

# Sur quelques conséquences des différents emplois du code des catégories socioprofessionnelles

par Jean-Pierre Briand \*

---

L'un des dangers les plus pernicious dans l'usage des données statistiques est peut-être que les mêmes intitulés peuvent ne pas recouvrir les mêmes réalités dans des sources différentes. Des exemples tirés de données par catégorie socioprofessionnelle sont à cet égard démonstratifs. En rapprochant des données sur les décès, le divorce, la scolarisation de celles provenant du recensement, on peut arriver à des taux par catégorie sociale visiblement anormaux.

De telles anomalies se comprennent lorsqu'on prend en compte les processus d'élaboration des données. Ainsi, les codeurs des administrations ne disposent

souvent que d'intitulés vagues sur les professions. Ils ont alors tendance, par prudence, à les classer dans les catégories moyennes, et notamment dans les « employés ».

Il est certes généralement admis que la signification des données statistiques dépend des conditions particulières de leur production. Mais pour que cet agrément ne reste pas de pure forme, il convient de développer les analyses concrètes de ces conditions. Ces analyses permettent non seulement d'établir l'interprétation des statistiques sur des bases plus solides, mais aussi d'éclairer l'ensemble du fonctionnement des institutions.

---

En raison du rôle central que joue l'INSEE dans la production de statistiques en France, le code des catégories socioprofessionnelles de l'INSEE s'est progressivement imposé, sous une forme plus ou moins simplifiée, aux divers organismes qui participent à cette production. Dès lors, la tentation est grande de rapprocher des données recueillies par différents organismes dans le cadre de la nomenclature commune. Cherche-t-on, par exemple, à déterminer les « taux d'accès » des enfants à un niveau scolaire selon la catégorie socioprofessionnelle de leur père, ou les « taux de divortialité » selon la catégorie socioprofessionnelle des

conjointes? Les statistiques scolaires ou judiciaires peuvent fournir les numérateurs de ces taux, et les recensements les dénominateurs.

---

\* Jean-Pierre Briand est maître-assistant de sociologie à l'Université de Paris VIII et membre du Centre de recherches socio-historiques sur l'éducation de cette Université.

Les nombres entre crochets, [ ], renvoient à la bibliographie en fin d'article.

Mais l'identité nominale de la nomenclature utilisée ne suffit évidemment pas à garantir l'identité des usages effectifs de celle-ci. Les déclarations d'une même situation peuvent en effet être différentes selon les sources; elles sont souvent de précision inégale, ce qui rend impossible l'application des mêmes principes de codage; enfin, même quand les déclarations sont identiques, les interprétations des codes peuvent diverger. Tout ceci peut entraîner des variations systématiques, dans les classements opérés d'une source à l'autre, qui sont du même ordre de grandeur que les différences entre catégories sociales que l'on croit mesurer au moyen des rapprochements effectués. En raison de ce problème, que les démographes de l'INSEE déclaraient « bien connu » dès 1957 [17], ceux-ci se sont abstenus de calculer des taux en rapprochant des répartitions par catégorie socioprofessionnelle issues de deux sources différentes. Par contre, certains sociologues ont élaboré, à partir de tels rapprochements, des théories sur les relations entre divers phénomènes sociaux, comme la scolarisation ou le divorce, et la structure de classe.

Je me propose d'abord de montrer ici que ces rapprochements conduisent aux mêmes anomalies que dans le cas de la mortalité précédemment étudié par l'INSEE [17], et d'analyser un des principaux mécanismes qui produisent celles-ci. Il convient toutefois de dépasser le point de vue strictement statistique qu'exprime le terme « d'anomalie », ou celui de « biais », afin de reconstituer les processus sociaux d'élaboration des données statistiques, dont ces « anomalies » sont en fait les conséquences normales<sup>1</sup>.

### Trois exemples

Pour illustrer la variété des formes sous lesquelles se présente le problème des rapprochements, deux types différents d'exemples ont été retenus<sup>2</sup>. Dans le cas de la scolarisation, les renseignements sur la profession sont recueillis et classés dans la liste des catégories socioprofessionnelles de l'INSEE par les établissements scolaires eux-mêmes, alors que, pour les décès et les divorces, les informations, recueillies par des administrations distinctes de l'INSEE, sont codées par cet organisme. (On verra que ces exemples s'opposent différemment sous un autre rapport.)

Pour comprendre le principe des divergences entre sources qui seront analysées, il convient de rappeler sommairement comment la catégorie socioprofessionnelle est déterminée dans les recensements, qui fournissent les données placées au dénominateur des taux calculés par rapprochement (les sources utilisées au numérateur seront décrites cas par cas). Les bulletins de recensement sont, en principe, remplis par les personnes concernées ou par une autre personne du ménage, sans intervention d'un enquêteur. Le répondant doit indiquer une série d'informations : profession individuelle; statut (salarié, indépendant, employeur, etc.); activité de l'établissement; nombre de salariés (pour les employeurs), qualification; grade, position

hiérarchique (pour les salariés). Les chiffreuses doivent ensuite déterminer la catégorie socioprofessionnelle à l'aide de répertoires de métiers et de codes de correspondance entre ceux-ci, combinés avec les divers autres critères énumérés, et les catégories socioprofessionnelles. Ou bien, selon une méthode plus récente, elles doivent déterminer la profession, ainsi que le statut, etc, et chaque intitulé devra être accepté par un programme informatique et classé par celui-ci en fonction des mêmes codes de correspondance. Des principes sont établis à l'avance pour classer certains types de déclarations très vagues (comme « fonctionnaire »). Les chiffreuses peuvent également, même si les consignes officielles ne le précisent pas, recourir à d'autres rubriques du questionnaire, par exemple à celle sur l'instruction, pour déterminer la catégorie socioprofessionnelle ou la profession. Le personnel affecté au codage est formé spécialement à cette tâche.

### L'exemple de la mortalité

À l'occasion des déclarations de naissance, de décès, ou des mariages, les mairies recueillent des renseignements d'état civil fournis, selon les cas, par les personnes concernées elles-mêmes, par des membres de leur famille, ou parfois par des agents des pompes funèbres ou des maternités. Des agents de la mairie sont chargés d'inscrire ces renseignements sur des bulletins statistiques spéciaux. En ce qui concerne la profession des décédés, ils doivent noter la profession proprement dite, l'activité de l'établissement où travaillait l'individu, et, pour les années cinquante sur lesquelles portent les données citées ci-dessous, la situation dans la profession : le bulletin indiquait en note les différents statuts visés par cette expression (par la suite, cette rubrique se réduira à la question : « Était-il salarié ? »). Il semble que les renseignements sur la profession soient souvent transcrits de manière assez simplifiée par les préposés.

Les données recueillies sont ensuite codées par l'INSEE. Les instruments utilisés pour déterminer la catégorie socioprofessionnelle ont d'abord été les mêmes codes des métiers et des catégories socioprofessionnelles que pour les recensements. Depuis 1975, des tableaux de correspondance spécifiques ont été dérivés des précédents. Ils indiquent aux chiffreuses des choix simplifiés, adaptés au caractère sommaire des renseignements recueillis. Ainsi, pour

---

1. Cette étude s'inscrit dans une démarche plus générale d'analyse des modes de production et d'utilisation de divers types de statistiques sociales, poursuivie depuis 1974 avec J. M. Chapoulie et H. Peretz [6, 7, 8].

2. On trouvera une analyse plus développée de ces trois exemples, et de quelques autres, ainsi que l'indication plus précise des sources et des modes de calcul dans [4] et [5].

TABLEAU 1

**Mortalité des hommes de 45 à 54 ans selon la catégorie socioprofessionnelle \***

En ‰

Taux de mortalité, année 1955		Quotients de mortalité, années 1956-1960		
		à partir de l'échantillon suivi par l'INSEE	en rapprochant les données de l'état civil et du recensement	
0 Agriculteurs.....	75	77	73	0 Agriculteurs.
1 Salariés agricoles.....	99	104	93	1 Salariés agricoles.
2 Patrons de l'industrie et du commerce.....	91	81	88	2 Patrons de l'industrie et du commerce.
3 Professions libérales et cadres supérieurs....	57	51	54	3 Professions libérales et cadres supérieurs, sauf officiers.
4 Cadres moyens.....	57			
		52	56	42 Instituteurs de l'enseignement public.
		61	57	43 Techniciens du privé.
		64	57	44-45 Cadres moyens.
5 Employés.....	97			
		79	104	51-52 Employés de bureau.
6 Ouvriers.....	101			
		84	88	60-61-62 Contremaîtres et ouvriers qualifiés.
		95	91	63-64 Ouvriers spécialisés.
dont 68 Manceuvres.....	130	131	108	68 Manceuvres.

\* Dans la première et la troisième colonnes, au numérateur, la catégorie socioprofessionnelle des personnes décédées est déterminée à partir des informations recueillies par les mairies; au dénominateur, la catégorie socioprofessionnelle est déterminée à partir des données recueillies aux recensements. Dans la deuxième colonne, au numérateur et au dénominateur, la catégorie socioprofessionnelle est déterminée à partir des données du recensement.

Les numéros des catégories socioprofessionnelles sont ceux des codes utilisés pour le recensement de 1954.

La première colonne est tirée d'un tableau publié dans [17], p. 41, la seconde et la troisième d'un tableau publié dans [23], p. 940. Pour plus de précisions sur le découpage des catégories utilisé à droite du tableau, voir [23] et [12], p. 10.

certaines intitulés de profession qui, dans les recensements, sont classés de manière variable selon les renseignements supplémentaires disponibles, ces documents fixent le classement en fonction des plus grandes fréquences observées dans les recensements. Ils prescrivent aussi, dans certains cas d'indétermination, des procédés conventionnels, comme par exemple de classer un intitulé alternativement dans chacune des deux catégories pouvant convenir.

Parmi les divers renseignements de l'état civil, ceux qui concernent les décès sont les plus pertinents par rapport à l'objet de cet article. La mort en elle-même constitue en effet un fait simple, dont la déclaration à l'état civil pose peu de problème de validité et d'interprétation. De plus, on dispose de deux types de mesure de la mortalité par catégorie socioprofessionnelle directement comparables.

D'une part, les études de M. Febvay et M. Aubenque [17] et de C. Lévy et S. Vallin [23] rapprochent la répartition des

décédés selon la catégorie socioprofessionnelle attribuée par l'INSEE d'après la profession déclarée à l'état civil, de la répartition socioprofessionnelle de la population du même âge dans les recensements. Leurs résultats peuvent être comparés, d'autre part, à ceux de l'observation des décès dans l'échantillon extrait par l'INSEE du recensement de 1954; les décès déclarés à l'état civil sont rapportés ici à la personne concernée dans l'échantillon, et l'analyse n'utilise que la catégorie socioprofessionnelle attribuée dans le recensement (tableau 1).

L'étude de M. Febvay et M. Aubenque a été effectuée avant la mise en place de l'échantillon extrait du recensement. Les auteurs commentent le caractère surprenant des taux de mortalité des cadres moyens et des employés, respectivement très faible et très élevé par rapport à ceux des autres catégories socioprofessionnelles. Ils mettent ce fait en relation avec la moins grande précision de la déclaration de la

# TRAVAIL ET EMPLOI

Revue trimestrielle éditée par  
le Ministère des Affaires Sociales  
et de la Solidarité Nationale  
Service des Études et de la Statistique

## AU SOMMAIRE DU NUMÉRO 20

### LA NÉGOCIATION COLLECTIVE

- La pratique de la négociation collective.
- La négociation collective de branche et d'entreprise dans la crise, modifications des enjeux, des stratégies et des pratiques.
- Aperçu sur la négociation salariale en 1983.
- Évolution comparée du SMIC et des rémunérations minimales garanties.

#### Dans ce même numéro

- Les cercles de qualité : la portée d'une implantation réussie.
- Contribution à une histoire des durées du travail dans l'industrie française du milieu du XIX<sup>e</sup> siècle à la seconde guerre mondiale.

#### CONSULTATIONS – VENTES :

LA DOCUMENTATION FRANÇAISE :  
29, quai Voltaire, 75007 PARIS;  
165, rue de Garibaldi, 69401 LYON.

Par correspondance :

124, rue Henri Barbusse 93308 AUBERVILLIERS.

MINISTÈRE DES AFFAIRES SOCIALES  
ET DE LA SOLIDARITÉ NATIONALE :  
1, place de Fontenoy 75700 PARIS

Abonnement : 110 F. Numéro : 30 F.

Abonnement jumelé avec BMST et son supplément : 320 F

profession à l'état civil qu'au recensement, et, par voie de conséquence, avec les modes d'application différents de la nomenclature à ces déclarations; en particulier, remarquent-ils, des déclarations vagues enregistrées par l'état civil, « fonctionnaire », « cheminot », « employé de l'EDF », risquent de faire classer les individus concernés parmi les employés, alors que des déclarations plus précises avaient pu faire classer ces mêmes individus, au recensement, parmi les « cadres moyens » ou les « ouvriers ».

C. Lévy et S. Vallin confirment l'existence d'anomalies semblables (ils disposaient, pour comparaison, des résultats de l'échantillon suivi par l'INSEE). Les auteurs admettent que le caractère flou de la catégorie « employés » provoque des transferts de sens contraire. Mais ils proposent comme principale explication des anomalies constatées une attitude prêtée aux déclarants : ceux-ci tendraient à « anoblir » la profession du défunt, ce qui entraînerait de proche en proche une sous-estimation des catégories inférieures au profit des catégories supérieures, surestimées. Toutefois cette explication ne s'accorde que très partiellement avec les données présentées, et ne rend pas compte d'un des écarts les plus nets, celui précisément qui concerne les « employés », sauf à supposer, ce qu'aucune des données présentées ne justifie, que la translation s'opérerait des « ouvriers » aux « employés », et des « cadres moyens » aux « cadres supérieurs », mais non des « employés » aux « cadres moyens ». Ce mode d'explication par les *attitudes* prêtées aux déclarants est très certainement moins pertinent ici que le mode d'explication par les propriétés des *situations* de déclaration et de codage auquel recourent M. Febvay et M. Aubenque.

### L'exemple du divorce

De 1970 à 1978, le ministère de la Justice a organisé une enquête permanente et exhaustive sur les tentatives de conciliation en divorce ou en séparation de corps engagées auprès des tribunaux. Les renseignements étaient recueillis par le personnel administratif des greffes, et leur exploitation a été assurée en commun par le ministère de la Justice et l'INSEE, qui a réalisé le chiffrage. La profession était demandée dans les mêmes termes que sur les bulletins statistiques de l'état civil à la même époque. Tout le processus était donc, dans ses grandes lignes, analogue au mode d'élaboration des données de l'état civil. Quelques particularités sont toutefois à signaler. Les greffiers recueillaient les renseignements auprès du demandeur uniquement, c'est-à-dire, dans la majorité des cas, auprès de l'épouse. Le taux de non-réponse est de ce fait assez élevé pour les hommes en ce qui concerne la profession. De plus, la réalisation de cette enquête a constitué pour le personnel un surcroît de travail qui paraît avoir été parfois mal accepté. Certains aspects de l'enquête, comme la question sur les professions, ne semblent pas non plus avoir reçu une large adhésion parmi les magistrats. Aussi certains tribunaux n'ont-ils pas communiqué les statistiques demandées, dont un au moins avait pour ressort une zone d'importante concentration ouvrière.

Des chercheurs du ministère de la Justice ont calculé des taux de divortialité en rapprochant la répartition par catégorie socioprofessionnelle des personnes engagées dans une tentative de conciliation une année, de celle des personnes qui se déclarent mariées dans les recensements. Ces taux peuvent être comparés à deux autres mesures de la divortialité, qui mettent chacune en œuvre, pour un individu donné, une seule et même catégorie socioprofessionnelle. Il s'agit des données sur l'état matrimonial selon la catégorie socioprofessionnelle des individus, recueillies dans les recensements, et des données sur les dissolutions par divorce des premiers mariages selon la catégorie socioprofessionnelle du mari, d'après les enquêtes « Familles » réalisées auprès d'un échantillon de femmes en même temps que les recensements de 1962 et de 1975 (tableau 2).

En raison des caractéristiques très différentes du mode de constitution des trois types de données utilisés ici, ainsi que de l'inexistence de repères que constitueraient quelques propriétés, grossières mais bien établies, de la divortialité selon les catégories sociales, l'interprétation de ces données doit être très prudente (voir notes du tableau 2).

Néanmoins, en comparant les taux des « employés » à ceux de l'ensemble ou des catégories les plus proches socialement (« ouvriers » et « cadres moyens »), on remarque qu'aucune enquête ne confirme l'ampleur de la sur-divortialité des « employés » des deux sexes que font apparaître les taux du ministère de la Justice, même si les « employés » constituent toujours une des catégories qui, après les « personnels de service » et les femmes des « autres catégories », se caractérisent par les chiffres les plus élevés. De même les taux obtenus par rapprochement sont les seuls à indiquer, dans l'ensemble, une divortialité plus élevée pour les « cadres moyens » que pour les « ouvriers » : les autres sources indiquent le plus souvent l'inverse.

Une deuxième conclusion peut être tirée de l'examen de toutes ces données, si on les utilise d'un autre point de vue, pour tester la proposition, avancée par A. Boigeol et J. Commaille [3], ainsi que par L. Roussel ([30], p. 148 et 152), d'une propension au divorce typique des « classes moyennes », représentées par les « cadres moyens » et les « employés ». D'après les calculs du ministère de la Justice, les taux de divorce des « cadres moyens » ne sont un peu plus élevés que ceux des « cadres supérieurs et professions libérales » et des « ouvriers » que pour les hommes, et le sentiment d'une sur-divortialité nette des « classes moyennes » n'est en réalité produit que par la considération du taux des « employés », très vraisemblablement surestimé. Dans les recensements, aucune différence nette n'oppose systématiquement les taux des « employés » et des « cadres moyens » à ceux des « ouvriers » ou des « cadres supérieurs et professions libérales ». Dans les enquêtes « Familles » enfin, les probabilités de divorces sont plus élevées pour les femmes d'« employés » que pour les femmes d'« ouvriers » ou de « cadres supérieurs et professions libérales », mais ce fait renvoie, au moins pour partie, au mode hétérogène de saisie de la profession, et les fréquences de divorce sont voisines pour les épouses de « cadres moyens » et celles d'« ouvriers ». Par ailleurs, lorsque les données sont ventilées par sous-catégorie d'« ouvriers », et qu'il s'agisse de la catégorie socioprofessionnelle individuelle ou de celle du mari, les chiffres de

## Une série INSEE présentant les résultats de travaux ou d'études spécialisés

# ARCHIVES & DOCUMENTS

Pour permettre à des utilisateurs spécialisés de disposer aisément et rapidement de résultats d'études qui n'entraient pas dans le champ général d'édition des « Collections de l'INSEE », l'INSEE a créé en 1980, la série « Archives et Documents ». Elle regroupe, dans des ouvrages à tirage limité, de présentation modeste, des informations chiffrées allant à un plus grand niveau de détail que celles publiées dans les « Collections », des études particulières réservées à un public de spécialistes, des résultats de travaux spécifiques...

### Derniers numéros parus :

Liste par secteur des principales concentrations réalisées dans l'industrie française : 2<sup>e</sup> semestre 1983 (n° 98, 98 pages, 30 F)

Tableaux normalisés par région RP 1982 - France métropolitaine : sondage au 1/20<sup>e</sup> (n° 99, 310 pages, 52 F)

Tableaux normalisés simplifiés par région et département RP 1982 - France métropolitaine : sondage au 1/20<sup>e</sup> (n° 100, 274 pages, 48 F)

Les enquêtes de conjoncture (n° 101, 60 pages, 23 F)

Comptes économiques de la Guadeloupe séries 1978-1981 - Système élargi de comptabilité nationale (n° 102, 130 pages, 39 F)

Les comptes nationaux trimestriels - Séries longues 1963-1983 (n° 103, 244 pages, 90 F)

RP 82 - Principaux résultats par catégorie de commune - France métropolitaine : sondage au 1/20<sup>e</sup> (n° 104, 246 pages, 47 F)

Les industries agricoles et alimentaires en 1983 - Séries statistiques 1976-1983 (n° 105, 112 pages, 34 F)

A la différence des Collections de l'INSEE, il n'est pas prévu d'abonnement à cette série.

CONSULTATION, VENTE

P 608

Dans les observatoires économiques régionaux de l'INSEE et chez les libraires spécialisés.



Institut National de la Statistique et des Études Économiques

TABLEAU 2

**Indices de divortialité selon la catégorie socioprofessionnelle \***

2 A. Taux annuels de divorce selon la catégorie socio-professionnelle individuelle (rapprochements)

En ‰

	Hommes			Femmes		
	1970	1975	1976-1977	1970	1975	1976-1977
Agriculteurs.....	0,9	1,2	1,2	0,5	0,5	0,5
Salariés agricoles....	3,1	5,3	5,6	6,5	5,6	7,0
Patrons du commerce et de l'industrie...	5,0	5,2	4,8	4,7	4,1	3,6
Professions libérales, cadres supérieurs..	7,2	7,1	8,4	12,5	8,2	11,3
Cadres moyens.....	8,7	8,9	9,7	11,0	7,7	10,0
Employés.....	11,0	11,9	14,0	18,9	12,6	15,6
Ouvriers.....	6,2	7,2	7,6	9,6	7,9	8,6
Personnel de service.	11,4	13,9	14,7	20,4	15,4	17,1
Autres catégories d'actifs.....	7,5	9,4	11,2	22,5	17,7	26,0
Inactifs.....	1,6	1,6	2,0	2,6	4,1	4,2
<b>Ensemble</b>	<b>5,4</b>	<b>6,2</b>	<b>6,7</b>	<b>5,4</b>	<b>6,2</b>	<b>6,7</b>

Sources : Pour 1970, les taux ont été établis par A. Boigeol et J. Commaille en rapportant, pour chaque catégorie, le nombre des hommes et des femmes engagés dans une tentative de conciliation au nombre des hommes et des femmes mariés dans le recensement de 1968 ([3], tableau 11, p. 14); les non-réponses sont ventilées proportionnellement aux catégories déterminées. Pour 1975 et 1976-1977, les taux ont été établis de manière analogue par D. Baillon et al. ([1], tome 1, tableau 9, p. 67).

divorce des « employés » ne dépassent sensiblement ceux des « ouvriers qualifiés » ou « spécialisés » que dans le cas des fréquences de divorce parmi les mariages de 1945 à 1975, et les chiffres des « cadres moyens » sont en général inférieurs à ceux de ces deux dernières sous-catégories d'« ouvriers ». Les données utilisées par A. Boigeol et J. Commaille, ainsi que par L. Roussel, les données du même type produites depuis, et les autres enquêtes comme l'enquête « Familles », ne permettent donc pas d'affirmer une surdivortialité des classes moyennes. Mais la critique développée ici ne démontre évidemment pas que cette surdivortialité n'existe pas<sup>3</sup>.

**L'exemple de la scolarisation**

Depuis 1947, le ministère de l'Éducation nationale effectue une enquête exhaustive sur l'origine sociale des élèves

2 B. Proportion de divorcés selon la catégorie socio-professionnelle (recensements)

En ‰

	Hommes		Femmes	
	1968	1975	1968	1975
Agriculteurs.....	6	8	6	8
Salariés agricoles.....	38	20	81	56
Patrons de l'industrie et du commerce.....	21	26	50	59
Professions libérales et cadres supérieurs.....	15	24	73	80
Cadres moyens.....	19	26	64	71
Employés.....	23	34	78	81
Ouvriers.....	23	28	78	76
dont :				
Contremaîtres.....	14		96	
Ouvriers qualifiés.....	22		75	
Ouvriers spécialisés et manœuvres.....	26		79	
Personnel de service.....	50	65	131	115
Autres catégories d'actifs.....	16	23	159	152
Inactifs.....	30	34	22	27
<b>Ensemble</b>	<b>22</b>	<b>28</b>	<b>36</b>	<b>45</b>

Source : recensements de la population; pour 1968, tableau DIV 589/V; pour 1975, tableau DG 65/C.

de l'enseignement secondaire public. D'abord annuelle et appliquée à l'ensemble des classes, cette enquête a vu sa périodicité s'espacer et son champ se restreindre aux classes des niveaux quatrième et seconde à partir des années soixante. Certaines années, elle a été étendue à l'enseignement privé.

Ce sont les établissements scolaires qui sont chargés de classer les professions des « chefs de famille » dans la liste des catégories socioprofessionnelles de l'INSEE. Cette liste est fournie avec, pour seules indications complémentaires, depuis 1967-1968, les critères du nombre de salariés pour distinguer parmi les « patrons », et du niveau d'instruction pour séparer « cadres moyens » et « supérieurs », ainsi que, depuis 1973-1974, quelques exemples de professions illustrant le contenu des catégories. Le travail de

3. Les autres données examinées dans [5] ne contredisent pas cette analyse.

\* NOTE GÉNÉRALE

2 C. Proportion de premiers mariages rompus pour les femmes interrogées dans les enquêtes « Familles », selon la catégorie socioprofessionnelle du mari \*\*

En ‰

	Enquête de 1962	Enquête de 1975	
	Probabilité de rupture sur 15 ans (tous mariages)	Probabilité de rupture sur 15 ans (mariages de 1945 à 1959)	Fréquence des ruptures parmi les mariages de 1945 à 1975
Agriculteurs.....	9	10	16
Salariés agricoles.....	37	38	53
Patrons de l'industrie et du commerce.....	34	40	58
Professions libérales et cadres supérieurs.....	35	32	46
Cadres moyens.....	39	43	56
Employés.....	51	47	66
Ouvriers.....	44	40	55
dont :			
Contremaîtres.....		11	20
Ouvriers qualifiés.....		44	59
Ouvriers spécialisés et manoeuvres.....		51	59
Personnel de service.....		60	175
Autres catégories d'actifs.....		70	
Inactifs.....		288	
<b>Ensemble</b>	<b>40</b>	<b>43</b>	<b>60</b>

\*\* En 1962, la probabilité de divorce pendant les quinze premières années du premier mariage est le complément, par rapport à 1, de la probabilité de survie du mariage dans le même délai, calculée d'après les quotients quinquennaux de dissolution par divorce de tous les premiers mariages enregistrés dans l'enquête ([16], tableau 25, p. 83). En 1975, la première colonne indique la probabilité de divorce, calculée de la même manière qu'en 1962, à partir des tableaux D 25 et D 34 de l'enquête, pour les premiers mariages contractés de 1949 à 1959. La deuxième colonne fournit une mesure d'une nature différente. Il s'agit simplement de la fréquence des divorces survenus parmi les premiers mariages conclus entre 1945 et la date de l'enquête (tableau D 25). Ces fréquences sont plus élevées que les probabilités des deux colonnes précédentes. Quelle que soit l'incidence de la différence de nature des mesures sur ce phénomène, il est très probable qu'il traduit au moins en partie l'accroissement de la divortialité, et notamment de la fréquence des divorces précoces, constaté depuis le milieu des années soixante.

Les indices de divortialité présentés dans les trois sous-tableaux sont de natures assez différentes. Dans les taux de divorcés d'après les recensements (2 B), et dans les fréquences de divorce d'après les enquêtes « Familles » (2 C), la catégorie socioprofessionnelle de chaque individu est, par construction, la même au numérateur et au dénominateur, alors que dans les taux de divortialité calculés par rapprochement de l'état civil et des recensements (2 A) le mode de détermination de la catégorie socioprofessionnelle est hétérogène. Mais la construction des données diffère aussi par d'autres aspects importants.

Le recensement et les enquêtes « Familles » dénombrent les personnes qui déclarent respectivement être divorcées ou avoir divorcé : la définition de ces notions par les répondants diffère évidemment, quoique de manière inconnue, de la définition légale des « tentatives de conciliation en divorce ou en séparation de corps » sur laquelle reposent les statistiques du ministère de la Justice.

Les données du recensement portent sur l'état de divorcé au moment du recensement, quelle que soit l'ancienneté du divorce, et elles concernent la profession de la personne au moment non du divorce mais du recensement. De ce fait, ces données se prêtent mal à une comparaison directe avec les taux de divortialité du ministère de la Justice qui portent sur le flux annuel d'événements que constituent les tentatives de conciliation. Pour faire une comparaison satisfaisante, il faudrait pouvoir tenir compte d'au moins trois facteurs : l'incidence du divorce lui-même sur le changement de catégorie socioprofessionnelle des femmes (certaines d'entre elles passent ainsi de l'inactivité à l'activité à la suite d'un tel événement); les probabilités différentielles de remariage selon les catégories; la mobilité professionnelle normale, notamment d'employé à cadre, car, compte tenu de l'âge souvent assez jeune auquel se produisent les séparations, et du fait qu'une partie des divorcés ne se remarient pas, des « divorcés d'employés » enregistrées par les tribunaux conduisent à des « cadres divorcés » saisis par les recensements.

Dans les enquêtes « Familles » la catégorie socioprofessionnelle du mari est celle au moment de l'enquête pour les femmes qui n'ont pas divorcé, mais celle au moment du premier mariage pour les femmes qui ont divorcé. En raison de la mobilité professionnelle, par exemple d'employé à cadre, la catégorie socioprofessionnelle des maris des femmes qui ont divorcé est ainsi rabaisée par rapport à celle des conjoints des femmes qui n'ont pas divorcé. De plus, les femmes qui ont divorcé sont à la fois moins disposées à parler de leur ancien mari, et moins à même d'indiquer avec précision sa profession : elles fournissent donc à ce sujet des réponses souvent vagues, qui conduisent fréquemment à des classements dans ces catégories intermédiaires que constituent les « employés » et les « cadres moyens ». Toutes choses égales par ailleurs, ces deux formes d'hétérogénéité dans l'identification de la catégorie socioprofessionnelle (différences d'époque et de précision) tendent donc simultanément à élever la proportion de maris de femmes qui ont divorcé dans les catégories les plus floues, principalement « employés » et secondairement « cadres moyens ».

TABLEAU 3

**Indices de scolarisation selon l'origine sociale**

3 A. Taux d'accès au niveau de la classe dite quatrième : rapprochement des statistiques d'origine sociale de l'Éducation nationale et de la répartition des enfants par catégorie socioprofessionnelle du chef de famille d'après les recensements de population \*

		En %										
		Agri- cul- teurs	Sala- riés agri- coles	Patrons	Profes- sions libé- rales, cadres supé- rieurs	Cadres moyens	Em- ployés	Ou- vriers	Per- sonnel de service	Autres caté- gories	Sans profes- sion	En- semble
1961-1962	Par référence, dans le recensement, aux générations de :											
	15 ans.....	35	28	83	77	93	107 <sup>1</sup>	41	49	97	14	55
	17 ans.....	36	31	79	72	99	113	44	44	97	13	55
	15 à 19 ans..... (sondage au 1/20)	35	31	80	75	94	109	44	46	95	14	55
	15 à 19 ans..... (exhaustif)	34	32	79	74	93	113	44	47	101	14	55
1973-1974	Par référence à la génération de 17 ans dans le recensement de 1975.....	92	99	122	92	112	140	94	140	162	57	102
1976-1977	Idem.....	82	80	118	105	119	137	96	143	159	63	103

1. L'apparition de pourcentages supérieurs à 100 % fait partie des propriétés de ces taux d'accès qui sont expliquées dans le texte.

\* Sources. — Sont placés au numérateur les effectifs d'élèves des classes de niveau quatrième selon la catégorie socioprofessionnelle du chef de famille, d'après les enquêtes de ministère de l'Éducation nationale sur les enseignements public et privé. Les classes « de niveau quatrième » sont les classes de quatrième proprement dites, les classes pré-professionnelles de niveau (CPPN) et les classes de première année des certificats d'aptitude professionnelle (CAP) en trois ans à temps plein. Est porté au dénominateur le nombre d'enfants de 14 ans révolus au premier janvier de l'année scolaire considérée (âge modal de l'ensemble des élèves au niveau quatrième) dans la population totale, selon la catégorie socioprofessionnelle du chef de famille. Ce nombre est obtenu en appliquant à la génération démographique de 14 ans à la même date, un coefficient de répartition des enfants par catégorie socioprofessionnelle du chef de famille, calculé d'après les tableaux de la série « SCO » des recensements de la population, pour une classe d'âge aussi voisine que possible de celle de 14 ans, soit 15 ans en 1962 et 17 ans en 1975.

L'âge des enfants dans les tableaux issus des recensements, y compris dans le tableau 3 B, est indiqué, conformément aux conventions de ces tableaux, en années révolues au premier janvier de l'année qui suit celle du recensement.

(Pour plus de précisions sur les sources de ce calcul, voir [4]).

Remarque. — Ces données sont inévitablement une approximation assez grossière. En effet, faute de renseignement, certaines classes du niveau quatrième ne peuvent être prises en compte : il s'agit des classes de fin d'études primaires, des sections spécialisées des cours complémentaires et de l'enseignement technique privé en 1961-1962; de l'enseignement spécial et de l'enseignement agricole sur toute la période. De plus, le taux de non-réponse des établissements scolaires est important en 1976-1977. Par ailleurs la répartition des enfants omis dans les tableaux « SCO », parce qu'ils ne sont pas rattachés à une « famille » d'un « ménage ordinaire » (de 5 à 10 % des enfants d'une génération), doit être supposée analogue à celle des enfants rattachés à une « famille » que ces tableaux prennent en compte. On remarquera toutefois que les taux d'accès, en 1961-1962 varient très peu, selon qu'on les calcule, à 15-19 ans, d'après l'exploitation exhaustive (SCO 7/TD) ou d'après le sondage au vingtième (SCO 4/1/SF) du recensement. Ils varient également assez peu selon qu'on prend comme base de calcul, dans le sondage de 1962 qui indique seul les répartitions par âge annuel, la classe d'âge de 15 ans ou celle de 17 ans, qui paraît a priori moins bien convenir : le choix forcé de la classe d'âge de 17 ans comme référence en 1975 (les tableaux SCO n'en comportant pas de plus jeune) n'introduit donc peut-être pas de distorsion importante.



3 B. Taux de scolarisation obtenus d'après des enquêtes homogènes (catégorie socioprofessionnelle identique au numérateur et au dénominateur) \*

	Agri- cul- teurs	Sala- riés agri- coles	Pa- trons <sup>1</sup>	Profes- sions libé- rales, cadres supé- rieurs	Cadres moyens	Em- ployés	Ou- vriers	Per- sonnel de service	Autres caté- gories	Inactifs	En- semble
En %											
1962											
Recensement de la population :											
Taux de scolarisation à 15 ans.....	55	56	80	95	93	92	67	70	87	66	71
Vers 1975-1977			A, PC	I, GC						Inactifs et divers	
« Panels » de l'Éducation nationale :											
Taux de passage au niveau 4° des élèves qui étaient en 6° en 1972, 1973 et 1974.....	87	90	97	99	99	100	97	99	99	98	98
1. Les « panels » de l'Éducation nationale distinguent « Artisans, Petits Commerçants » et « Industriels, Gros Commerçants ».											

\* Sources. — 1962 : tableau SCO/1/SF du recensement.

1975-1977 : [26]. Dans cet échantillon représentatif des enseignements public et privé, les effectifs d'élèves présents en sixième ou en première année des sections d'enseignement spécialisé des collèges, en 1972, 1973 et 1974, fournissent la base du dénominateur. Cette définition du point de départ des cohortes suivies correspond à une étape de la scolarisation par où passe à peu près la totalité d'une génération démographique. Le numérateur résulte de l'addition des taux d'accès des enfants de ces cohortes en quatrième ([26], p. 34), en classe pré-professionnelle de niveau et en première année des classes préparant au Certificat d'aptitude professionnelle en trois ans ([26], p. 19). Les élèves qui ont fréquenté plus d'un de ces trois types de classe sont comptés autant de fois, mais les redoublements dans un même type de classe sont comptés une seule fois.

Remarque. — La catégorie socioprofessionnelle du chef de famille dans le tableau 3 B est identique pour un même individu au numérateur et au dénominateur. Mais la construction de ces données diffère également de celle des données du tableau 3 A par la définition des situations scolaires des enfants que saisissent les enquêtes. Le champ des filières scolaires prises en compte est plus vaste dans le recensement — mais dans les « panels », il est à peu près le même — que dans l'enquête exhaustive de l'Éducation nationale. L'indice retenu pour mesurer l'intensité de la scolarisation est également d'une nature un peu différente, comme on peut le voir aisément. De plus, les taux de scolarisation indiqués par le recensement reposent sur la déclaration de la situation scolaire des enfants fournie par les familles, alors que les enquêtes de l'Éducation nationale, y compris l'enquête « panel », reposent sur la déclaration des établissements scolaires.

classement est généralement effectué par le personnel administratif ou de surveillance des établissements, à l'aide de cette seule liste. Les renseignements utilisés sont les professions qu'ont indiquées les familles sur les fiches ou les registres d'inscription des élèves qui servent à la gestion courante. Dans ces documents administratifs, la seule rubrique se rapportant à la catégorie socioprofessionnelle se résume en général à la formule « Profession », pour le père, la mère et éventuellement le tuteur. Dans les divers établissements où j'ai consulté les fiches, la réponse à cette question manque très rarement.

Les données issues de ces enquêtes ont été utilisées par les sociologues de l'éducation, qui les ont parfois rapprochées de la répartition socioprofessionnelle de la population globale donnée par les recensements, pour calculer

des taux d'accès à une filière ou à un niveau d'enseignement. On trouvera ici les résultats d'un calcul de ce type effectué, pour les élèves du niveau quatrième, les années où l'enquête de l'Éducation nationale a inclus l'enseignement privé.

Pour comparaison, on trouvera aussi d'autres mesures de l'intensité de la scolarisation par origine sociale, qui mettent chacune en œuvre, pour un individu donné, une seule et même origine sociale. Il ne fait aucun doute que les propriétés de la scolarisation exprimées par ces données sont, dans leurs grandes lignes, plausibles (tableau 3 B). Il est plausible en effet qu'en 1962 les catégories sociales dont les enfants de 14-15 ans sont les plus scolarisés soient les cadres, celles dont les enfants sont les moins scolarisés les paysans et les ouvriers, et que les enfants des employés ne soient pas plus scolarisés que ceux des cadres. De même,

vers 1975, la scolarité étant en particulier devenue obligatoire jusqu'à 16 ans, il est plausible que les enfants de cet âge soient scolarisés dans leur quasi-totalité pour toutes les catégories socioprofessionnelles. Seuls les enfants d'agriculteurs semblent faire exception, mais c'est que beaucoup d'entre eux fréquentent des établissements relevant du ministère de l'Agriculture et ne sont donc pas recensés par l'Éducation nationale.

Par contre, le calcul des taux d'accès par rapprochement fait apparaître quelques résultats surprenants (tableau 3 A). Tout d'abord certains taux dépassent 100 %. Lorsqu'il concerne exclusivement certaines catégories socioprofessionnelles, comme en 1961-1962, ce dépassement traduit seulement l'effet des divergences entre les deux sources dans la manière de recueillir et de classer les professions (voir ci-dessous). Mais en 1973-1974 et en 1976-1977, le taux d'accès global lui-même excède 100 %. Cet excès est d'autant plus remarquable que les enquêtes de l'Éducation nationale sur l'origine sociale ne sont pas réellement exhaustives (voir notes du tableau 3). Il s'explique par deux phénomènes d'une autre nature. D'une part, dans les années soixante-dix, les recensements de l'Éducation nationale dénombrent chaque année nettement plus d'élèves au niveau quatrième que d'élèves de 14 ans; cette concentration d'élèves au niveau quatrième, quelle qu'en soit la raison, explique l'essentiel de l'excès, par rapport à 100 %, du taux d'accès de l'ensemble. D'autre part — le fait est bien connu — une partie des chefs d'établissement majorent légèrement leurs effectifs, alors qu'à l'inverse les recensements sous-estiment peut-être le nombre des enfants; pour les classes d'âge qui sont totalement scolarisées, les effectifs d'élèves dénombrés par l'Éducation nationale sont d'ailleurs régulièrement un peu supérieurs au nombre d'enfants du même âge recensés par l'INSEE [2].

Les autres propriétés les plus surprenantes *a priori* concernent certaines catégories socioprofessionnelles. Ainsi, le taux d'accès des enfants d'« employés » est nettement supérieur à celui des enfants de « cadres moyens », lequel est nettement supérieur au taux d'accès des enfants de « professions libérales et cadres supérieurs ». Par ailleurs, le taux d'accès des enfants des « autres catégories » est très élevé, dans l'absolu en 1973-1974 et en 1976-1977, mais aussi en 1961-1962 si on compare le rapport entre ce taux et celui de l'ensemble avec le rapport homologue entre indices de scolarisation tirés du recensement. Enfin, le taux d'accès de la catégorie « sans profession » est particulièrement bas, si on le compare, en 1961-1962, avec le taux de scolarisation de la même catégorie d'après les recensements, et, en 1973-1974 et 1976-1977, compte-tenu du fait que la scolarisation est devenue obligatoire jusqu'à 16 ans.

Pour expliquer ces résultats, j'ai étudié l'ensemble du processus de production des données sur l'origine sociale des élèves. En particulier dans quatre établissements, deux situés en proche banlieue parisienne et deux dans deux petites villes de province, j'ai codé les renseignements sur les professions des parents contenus dans les fiches d'inscription, et comparé mes résultats à ceux obtenus par les codeurs des établissements d'après les mêmes fiches. Pour ce codage, je me suis servi du *Code des catégories socio-*

*professionnelles* et des *Codes (analytique/alphabétique) des métiers* utilisés par l'INSEE pour le chiffrage des recensements. Toutefois, afin de ne saisir que les différences les plus claires entre la pratique des codeurs des établissements et les principes officiels des codes de l'INSEE, j'ai suivi les instructions des documents de l'Éducation nationale quand elles étaient contraires à ces codes, et par ailleurs, dans les cas indécidables, j'ai choisi le classement qui allait dans le sens de celui réalisé par les codeurs des établissements <sup>4</sup>.

Les conclusions tirées de cette étude permettent d'interpréter les propriétés des taux d'accès énumérées plus haut.

Dans la plupart des quatre établissements et au total, je dénombre plus de « professions libérales et cadres supérieurs » et moins de « cadres moyens » que les codeurs des établissements. Ce fait, qui est cohérent avec les taux d'accès nationaux, s'éclaire par les témoignages recueillis. D'une part, les codeurs jugent souvent que les « cadres » se surestiment, et on peut supposer qu'ils tendent donc à les rabaisser. D'autre part, ils exposent spontanément diverses solutions prudentes adoptées (« quand j'hésite, je le mets dans « cadre moyen », on ne se trompe pas trop en le mettant là »).

Le résultat le plus significatif de cette enquête concerne les « employés ». Contrairement à mon attente, je n'ai jamais trouvé moins d'« employés » que les codeurs, mais presque toujours plus. L'explication de ce résultat est en fait très simple. Comme les codeurs, mais suivant quant à moi à la lettre les indications des codes utilisés par l'INSEE pour les recensements, j'ai rangé dans cette catégorie toutes les mentions « employé », « agent » et « fonctionnaire » sans autre indication, qui représentaient environ *la moitié* de l'ensemble des cas classés comme « employés ». C'est donc le caractère vague des renseignements qui conduit les codeurs des établissements scolaires, comme toute autre personne placée dans la même situation qu'eux, à gonfler massivement les effectifs de la catégorie « employés ».

J'ai également obtenu à peu près les mêmes résultats que les codeurs pour les « autres catégories » (artistes, clergé, armée et police). Sur ce point, je ne peux que conclure globalement à la non-coïncidence des résultats de mon enquête avec les propriétés des taux d'accès nationaux. Ces taux ne laissent pourtant, on l'a vu, aucun doute quant au gonflement de cette catégorie dans les statistiques de l'Éducation nationale. On peut trouver ailleurs une confirmation de ce que la raison de ce gonflement est bien, comme il est facile de l'imaginer, le rôle de catégorie-rebut que son intitulé fait jouer à cette rubrique [4]. En 1967-1968 en effet, les formulaires de l'enquête de l'Éducation nationale, qui ne concernaient que l'enseignement public, ont comporté exceptionnellement une rubrique supplémentaire « autres professions (catégories non mentionnées ci-dessus) », distincte de la rubrique « autres catégories » de l'INSEE. Cette année-

---

4. Un compte rendu détaillé de cette étude sera publié ultérieurement.

là, le taux d'accès des élèves issus des « autres catégories » (au sens de l'INSEE) est inférieur à leur taux d'accès dans l'enseignement public en 1961-1962, alors que les taux d'accès des élèves issus des autres catégories socioprofessionnelles sont tous nettement supérieurs. Tous, sauf un, celui des « employés », qui, comme celui des « autres catégories », est inférieur : ceci confirme le rôle de refuge pour les appellations vagues que joue aussi la catégorie « employés ».

Les codeurs des quatre établissements étudiés avaient classé tous les cas, et évité de ranger les plus difficiles dans la catégorie « sans profession », sans doute parce que cela aurait pu apparaître comme une négligence dans leur travail. Ce fait n'est pas incompatible avec le faible taux d'accès national calculé ici pour cette dernière catégorie (« inactifs » dans la terminologie de l'INSEE), mais il ne suffit pas à l'expliquer. Je suppose — bien que cette hypothèse ne soit confirmée dans aucun des quatre établissements — que dans l'ensemble les codeurs utilisent assez souvent la profession de la mère lorsque la situation du père n'apparaît pas constituer vraiment une profession, par exemple lorsqu'il est « retraité », « invalide », « chômeur », ou « étudiant » (toutes ces mentions apparaissent dans les fiches d'inscription, mais les dernières citées moins souvent que les premières, bien entendu). Ils adopteraient ainsi une autre logique que celle sur laquelle reposent les tableaux de l'INSEE. L'INSEE définit en effet *a priori* le critère « chef de famille » (déclaré), qui désigne impérativement la personne à classer, fût-ce dans une catégorie d'« inactifs » qui ne correspond pas à une position sociale. Les codeurs, au contraire, chercheraient par tous les moyens disponibles à assigner une origine sociale aux élèves, donc à trouver une profession proprement dite, chez la mère à défaut du père.

Ces résultats peuvent être comparés à ceux de deux enquêtes sur le processus de codage des origines sociales des étudiants dans les universités de Paris I et de Nantes. Deux différences doivent toutefois être notées entre ce processus et celui en vigueur dans l'enseignement secondaire. Dans les tableaux de ces universités, les rubriques « catégories non mentionnées » et « non-réponse » regroupent environ 20 % des inscrits, alors que la première n'a existé qu'en 1967-1968 dans le secondaire, et la seconde jamais. Par ailleurs, à Paris I, les étudiants se classent eux-mêmes dans la liste des catégories socioprofessionnelles. D. Merllié [25] et J. Peneff [27] ont interrogé les étudiants sur la profession de leur père, avec un questionnaire analogue à celui du recensement. Ils ont classé les réponses par catégorie socioprofessionnelle en utilisant les codes de l'INSEE. J. Peneff a également recodé, à l'aide des mêmes instruments, les renseignements sur la profession contenus dans les dossiers d'inscription des étudiants et antérieurement codés par l'université. Par rapport aux résultats ainsi obtenus, il apparaît notamment que les statistiques universitaires surestiment le nombre des enfants d'« employés » et sous-estiment celui des enfants de « professions libérales et cadres supérieurs ». Les écarts sont très importants par référence à l'enquête directe réalisée auprès des étudiants. Ils sont moins importants lorsque le point de comparaison est le recodage des dossiers d'inscription.

## Conclusions

On peut tirer des analyses précédentes plusieurs types de conclusions, qui concerneront d'abord les processus conduisant au gonflement de la catégorie des « employés ».

Les données rassemblées ici ne confirment pas l'idée, fréquemment avancée, selon laquelle l'un des biais principaux de toute enquête déterminant la profession sur la seule base de la déclaration des répondants serait la propension généralisée de ceux-ci à rehausser leur position sociale. Les enquêtes de D. Merllié [25] et de J. Peneff [27] décèlent à l'inverse une tendance des étudiants, du moins de ceux issus des classes supérieures, à rabaisser leur origine sociale lors de l'inscription universitaire. Même si l'on admet l'existence, dans certaines circonstances, d'une propension à la surestimation, son incidence paraît moins importante que celle des *propriétés des situations*. Dans les exemples analysés, la situation de recueil n'incite qu'à des déclarations sommaires, et la situation de codage, compte tenu de l'insuffisance de l'information, conduit au gonflement de catégories sociales moyennes, et non supérieures<sup>5</sup>.

La surestimation constante du nombre des « employés » dans les données issues de sources administratives exprime ainsi de manière typique l'un des principaux problèmes qu'affrontent, sous des formes un peu différentes, les personnels chargés du codage dans les organismes statistiques et dans les autres administrations. Pour les codeurs des organismes statistiques, le problème vient de la définition assez floue de cette catégorie des « employés » par une série de négations (en simplifiant : ni ouvriers, ni cadres), et de la fréquence d'intitulés sommaires et vagues dans les sources administratives — parmi lesquels en particulier les expressions homonymes de l'intitulé de la catégorie (« employé de » ou « chez ») (voir [17], p. 44; [19], p. 27; [21], p. 35-37). Témoignent de ce problème les anomalies des taux de mortalité et de divortialité calculés par rapprochement. Le problème ainsi posé ne peut recevoir de solution satisfaisante, en dépit de la conscience qu'en ont les chiffreuses professionnelles ([28], p. 127 et 167; [21], p. 150). Pour les codeurs occasionnels des administrations, ceux des établissements scolaires par exemple, ces difficultés de classement sont accrues par leur moindre connaissance du code des catégories socioprofessionnelles de l'INSEE. Ils tendent, par prudence, à ranger un grand nombre de cas mal déterminés à leurs yeux dans des catégories dont la position dans la structure sociale semble « moyenne » ou dont l'intitulé est ambigu — soit principalement les « employés », « cadres moyens » et « autres catégories ». (Ceci est encore accentué par l'absence d'une catégorie explicitement résiduelle.)

Cet examen des problèmes rencontrés par les codeurs et des solutions qu'ils adoptent pourrait être poursuivi dans le contexte plus large d'une analyse des propriétés de l'orga-

5. Je partage sur ce point l'appréciation de D. Merllié [25], p. 15.

nisation des établissements concernés — écoles, mairies, tribunaux — et du travail des personnels impliqués dans le codage. Toutefois, pour l'objet étudié ici, cette perspective de recherche ne serait probablement pas des plus intéressantes. Le classement des professions est en effet une tâche de faible conséquence à la fois pour les organisations et pour les personnels qui en sont chargés. L'analyse ainsi envisagée ne serait donc guère susceptible d'éclairer des aspects essentiels du fonctionnement des institutions et du comportement de leurs agents, comme cela est au contraire le cas pour l'analyse de ces problèmes centraux que sont les effectifs pour l'institution scolaire [8] la caractérisation des « cas » pour la police ou la justice [10] ou pour les hôpitaux ([18], p. 186-207). La détermination de la catégorie socioprofessionnelle elle-même peut aussi revêtir des conséquences plus importantes dans d'autres contextes que ceux étudiés ici : par exemple dans les entreprises, lorsque la classification socioprofessionnelle constitue un enjeu pour les employeurs, les salariés, les administrations de tutelle ([21], p. 29-38 et 83-103) et dans les enquêtes statistiques et sociologiques, lorsque le recueil et le classement de renseignements divers, dont ceux concernant la profession, constituent la tâche principale du travail ordinaire des agents [29].

On pourrait également poursuivre l'analyse en cherchant à caractériser plus systématiquement les *catégories de perception* de la structure sociale qui sont propres aux codeurs, et influencent évidemment les classements réalisés. Cette perspective de recherche semble toutefois peu susceptible d'aboutir à des résultats systématiques, en raison du caractère diffus de ces catégories et composite du personnel, notamment dans les institutions autres que les organismes statistiques. Les obstacles de ce type seraient encore plus grands si l'on cherchait aussi à caractériser systématiquement les *attitudes des déclarants* — comme l'exigerait en principe une analyse du processus d'ensemble de production des statistiques socioprofessionnelles.

## Le point de vue de l'utilisateur

J'ai concentré l'analyse sur les « employés » parce que les effectifs de cette catégorie sont surestimés dans toutes les sources administratives étudiées ici. Mais il ne faut pas en conclure que les problèmes se limiteraient à cette catégorie, ne serait-ce que parce les excédents dont elle est créditée sont évidemment prélevés dans d'autres catégories (encore que les données rassemblées ici ne permettent d'identifier aucune structure générale des transferts). Et, même si, mis en garde par les anomalies que font apparaître les calculs de taux par rapprochement de deux sources distinctes, on se limite à des calculs internes à chaque source pour mesurer la fréquence de certains comportements par catégorie socioprofessionnelle, une autre précaution s'impose. Encore faut-il en effet retenir que, comme le révèlent les taux obtenus par rapprochement, les catégories n'ont pas la même composition dans les diverses sources que l'on sera tenté de mettre en rapport, sinon par le calcul d'un taux, du moins par l'emploi d'un mot, en reliant entre elles les propriétés d'une « même » catégorie relevées dans des sources

différentes. (De ce point de vue, le calcul de taux par rapprochement de sources distinctes garde une certaine utilité, puisqu'il peut permettre de mettre en évidence l'hétérogénéité de catégories socioprofessionnelles nominale-ment identiques.)

Par ailleurs, en prenant pour point de comparaison les données d'enquêtes homogènes quant à la catégorie socioprofessionnelle, je n'ai pas voulu suggérer que celles-ci révèlent les « véritables » taux de chaque catégorie. La vraisemblance ou l'invraisemblance des différents types de données statistiques ne peut être établie que par le recoupement de preuves d'origines diverses. Sous ce rapport, la situation n'est pas la même dans les trois cas examinés : les anomalies de taux calculés par rapprochement sont évidentes dans les exemples de la mortalité et de la scolarisation, parce qu'il existe d'autres données beaucoup plus plausibles; en l'absence de telles données pour le divorce, les anomalies peuvent paraître en ce cas moins évidentes.

Les enquêtes auprès des individus, en particulier, ne sont évidemment pas dépourvues d'incertitudes. Celles-ci sont évidentes pour les enquêtes d'opinion, de motivation ou d'intention, qui ne saisissent dans une large mesure que des phénomènes produits par l'enquête elle-même. Les incertitudes peuvent sembler moins fortes dans les enquêtes qui cherchent à atteindre des comportements, des statuts ou des possessions, c'est-à-dire des caractéristiques qui existent indépendamment de l'enquête, et qui seraient vérifiables par observation directe ou consultation de documents officiels. Il est cependant facile d'apercevoir que, sur certains sujets, les répondants n'ont pas de raisons ou pas de catégories pour saisir et mémoriser leur propre expérience; leurs catégories ne coïncident pas nécessairement avec celles que mettent en œuvre les enquêtes; ils peuvent enfin avoir des raisons de ne pas mentionner, ou n'avoir pas de raison de mentionner, tel aspect de leur expérience dans les circonstances de l'enquête, face à un enquêteur d'un sexe, d'un âge, d'une classe sociale déterminés.

## Les sociologues et la statistique

Les sociologues sont tributaires de ces repères sur la différenciation sociale que constituent les statistiques produites à grande échelle par des organismes dont ils n'influencent pas, ou très peu, les activités. Les données statistiques qu'ils peuvent utiliser sont largement déterminées, souvent à leur insu, par les conditions particulières de leur production. Il est donc fréquent qu'ils ignorent le degré d'approximation des enquêtes utilisées, en particulier pour la signification concrète des catégories de classement mises en œuvre, et, par voie de conséquence, le degré de cohérence entre elles des données juxtaposées dans l'analyse. De plus, on ne connaît que rarement les comportements de catégories sociales délimitées, possédant une assez grande homogénéité, et susceptibles par là de fournir des repères plus sûrs pour interpréter les caractéristiques manifestées

par des agrégats statistiques composites. Dans ces conditions, il paraît inutile d'employer des techniques d'analyse des données formellement sophistiquées, dont les résultats ne sont pas contrôlables, et il est aventureux de chercher à développer des théories sociologiques concernant les propriétés des classes sociales sur la base des seules données statistiques.

La contribution propre des sociologues peut mieux s'exercer par d'autres voies. Ils peuvent contribuer à éclairer le contenu des catégories socioprofessionnelles et les propriétés statistiques des groupes qu'elles désignent, grâce à l'étude monographique de populations circonscrites, par exemple tel type d'employé des banques ou des postes. Ils peuvent aussi étudier les rapports sociaux dans des institutions particulières, au moyen de ces formes d'étude directe que constituent aussi bien l'observation *in situ* des activités que l'analyse des documents liés au fonctionnement des organisations. Les sociologues peuvent enfin travailler à développer l'analyse sociologique des processus de production de statistiques, en s'appuyant d'ailleurs sur des travaux déjà réalisés dans divers contextes, que je rappellerai brièvement.

Un premier centre d'intérêt concerne la production de statistiques par les institutions pour lesquelles cette production n'est pas la finalité principale, mais constitue un instrument de leur propre gestion. Le système scolaire, la justice, la police, le système hospitalier ou la sécurité sociale en sont des exemples. Il serait utile de disposer pour plusieurs de ces institutions de descriptions systématiques de la manière dont elles construisent des données, qui prennent finalement la forme de statistiques, au sein même du processus par lequel elles définissent les populations qu'elles traitent. On apercevrait ainsi que les statistiques établies mesurent en première analyse non des propriétés de ces populations, mais des aspects de l'activité des institutions, aspects saisis à travers les catégories propres à ces institutions. Ces statistiques constituent, par conséquent, une voie d'accès à la connaissance des catégories en question ([8], p. 684-685).

Les premiers exposés systématiques de cette perspective de recherche, parus dans les années soixante, sont dus à H. Garfinkel, J. Kitsuse et A. V. Cicourel [18; 22; 10]. Les premières enquêtes empiriques détaillées mettant en œuvre ce point de vue ont été réalisées à la fin des années cinquante et au début des années soixante, et sont rapportées dans les livres de H. Garfinkel [18] et d'A. V. Cicourel [10]. Mais ces recherches, ainsi que d'autres relevant du même courant « ethnométhodologique », sont en général

plus attentives au détail des procédures de caractérisation des cas qui sont à la base de la constitution des statistiques, qu'à la mesure des effets de ces procédures dans les statistiques finales. Toutefois, quelques études postérieures réalisées aux États-Unis, s'inspirant de la perspective tracée par l'article de Kitsuse et Cicourel, et consacrées le plus souvent aux statistiques de la police, ont mis en évidence l'incidence de l'organisation du travail et des caractéristiques des agents sur les propriétés des séries statistiques établies par les administrations au niveau de circonscriptions locales (voir par exemple [24]).

En France aussi, des études ont été consacrées, principalement depuis le milieu des années soixante-dix, à la production de statistiques. Mais leur centre d'intérêt est un peu différent. Ces études ont été effectuées, pour l'essentiel, par des statisticiens de l'INSEE et concernent les activités de cet organisme. La plupart de ces recherches se sont attachées soit à décrire l'élaboration et l'évolution des nomenclatures d'activité économique et professionnelle, soit à mesurer les variations de la détermination de l'activité professionnelle en fonction des catégories de classement adoptées, de la formulation des questions posées dans les questionnaires et des conditions du codage (voir notamment [13] et [21]). Toutefois, aucune étude n'a cherché à décrire systématiquement l'ensemble du processus de production d'un type de données, par exemple d'un recensement. Certes, il a plusieurs fois été remarqué que la représentation des phénomènes que fournit une enquête est influencée par les caractéristiques des situations et des agents, enquêtés compris, à chaque phase de l'enquête (voir par exemple [14]). Mais ce schéma d'analyse a toujours été illustré par des exemples ponctuels empruntés à des enquêtes diverses. À l'inverse, les recherches de M. Volle [31; 32] sur le système de statistique industrielle décrivent bien un grand nombre d'aspects de ce système. Mais elles n'examinent que succinctement la mise en œuvre concrète des catégories de classement et l'influence des caractéristiques de la division du travail sur les séries statistiques finalement produites.

Un aspect reste peu étudié en France : l'ajustement des enquêtés et des enquêteurs à la situation d'entretien ou de passation d'un questionnaire, et l'interaction entre enquêtés et enquêteurs dans cette situation (Pour quelques exemples récents d'analyse de cet aspect, voir [21], p. 55-56; [14], p. 13; [9], p. 231-299 et p. 579-602). Les recherches, parfois assez anciennes, menées par des sociologues américains se rattachant à des courants divers, montrent pourtant de manière convaincante que cet aspect ne devrait pas être négligé [20; 11; 15].

## RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] D. BAILLON, M. COSTECALDE, G. GODIN, B. MUNOZ-PEREZ : « Le divorce en France », *Les Collections de l'INSEE*, n° D 85-86.
- [2] L. BOCQUET : « Scolarité de générations successives », *Études et Documents*, n° 81-1, ministère de l'Éducation nationale.
- [3] A. BOIGEOL et J. COMMAILLE : « Le divorce en France », *Économie et statistique*, n° 53, février 1975.
- [4] J.-P. BRIAND : *Conditions et résultats du travail de classement des origines sociales dans les statistiques scolaires*, Centre de recherches socio-historiques sur l'éducation de l'Université de Paris VIII, multigraphie, 1982.
- [5] J.-P. BRIAND : *Anomalies statistiques et usages du code des catégories socioprofessionnelles*, Centre de recherches socio-historiques sur l'éducation de l'Université de Paris VIII, multigraphie, 1983.
- [6] J.-P. BRIAND, J.-M. CHAPOULIE et H. PERETZ : *L'inculcation en théorie et le calcul en pratique*, Université de Paris VIII, multigraphie, 1974.
- [7] J.-P. BRIAND, J.-M. CHAPOULIE et H. PERETZ : « L'inculcation en théorie et le calcul en pratique », *La Pensée*, n° 213-214, juillet-septembre 1980.
- [8] J.-P. BRIAND, J.-M. CHAPOULIE et H. PERETZ : « Les statistiques scolaires comme représentation et comme activité », *Revue française de sociologie*, t. XX, n° 4, octobre-décembre 1979.
- [9] J.-M. CHAPOULIE : *Le corps professoral de l'enseignement secondaire en France vers 1975*, thèse de doctorat, Université de Paris IV, 1984.
- [10] A. V. CICOUREL : *The social organization of juvenile justice*, New York, Wiley, 1968.
- [11] A. V. CICOUREL : *Theory and method in a study of Argentine fertility*, New York, Wiley, 1974.
- [12] G. DESPLANQUES, « La mortalité des adultes suivant le milieu social, 1955-1971 » *Les Collections de l'INSEE*, D 44.
- [13] A. DESROSIÈRES : « Éléments pour l'histoire des nomenclatures socioprofessionnelles », in : *INSEE, Pour une histoire de la statistique*, t. I, 1977.
- [14] A. DESROSIÈRES : « La portée sociologique des diverses phases du travail statistique », *Cahiers Analyse-Epistémologie-Histoire*, Université de Lyon II, 1984, n° 2.
- [15] I. DEUTSCHER, *What we say/What we do*, Glenview (Illinois), Scott, Foresman, 1973.
- [16] J.-C. DEVILLE, « Structure des familles. Enquête de 1962 », *Les Collections de l'INSEE*, n° D 13-14.
- [17] M. FEBVAY et M. AUBENQUE : « La mortalité par catégorie socioprofessionnelle », *Études statistiques*, n° 3, juillet-septembre 1957.
- [18] H. GARFINKEL : *Studies in ethnomethodology*, Englewood Cliffs (N. J.), Prentice-Hall, 1967.
- [19] A. GIRARD, L. HENRY, R. NISTRÉ : « Facteurs sociaux et culturels de la mortalité infantile », INED, *Cahier « Travaux et documents »*, n° 36.
- [20] H. H. HYMAN et alii, *Interviewing in social research*, University of Chicago Press, 1954.
- [21] INSEE : « Les catégories socioprofessionnelles et leur repérage dans les enquêtes », *Archives et Documents*, n° 38.
- [22] J. I. KITSUSE et A. V. CICOUREL, « A note on the uses of official statistics », *Social Problems*, vol. 11, n° 2, fall 1963.
- [23] C. LEVY et S. VALLIN : « La mortalité par catégorie socioprofessionnelle. Un essai de calcul direct », *Population*, vol. 36, n° 4, juillet-octobre 1981.
- [24] R. Mc CLEARY, B. C. NIENSTEDT, J. M. ERVEN : « Uniform Crime Reports as organizational outcomes : Three time series experiments », *Social Problems*, vol. 29 n° 4, avril 1982.
- [25] D. MERLLIÉ, « Une nomenclature et sa mise en œuvre. Les statistiques sur l'origine sociale des étudiants », *Actes de la recherche en sciences sociales*, n° 50, novembre 1983.
- [26] Ministère de l'Éducation nationale, SEIS : « L'entrée dans le second cycle long des élèves admis en sixième en 1972, 1973 et 1974 », *Document de travail*, n° 254.
- [27] J. PENEFF : « La fabrication statistique ou le métier du père », *Sociologie du travail*, n° 2-84, avril-mai-juin 1984.
- [28] L. PINSKY, R. KANDAROUN, G. LANTIN : « Le travail de saisie-chiffrement sur terminal d'ordinateur », *Collection de physiologie du travail et d'ergonomie* du CNAM, n° 65.
- [29] J. ROTH : « Hired hand research », *American Sociologist*, 1 august 1966.
- [30] L. ROUSSEL : « Le mariage dans la société française », *Cahier « Travaux et Documents »* de l'INED, n° 73.
- [31] M. VOLLE : « Statistique fétichisée et statistique réelle », *Le mouvement social*, n° 104, juillet-septembre 1978.
- [32] M. VOLLE : *Histoire de la statistique industrielle*, *Economica*, 1982.