en couple avec le père de l'enfant (%)
Enseignantes	17
CPIS (hors enseignantes)	24
Professions intermédiaires administratives, commerçantes ou techniciennes	19
Professions intermédiaires de la santé ou du travail social	29
Salariées d'exécution qualifiées	37
Salariées d'exécution non qualifiées	32
Sans emploi	18

Source : Base Cohortes

Tableau 3. Stabilité de l'emploi et profession du père

Profession du père	Probabilité de se déclarer chômeur à un moment de la trajectoire scolaire de l'élève (%)
Enseignants (tous niveaux)	0
Cadres et professions intellectuelles supérieures (hors enseignants)	0,9
Professions intermédiaires de santé ou travail social	0
Professions intermédiaires administratives	10,5
Techniciens et agents de maîtrise	4,2
Artisans et commerçants	9,6
Salariés d'exécution qualifiés	13,4
Salariés d'exécution non qualifiés	17,7
Ensemble	8,5

Figure 1 : La diversité des quartiers à Vimont



Source : recensement 1999, INSEE, données fournies par le réseau Quételet.

Figure 2 : Street view révèle...

(types d'habitation utilisés)



Une maison individuelle



Une habitation collective « basse »



Une habitation collective « haute »



Une tour d'habitation

Tableau 4 : Types d'habitation selon la profession du père (en % ligne)

Professions du père	Types d'habitation			
	Tour d'habitation	Immeuble collectif haut	Immeuble collectif bas	Maison individuelle
Enseignants	4	33	21	42
CPIS (hors enseignants)	7	16	19	58
Professions intermédiaires, artisans et commerçants	13	28	26	34
Salariés d'exécution qualifiés	22	30	27	22
Salariés d'exécution non qualifiés	34	31	28	8
Sans profession	47	42	5	5
Total	19	27	24	30

La situation à l'égard de l'emploi atteste une fois encore de la persistance de la domination, d'une part, et des processus de segmentation des classes populaires, d'autre part. Sur ce point, nos données présentent l'originalité de prendre en compte non seulement la profession des parents mais également son évolution tout au long de la trajectoire scolaire de l'enfant. Par rapport aux enquêtes nationales, cette information autorise un point de vue différent sur les risques de chômage en fonction de la catégorie socioprofessionnelle. Alors que l'INSEE présente le plus souvent une photographie de la situation à un temps t en faisant le lien entre la dernière profession déclarée et la situation professionnelle actuelle, les dossiers scolaires permettent, sur une durée de sept années en moyenne, de repérer les périodes de chômage ou d'inactivité déclarées. On observe, sur nos données, que près de 9 % des pères alternent entre périodes d'emploi et de non emploi. Or, cette proportion est nulle ou quasi nulle pour les enseignants et cadres alors qu'elle frôle les 20 % pour les salariés d'exécution non qualifiés (tableau 3). Ainsi, nos données démontrent de nouveau l'intérêt de complexifier l'étude de la question sociale en prenant en compte à la fois la profession et le risque qu'elle induit de connaître une période de chômage. [...]

Les éléments jusqu'ici présentés sont également, au moins en partie, observés sur des données nationales. L'intérêt de les exposer ici est d'établir qu'ils s'appliquent, d'une part, à une population circonscrite – celle des familles dont un enfant rentre au collège public – saisie de façon exhaustive, et d'autre part, à l'échelle locale. La démonstration des inégalités sociales n'en est que plus forte lorsque l'on constate que des tendances générales s'observent tout autant sur un « bout de territoire ». Cependant, là où l'analyse localisée s'avère essentielle, c'est pour rendre compte de différences géographiques.

Classes populaires de cité, de centre-ville et de pavillons

En plus de se différencier selon leurs situations familiale, professionnelle et économique ainsi que leur histoire migratoire, les familles des élèves se distinguent également entre elles par leur habitation. Le lieu de résidence prend ici tout son sens puisque l'on sait que l'attachement au quartier est une caractéristique ancienne du style de vie des classes populaires, surtout des jeunes (Hoggart, 1970 [1957]), et que la sociabilité locale est sans doute « une des spécificités culturelles les plus incontestables » (Schwartz, 1998, p. 80) de ce groupe social, le quartier occupant une place centrale dans la structuration des univers de vie.

Présentation de l'ACM

Grâce à l'ensemble des variables présentes dans les dossiers scolaires, j'ai d'abord construit un espace des positions sociales (Figures 3 et 4) grâce à une analyse des correspondances multiples (ACM) spécifique. Cet espace permet de rendre compte des articulations les plus récurrentes en rapprochant les variables les plus fréquemment réunies pour un même élève, en éloignant celles dont l'articulation est plus rare (Duval, 2010).

Trois axes rendent compte des principales oppositions entre les 42 modalités dichotomiques traitées en variables actives. L'axe 1 est un axe de volume du capital global, qui prend en compte de nombreuses variables. De l'ouest vers l'est, on lit les modalités des moins favorisés aux plus favorisés. La situation d'emploi du père – instable ou sans emploi – contribue fortement à construire l'axe du côté défavorisé. L'absence d'emploi de la mère y participe également. Cette opposition se cumule avec une différence selon le type d'habitation et de quartier : d'un côté (à l'ouest), le fait de résider en cité et en tour d'habitation, de l'autre (à l'est), le fait d'habiter une maison et de vivre à l'extérieur de Vimont. L'opposition boursier/non boursier (sur critères sociaux) est également forte sur cet axe. S'y ajoute une opposition entre, à l'ouest, les fratries de quatre enfants et plus, les prénoms maghrébins et les mères salariées d'exécution non qualifiées ; à l'est, les parents cadres, travaillant dans le secteur public et les enfants portant un prénom « autochtone ».

Variables professionnelles

Profession du père à l'entrée en 6e en sept modalités : 1/ pères cadres, de profession intellectuelle supérieure ou enseignants (26 %) ; 2/ pères de profession intermédiaire administrative, de la santé ou du travail social ou commerçants (10 %) ; 3/ pères de profession intermédiaire manuelle (techniciens, contremaîtres, agents de maîtrise, artisans) (18 %) ; 4/ pères salariés d'exécution qualifiés (22 %) ; 5/ pères salariés d'exécution non qualifiés (12 %) ; 6/ pères sans emploi* 2 (4 %) ; 7/ pères absents* (8 %).

Profession de la mère à l'entrée en 6e en cinq modalités : 1/ mères cadres, de profession intellectuelle supérieure ou enseignantes (18 %) ; 2/ mères de profession intermédiaire (hors professeures des écoles) ou commerçantes (24 %) ; 3/ mères salariées d'exécution qualifiées (21 %) ; 4/ mères salariées d'exécution non qualifiées (22 %) ; 5/ mères sans emploi* (15 %) ou absente* (quatre individus).

Secteur d'emploi du père en trois modalités : 1/ pères dans le secteur public (18 %) ; 2/ pères dans le secteur privé (70 %) ; 3/ pères sans profession ou absents* (12 %).

Secteur d'emploi de la mère en trois modalités : 1/ mères dans le secteur public (36 %) ; 2/ mères dans le secteur privé (49 %) ; 3/ mères sans profession* ou mères absentes* (15 %).

Situation professionnelle du père en quatre modalités : 1/ pères sans emploi à chaque déclaration durant toute la scolarité secondaire de l'élève (6 %) ; 2/ pères en emploi à chaque déclaration durant toute la scolarité secondaire de l'élève (70 %) ; 3/ pères déclarant parfois des périodes d'emploi, parfois des périodes de chômage ou de non emploi (8 %) ; 4/ inconnue* ou père absent* (15 %).

Situation professionnelle de la mère en trois modalités : 1/ mères sans emploi à chaque déclaration durant toute la scolarité secondaire de l'élève (16 %) ; 2/ mères en emploi à chaque déclaration durant toute la scolarité secondaire de l'élève (65 %) ; 3/ mères déclarant parfois des périodes d'emploi, parfois des périodes de chômage ou de non emploi (15 %) ; 4/ inconnue* ou mère absente*(3 %).

Variables familiales

Situation familiale en deux modalités : 1/ parents déclarés en couple tout au long de la scolarité secondaire de l'élève (73 %) ; 2/ parents déclarés séparés, divorcés ou seuls à au moins un moment durant la scolarité secondaire de l'élève (27 %).

Nombre d'enfants dans la fratrie en cinq modalités : 1/ un enfant (17 %) ; 2/ deux enfants (43 %) ; 3/ trois enfants (24 %) ; 4/ quatre enfants et plus (13 %) ; 5/ inconnu* (3 %).

Variable économique

Statut de boursier en trois modalités : 1/ n'a jamais été boursier pendant sa scolarité secondaire (67 %) ; 2/ a été boursier au moins une année pendant sa scolarité secondaire (27 %) ; 3/ inconnu* (6 %).

Variable d'origine migratoire

Origine culturelle du prénom en six modalités : 1/ prénom français (65 %) ; 2/ autre prénom européen (8 %) ; 3/ prénom maghrébin (15 %) ; 4/ prénom d'Afrique subsaharienne (4 %) ; 5/ prénom asiatique (4 %) ; 6/ prénom mixte (3 %).

Variables de lieu d'habitation

Types de quartier en cinq modalités : 1/ cités ou quartiers privés défavorisés (25 %) ; 2/ centre-ville (11 %) ; 3/ quartier Nord (24 %) ; 4/ quartier Sud (22 %) ; 5/ hors Vimont (18 %).

Types d'habitation en quatre modalités : 1/ maison individuelle ou maison de ville (32 %) ; 2/ immeuble collectif de quatre étages et moins (24 %) ; 3/ immeuble collectif de cinq à huit étages (25 %) ; 4/ tours d'habitation (18 %).

1 Cette ACM spécifique a été réalisée sur 467 individus (soit 88 %) de la base Cohortes. Les individus pour lesquels les données sont lacunaires sont exclus de la construction de l'ACM.

Figure 4 : L'espace des positions sociales (axes 1 et 3)

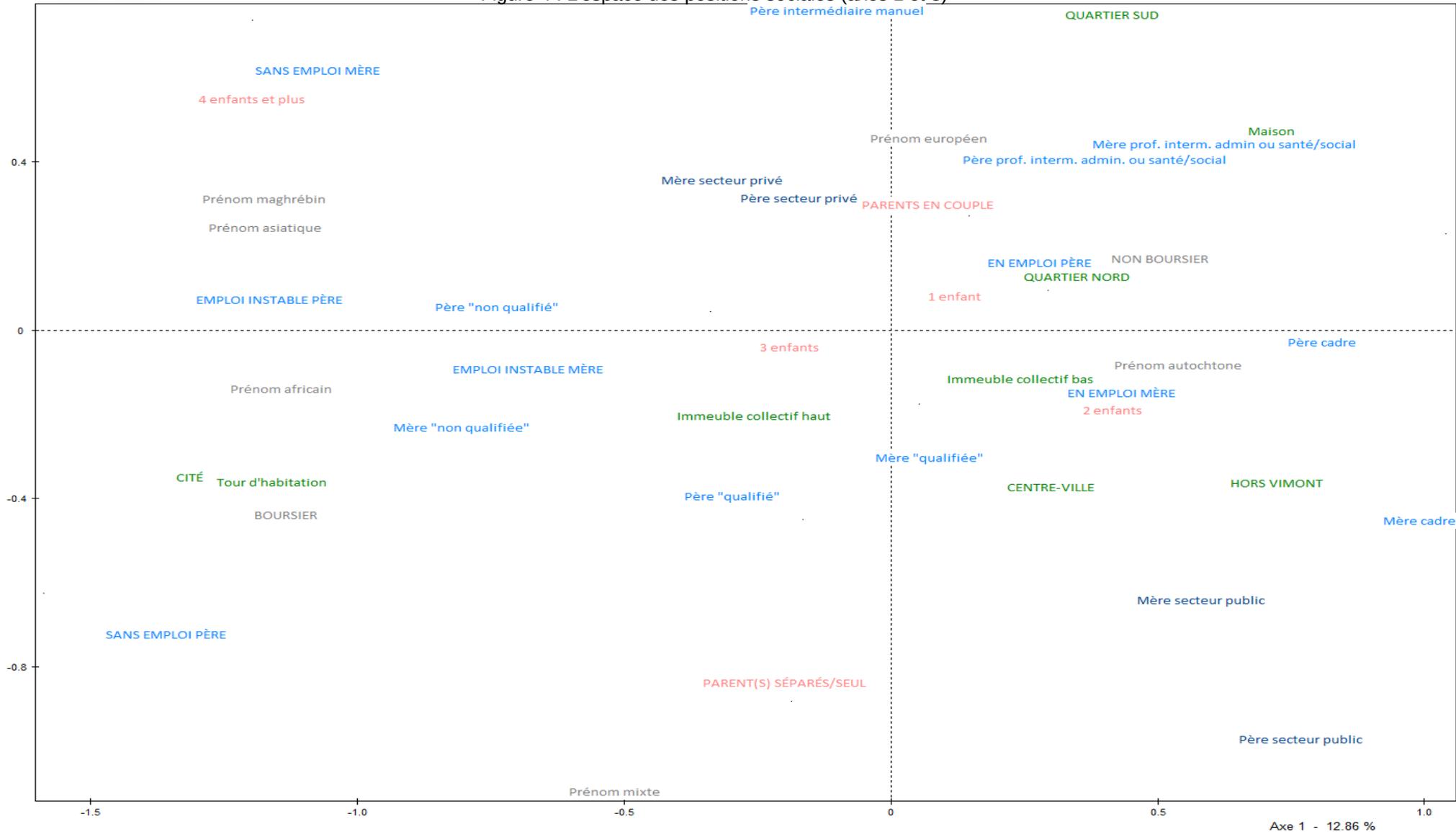
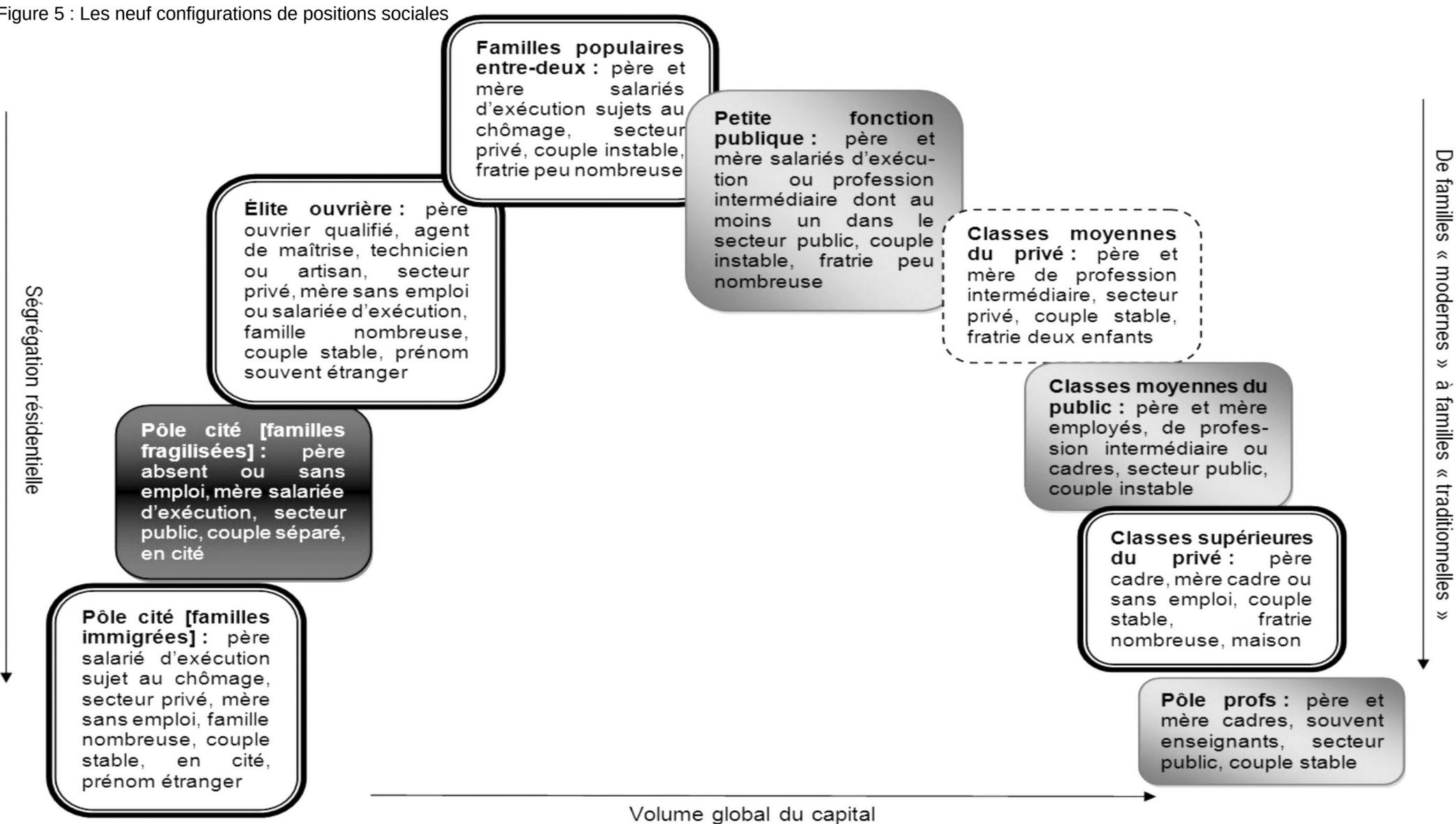


Figure 5 : Les neuf configurations de positions sociales



Annexe 2 : Répartition de la population en neuf configurations de positions sociales

Configurations de positions sociales	Fréquence	%
Pôle profs	40	7,6
Classes supérieures du privé	61	11,5
Classes moyennes du public	48	9,1
Classes moyennes du privé	74	14,0
Petite fonction publique	49	9,3
Familles populaires entre-deux	88	16,6
Élite ouvrière	58	10,9
Pôle cité [familles fragilisées]	64	12,1
Pôle cité [familles immigrées]	48	9,1

Principe général

- Étude systématisée de variables corrélées entre elles / séparation d'effets
- Variables d'intérêt en *actives*
- Variables usuelles en *supplémentaires*...
- ... ainsi que les modalités rares, susceptibles d'apparaître comme les plus discriminantes
- Utilisation combinée avec des méthodes de classification → très courante (construction de typologies)

Spécificités de cette ACM

- Revisiter ici un objet classique (les appartenances sociales)
- Éclairage ethnographique
- Données administratives