

LA DISPERSION DES DÉPENSES ET LE REVENU DES MÉNAGES

Suite de l'analyse
d'un sous-échantillon de l'enquête de 1956

par

Michel GUILLOT

I. INTRODUCTION

On sait que le niveau des dépenses dans un groupe de ménages obéit à l'action de facteurs que les analyses classiques de budgets familiaux s'efforcent de repérer. Autour de leur moyenne les dépenses des différents ménages sont souvent très dispersées même lorsque les ménages sont de conditions sociales et économiques voisines. L'objet de cet article est d'examiner si cette dispersion des dépenses, un peu à la façon des moyennes, varie d'un groupe de ménages à l'autre en raison des différentes caractéristiques de ces groupes : la taille des familles qui les composent ou la moyenne de leurs revenus par exemple. L'étude des dispersions peut contribuer à une meilleure connaissance du comportement des consommateurs, mais il n'est généralement pas possible de l'aborder pour la raison évidente qu'elle exige le recours aux observations individuelles et non plus aux données groupées, comme il est fait le plus souvent afin d'alléger les procédures de calcul ⁽¹⁾.

Chaque type de dépenses pourrait faire l'objet de cet examen mais les données disponibles limitent nos possibilités. Ces données consistent uniquement en un sous-échantillon de 983 ménages, extrait de l'enquête sur les budgets familiaux de 1956 et déjà utilisé par J. Cramer dans l'article précédent. Sans qu'il s'agisse d'autre chose que d'une division a priori, nous nous sommes efforcés de regrouper dans des catégories distinctes les dépenses qui diffèrent soit par le rythme auquel elles sont effectuées, soit par le processus de décision auquel elles obéissent. Nous divisons ainsi la dépense totale en quatre catégories :

— **les dépenses alimentaires,**

— **les dépenses permanentes,** qui représentent des dépenses à peu près irréversibles au cours d'une année et qui hypothèquent par conséquent une partie du revenu : le loyer du logement, les frais d'éducation, les pensions, les mensualités pour l'achat de biens durables...

(1) Pour réaliser cette étude nous avons bénéficié de l'aide de Danièle PRANGÈRE qui a assumé la responsabilité de toutes les opérations techniques, calculs et tests, qui ont été nécessaires.

G. VANGREVELINGHE, Administrateur à l'I.N.S.E.E., a bien voulu revoir le manuscrit et faire plusieurs remarques dont nous le remercions.

— **les dépenses exceptionnelles** telles que les frais de réparation du logement, les achats au comptant de gros biens durables, de bijoux, ...

— enfin les autres dépenses, dénommées ici **dépenses courantes**, qui comprennent par exemple les achats d'habillement, de produits d'entretien, les frais d'utilisation de véhicules... ⁽¹⁾.

Les 983 ménages que compte l'échantillon sont uniquement des ménages urbains de province, composés de deux adultes avec un ou plusieurs enfants. Ils sont répartis en **trente sous-groupes** selon le double classement de la catégorie socio-professionnelle du chef de ménage et du type de famille ⁽²⁾. De façon secondaire, nous avons recours à une autre division qui dépend uniquement des revenus déclarés ; les 983 ménages sont alors rangés en 11 sous-groupes dans l'ordre croissant des revenus ⁽³⁾.

Dans toute cette étude, c'est par leur variance que nous mesurons la dispersion des dépenses ; mais cet indicateur commode est évidemment bien moins riche que la distribution sur laquelle il est calculé et son sens concret n'apparaît vraiment qu'après examen de celle-ci. Il n'est donc pas inutile, en commençant, de faire quelques remarques sur les formes empiriques que présentent les distributions de chaque catégorie de biens. C'est ce que nous faisons dans la section 2.

Nous considérons ensuite rapidement au début de la section 3 la variance totale de chaque catégorie de dépenses pour les trente sous-groupes de ménages. Il s'agit là seulement d'un examen brut des données. Mais à l'intérieur de chaque sous-groupe, tous les ménages n'ont pas le même revenu ; ces différences expliquent une partie de la dispersion des dépenses, superposant leur action aux influences éventuelles des caractéristiques du sous-groupe (catégorie socio-professionnelle, type de famille, revenu moyen). Or ces dernières influences seules nous intéressent. Pour mieux juger de leur existence nous éliminons donc la part du revenu dans la dispersion à l'intérieur de chaque sous-groupe et nous comparons alors entre eux les différents sous-groupes de ménages. Le passage des variances totales aux variances résiduelles est effectué dans la suite de la section 3 et les comparaisons de la variance résiduelle entre les sous-groupes dans les sections 4 et 5 qui sont les plus importantes de l'article ; la section 4 est consacrée aux variances résiduelles des quatre catégories de consommation ; la section 5, à la variance résiduelle de la dépense totale.

2. LES DISTRIBUTIONS DES DIFFÉRENTES DÉPENSES

L'examen des distributions montre que, dans tous les sous-groupes, il faut distinguer nettement le cas des dépenses alimentaires et des dépenses courantes d'une part, celui des charges permanentes et des dépenses exceptionnelles d'autre part. Pour les premières, les ménages se distribuent de façon à peu près symétrique autour de la dépense moyenne du sous-groupe. Pour les charges permanentes et les dépenses exceptionnelles, il n'y a aucune symétrie : les niveaux de dépenses les plus souvent observés sont les niveaux très faibles ; les moins souvent observés sont les

(1) La définition de ces quatre catégories est donnée avec plus de précisions en annexe. Cette définition renvoie à la nomenclature de biens et services employée dans le compte-rendu de l'enquête de 1956 (*Consommation*, nos 1 et 2 de 1960). En se reportant à ce compte rendu, on peut donc voir quel est approximativement le poids des différentes dépenses rassemblées ici dans une même catégorie.

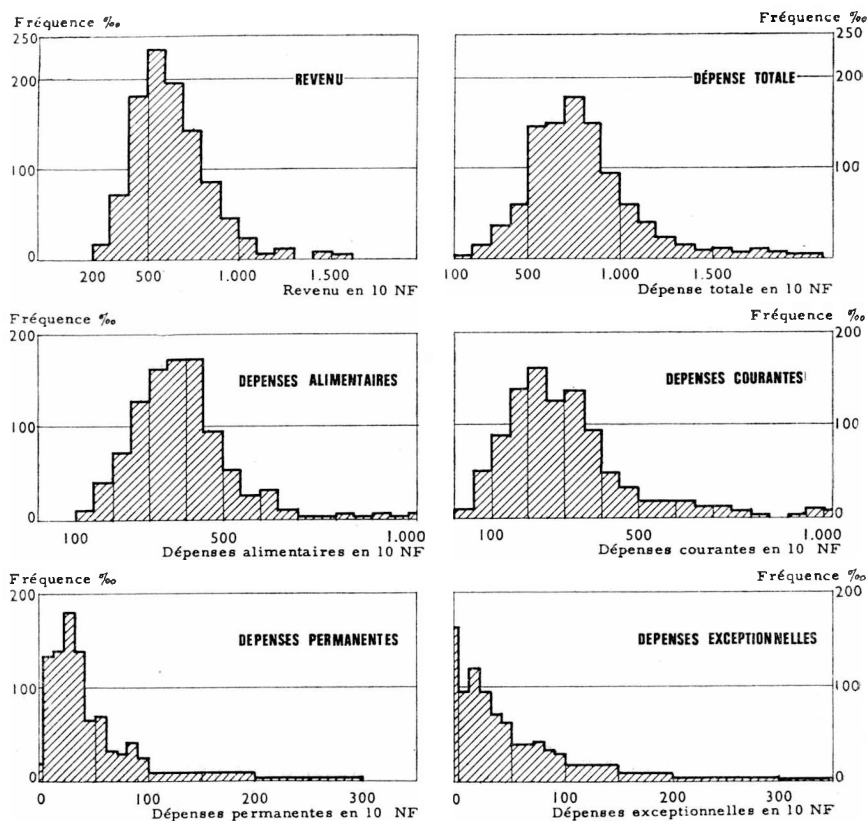
(2) L'échantillon et ce classement sont définis en annexe. Voir aussi dans l'article précédent, page 7, tableau 1, la répartition des 983 ménages selon la catégorie socio-professionnelle du chef de ménage et le type de famille.

(3) 10 sous-groupes de 90 ménages et un de 83 ménages.

niveaux supérieurs à la moyenne. Les graphiques I illustrent ces remarques ; ils présentent ces distributions pour les 522 ménages ouvriers

GRAPHIQUES I

Distribution des ménages ouvriers (522 ménages)



de notre échantillon ⁽¹⁾. On peut admettre que les distributions sont **approximativement** gaussiennes pour les dépenses alimentaires et les dépenses courantes, mais cette hypothèse doit être absolument rejetée pour les deux autres catégories ⁽²⁾. En conséquence dans le premier cas seulement,

(1) On a fait figurer aussi la distribution des revenus et celle des dépenses totales afin de permettre des comparaisons. Les 522 ménages considérés proviennent de la réunion des six sous-groupes d'ouvriers, un par type de famille ; ces sous-groupes seront habituellement distingués dans la suite de l'article.

(2) D'après les graphiques I, la transformation des variables en leur logarithme serait recommandée pour les dépenses alimentaires et les dépenses courantes ; en effet ces distributions sont plus proches de la forme log-normale que de la forme normale. Mais cette transformation de variables n'aurait aucun avantage pour les charges permanentes et les dépenses exceptionnelles : leurs distributions resteraient aussi dissymétriques. Nous ne ferons aucune transformation de variables dans cet article.

moyenne et écart-type suffisent à la rigueur pour décrire entièrement les distributions; ils permettent par exemple de savoir qu'environ 95 % des ménages d'un sous-groupe ont une dépense alimentaire comprise entre deux valeurs, égales à la dépense moyenne du sous-groupe, plus ou moins deux écarts-types. Il n'en va plus de même pour les charges permanentes et les dépenses exceptionnelles. On saisit ici une des limites de l'étude des variances : elle gagnerait à être prolongée par un examen de la forme analytique des distributions et des processus aléatoires susceptibles d'en rendre compte.

TABLEAU I

Fréquence (q) des ménages ayant déclaré une dépense nulle et revenus moyens

(Les ménages sont rangés en 11 sous-groupes d'effectif égal, dans l'ordre croissant des revenus déclarés)

N° DU SOUS- GROUPE	REVENU MOYEN NF/an	q		N° DU SOUS- GROUPE	REVENU MOYEN NF/an	q	
		Dépenses permanentes	Dépenses exceptionnelles			Dépenses permanentes	Dépenses exceptionnelles
1	368	0,02	0,31	7	731	0,02	0,21
2	467	0,03	0,19	8	806	0,01	0,13
3	518	0,01	0,19	9	902	—	0,10
4	568	0,01	0,12	10	1 071	—	0,08
5	617	—	0,12	11*	1 731	—	0,12
6	667	0,01	0,18				

* Le revenu moyen de ce sous-groupe est estimé après exclusion de deux valeurs extrêmes.

Un certain nombre de ménages n'ont déclaré aucune dépense correspondant aux charges permanentes ou aux achats exceptionnels. On voit sur les graphiques 1 que leur fréquence est respectivement de 2 et 16 % dans le groupe des ouvriers. Or toutes les variances considérées dans cet article se rapportent à l'ensemble des ménages, qu'ils présentent ou non dans l'enquête une dépense nulle ⁽¹⁾. Cette variance est un peu plus faible que celle que l'on obtiendrait pour les seuls ménages qui ne présentent pas une dépense nulle ⁽²⁾. La réduction est très modeste cependant. Le calcul

(1) Nous ne pouvions pas en effet, sans perdre toute possibilité de confrontation commode, définir plusieurs populations, se recoupant en partie, selon l'absence ou la présence d'une dépense de tel type dans le questionnaire rempli pour chaque ménage.

(2) Cela tient aux caractères particuliers des distributions des charges permanentes et des dépenses exceptionnelles. Pour d'autres distributions on pourrait avoir le résultat inverse. De façon plus précise, si nous notons V la variance pour l'ensemble des ménages (dépenses nulles et dépenses non nulles), p la proportion des ménages qui ont déclaré une dépense non nulle, m la moyenne de ces dépenses non nulles et v la variance des dépenses non nulles on voit que :

$$V = p \cdot v + p(1 - p)m^2$$

c'est-à-dire que la variance pour l'ensemble des ménages (V) est plus petite que la variance pour les seuls ménages à dépenses non nulles à condition que :

$$V/m^2 > p.$$

montre qu'elle est d'un ou deux millièmes pour les charges permanentes et de 5 % environ pour les dépenses exceptionnelles. Ces pourcentages de réduction sont un peu plus forts dans les sous-groupes où la fréquence des dépenses nulles est la plus élevée, un peu plus faibles dans les autres ; mais au total il n'y a là aucune cause de différence systématique des variances entre les sous-groupes.

Cela tient largement au fait que la fréquence des dépenses nulles n'est elle-même l'objet d'aucune variation régulière, ni avec le revenu moyen des sous-groupes (voir tableau 1), ni avec la catégorie socio-professionnelle ou le type de famille ⁽¹⁾. Nous poursuivons donc l'étude des variances en enregistrant le fait que les dépenses nulles n'influent pratiquement pas sur leur valeur ⁽²⁾.

3. LA MÉTHODE UTILISÉE POUR L'EXAMEN DES VARIATIONS SYSTÉMATIQUES DE LA DISPERSION

Le tableau 2 fournit l'ensemble des données sur lesquelles se fonde la suite de cette étude. Il montre que la variance estimée pour une même catégorie de dépenses diffère beaucoup entre les sous-groupes de ménages ; la variance de la dépense alimentaire par exemple prend des valeurs comprises entre 21 et 381, celle de la dépense courante des valeurs comprises entre 92 et 8 649 ⁽³⁾.

Si l'on considère non plus les variances, mais les coefficients de variation, ces différences deviennent plus faibles mais elles ne s'annulent pas. Il semble en particulier que les coefficients de variation croissent avec les moyennes (voir le tableau 2, colonnes 15 à 18). La constance des coefficients de variation d'une dépense entre les sous-groupes signifierait que les variances de cette dépense sont proportionnelles à la dépense moyenne dans chaque sous-groupe ⁽⁴⁾ ; ou, en termes plus précis, que l'action sur le niveau des dépenses des facteurs tels que le revenu, la composition des familles... se traduit non seulement par une dispersion des moyennes, mais aussi par une dispersion équivalente des variances. Si les coefficients de variation croissent vraiment avec les moyennes, on doit penser que l'action de ces facteurs se répercute plus fortement sur les variances que sur les moyennes. On peut donc espérer la déceler, et se mettre sur la voie d'une interprétation économique des différences observées entre les

(1) Nous ne publions pas ces fréquences observées dans les trente sous-groupes classés par C.S.P. et types de familles. L'absence de toute variation systématique y est encore plus nette que dans le classement selon les revenus qui figurent au tableau 1.

(2) Les réflexions qui peuvent être faites à propos de la distribution des dépenses et de la fréquence des dépenses nulles sont limitées par le regroupement en quatre catégories. Chaque catégorie rassemble des consommations qui correspondent à des dépenses d'importance et de fréquences inégales. On songe par exemple, dans la dépense courante, aux achats de costumes ou de pardessus d'une part et aux achats de mercerie d'autre part.

(3) Le tableau 2 fournit aussi les moyennes, variances et coefficients de variation des revenus et de la dépense totale afin de permettre des comparaisons. **Toutes les variances, dans le texte comme dans les tableaux sont en (100 NF)².**

(4) Ce type d'hétéroscédasticité a été observé plusieurs fois. On en trouve un exemple typique dans : PRAIS et HOUTHAKKER, *The analysis of family budgets* ; Cambridge, 1955, page 56.

TABLEAU 2

Moyennes, variances et coefficients
Pour la définition des sous-groupes de ménages,

SOUS-GROUPE		NOMBRE D'OBSER- VATIONS	MOYENNES EN 10 NF						VARIANCES	
C.S.P.	Ty		Revenu	Dépense totale	Dépenses alimentaires	Dépenses courantes	Dépenses permanentes	Dépenses exceptionnelles	Revenu	Dépense totale
			1	2	3	4	5	6	7	8
1	1	12	987	1 468	377	575	120	394	1 241	10 756
	2	14	1 574	1 742	485	880	227	149	24 607	19 082
	2**	13	1 184	1 399	455	653	173	118	3 535	2 806
	3	8	1 768	1 943	461	929	235	318	26 231	16 251
	4	9	1 128	1 127	424	468	88	146	1 433	1 270
	5	10	1 320	1 447	461	656	164	165	4 236	2 675
	6	26	1 400	1 674	560	699	177	238	3 141	3 440
2	1	11	770	1 038	326	335	284	94	832	1 003
	2	28	765	1 034	419	402	120	93	1 166	2 674
	3	10	1 471	996	342	375	95	184	52 647	1 586
	3**	9	747	902	342	326	66	166	328	783
	4	12	830	1 011	397	392	130	92	802	1 889
	5	14	1 272	1 255	428	494	217	117	13 347	4 898
	6	15	1 079	1 522	525	698	188	111	4 198	7 963
3	1	44	781	1 004	360	450	75	119	1 002	2 337
	2	64	857	1 053	378	446	79	150	3 970	3 808
	3	26	620	839	358	342	47	92	222	640
	4	33	783	1 070	435	441	80	111	908	2 235
	5	29	799	1 192	422	506	103	160	919	3 892
	6	47	897	1 147	473	448	145	81	898	2 400
4	1	99	571	740	316	290	50	84	384	935
	2	99	569	716	357	269	36	53	351	859
	3	59	616	786	363	300	54	69	314	1 094
	4	49	603	764	392	281	41	50	204	555
	5	59	661	867	402	345	51	68	481	997
	6	157	724	869	445	321	52	51	364	861
5	1	7	446	569	282	170	90	27	59	139
	2	11	754	943	387	374	67	114	620	1 673
	3	5	764	672	260	217	88	107	157	210
	4	9	631	834	352	358	53	71	103	476
	5	7	712	842	387	357	63	35	109	56
	6	10	910	1 071	482	392	108	90	615	224

* Le coefficient de variation est défini à l'aide du rapport de l'écart-type à la moyenne $\times 100$.

** Même sous-échantillon que le précédent, mais après exclusion d'une observation extrême pour le revenu.

de variation du revenu et des dépenses

TABLEAU 2

voir dans l'article précédent le tableau 1, p. 7.

EN (100 NF) ²				COEFFICIENTS DE VARIATION*					
Dépenses alimentaires	Dépenses courantes	Dépenses permanentes	Dépenses exceptionnelles	Revenu	Dépense totale	Dépenses alimentaires	Dépenses courantes	Dépenses permanentes	Dépenses exceptionnelles
9	10	11	12	13	14	15	16	17	18
105	2 597	107	4 613	36	71	27	88	86	172
292	8 649	662	248	100	79	35	106	113	105
176	1 560	264	121	50	38	29	60	94	93
167	3 924	933	1 631	92	66	28	67	130	127
244	565	27	148	34	32	37	51	59	83
190	1 020	271	376	49	36	30	49	100	117
234	956	650	431	40	35	27	44	144	87
56	288	641	141	37	30	23	51	89	127
381	569	186	121	45	50	47	59	114	118
123	344	102	456	156	40	32	50	106	116
138	136	15	476	24	31	34	36	59	130
119	320	48	362	34	43	27	46	53	206
81	365	1 665	267	91	56	21	39	188	140
299	3 627	410	429	60	59	33	86	108	186
161	844	52	294	41	48	35	65	96	144
128	1 025	89	813	74	59	30	72	119	191
75	309	9	150	24	30	24	51	63	133
276	1 013	73	267	39	44	38	72	106	148
138	1 214	91	632	38	52	28	69	92	157
214	606	910	147	33	43	31	55	207	150
71	707	24	113	34	41	27	92	99	126
148	256	17	73	33	41	34	59	115	160
118	408	32	68	29	42	30	67	105	119
101	181	16	28	24	31	26	48	98	106
122	435	36	64	33	36	27	60	119	118
253	247	48	37	26	34	36	49	133	120
81	62	39	8	17	21	32	46	69	107
152	345	37	282	33	43	32	50	90	147
45	92	26	49	16	22	26	44	58	66
115	211	11	134	16	26	30	41	64	164
21	124	49	37	15	9	12	31	111	173
144	152	55	175	27	14	25	31	69	147

variances d'un même type de biens dans différents groupes sociaux ⁽¹⁾. Pour procéder à cet examen qui est au cœur de notre sujet il est utile de recourir maintenant à une nouvelle définition de la dispersion des dépenses.

Élimination de l'influence du revenu dans chaque sous-groupe

Il existe une liaison statistique assez faible mais significative entre les consommations et le revenu ⁽²⁾. A l'intérieur de chaque sous-groupe, cette relation a une influence sur la dispersion des dépenses : toutes choses égales par ailleurs, plus grands sont les écarts du revenu entre les ménages et forte l'action du revenu sur la consommation, plus la variance de cette dernière est élevée. Or cette relation peut différer selon les sous-groupes. Elle est donc susceptible d'entraîner une distinction des variances entre les sous-groupes ; l'étude d'une distinction de cette sorte n'a pas d'intérêt en elle-même ; elle renvoie à l'analyse des relations entre les dépenses et le revenu dans chaque sous-groupe. Nous ne considérons plus par conséquent la dispersion totale des dépenses, mais cette dispersion une fois éliminée la part qui est due à la relation dépenses-revenu. Nous faisons cette élimination à l'intérieur de tous les sous-groupes. Précisons rapidement comment elle s'effectue ainsi que l'hypothèse particulière que nous devons introduire pour donner un sens concret à la nouvelle mesure de la dispersion à laquelle elle conduit.

Nous supposons d'abord que la liaison statistique entre chaque catégorie de dépenses et le revenu est de forme linéaire ⁽³⁾ :

$$X_{ik} = \alpha_k Y_{ik} + \beta_k + U_{ik}$$

où X_{ik} représente la dépense du ménage i , appartenant au groupe k , dans l'une des quatre catégories de biens ; Y_{ik} le revenu de ce même

(1) Le tableau 2 invite aussi à faire quelques comparaisons entre les variances des quatre dépenses : a) les montants moyens des dépenses alimentaires et des dépenses courantes sont d'environ 4 000 NF par an et par ménage, tandis que ceux des charges permanentes et des dépenses exceptionnelles leur sont très inférieurs ; en effet, ils dépassent rarement 2 000 NF. Ces différentes positions sont importantes si l'on veut comparer les variances des quatre catégories de biens puisque la variance est un nombre qui dépend des dimensions du phénomène auquel il se rapporte.

b) Pour faciliter ces comparaisons, il est plus commode cependant de se servir des nombres sans dimension que sont les coefficients de variation. On constate sur le tableau 2 que les coefficients de variation des quatre catégories de dépenses se classent de la même façon dans la plupart des trente sous-groupes : le plus faible est généralement celui des dépenses alimentaires avec une valeur de 30 % environ ; vient ensuite le coefficient des dépenses courantes (100 %) environ, enfin celui des dépenses exceptionnelles (120 % environ). Les variations du niveau des charges permanentes (loyers, frais d'éducation...), presque aussi importantes que celles des dépenses exceptionnelles, méritent d'être remarquées. Des coefficients de variation, relatifs aussi à quelques sous-groupes seulement de l'échantillon total, ont été calculés sur l'enquête de 1951 effectuée aux U.S.A. Il est intéressant de les comparer avec ceux que nous publions. On les trouve en particulier dans le tableau 5 (p. 88) du compte rendu méthodologique de cette enquête : *Study of consumer expenditures incomes and savings. Methodology of the survey of consumer expenditures in 1950*. H. Humes Lamale, University of Pennsylvania, 1959.

c) Dans presque tous les sous-groupes, la variance de chaque dépense est inférieure à celle du revenu. Cette constatation infirme une idée simple à laquelle on aurait pu songer pour expliquer que la variance de la dépense totale dépasse largement dans la plupart des cas celle du revenu. Selon cette idée, les dépenses relatives à une catégorie de consommations seraient considérées dans les budgets comme un solde, une fois les autres dépenses décidées ; ainsi la dispersion dans cette catégorie serait forcément plus grande que dans les revenus puisqu'elle accumulerait les effets de la dispersion des revenus et ceux des autres dépenses. Aucune des catégories que nous avons distinguées n'est systématiquement l'objet d'une conduite de ce genre.

(2) Dans le même numéro, l'article de J. Cramer analyse cette liaison dans le cas des dépenses alimentaires. Pour les trois autres catégories nous en avons vérifié l'existence par des tests classiques.

(3) Nous ne nous posons pas le problème de savoir si cette liaison a une signification au regard de la théorie de la demande du consommateur. Au point de vue où nous nous plaçons, il suffit que cette liaison soit empiriquement constatée pour qu'il soit utile d'en tenir compte.

ménage et U_{ik} la distance aléatoire entre la dépense du ménage et le montant :

$$\alpha_k Y_{ik} + \beta_k$$

considéré par le modèle comme une estimation de la dépense des ménages de même revenu. U_{ik} est une variable aléatoire caractéristique du ménage i . Or, si nous observons la dépense de ce ménage à un moment donné, nous n'obtenons de cet aléatoire qu'une valeur particulière qui ne dit rien sur la distribution dont elle est issue. Nous introduisons donc une nouvelle hypothèse : **la distribution des résidus U_{ik} est la même pour tous les ménages du groupe k** . Nous admettons cette hypothèse parce que nous travaillons sur des sous-groupes assez étroitement définis, à l'intérieur desquels en particulier le champ de variation des revenus est généralement restreint ⁽¹⁾.

Moyennant cette hypothèse, si nous observons les N ménages qui composent le groupe k , nous obtenons N valeurs de la même variable aléatoire ($u_{1k}, u_{2k}, \dots, u_{ik}, \dots, u_{Nk}$) et disposons ainsi d'une estimation de sa loi de répartition. En particulier, et c'est seulement ce point qui nous intéresse ici, la variance des N valeurs permet d'estimer la variance de cette loi ou la dispersion des dépenses des ménages du groupe k , une fois éliminé l'effet des différences de revenu à l'intérieur de ce groupe.

Cette estimation est celle de la variance résiduelle ; pour le sous-groupe k , elle est égale à :

$$s_k^2 = \frac{1}{n-2} \sum_{i=1}^n (X_{ik} - \hat{\alpha}_k Y_{ik} - \hat{\beta})^2$$

où $\hat{\alpha}_k$ et $\hat{\beta}_k$ sont les estimations de α_k et β_k obtenues par la méthode des moindres carrés et n l'effectif du sous-groupe k dans l'échantillon.

La comparaison des variances résiduelles entre les sous-groupes

Après cette élimination de l'influence du revenu à l'intérieur de chaque sous-groupe, les différences que l'on peut observer entre les sous-groupes dans la dispersion d'une même catégorie de dépenses, si elles sont systématiques, traduisent les effets d'un ou plusieurs des traits distinctifs de ces sous-groupes :

- la catégorie socio-professionnelle,
- le type de famille,
- et le revenu moyen puisque lui aussi diffère d'un sous-groupe à l'autre ⁽²⁾.

Nous commençons par l'examen de l'influence des deux premières caractéristiques. Si la variance résiduelle, pour une catégorie donnée de

(1) Outre la simplification des calculs, nous invoquons ces mêmes raisons pour justifier l'ajustement linéaire de la relation dépenses-revenu.

(2) Ce sont là les traits distinctifs que la présentation des données permet de repérer. La distinction selon la catégorie de communes serait pratiquement sans portée ici puisque l'échantillon comprend uniquement des urbains de province.

consommations, dépend de la catégorie socio-professionnelle et ne dépend pas du type de famille, tous les types de familles de la même catégorie socio-professionnelle doivent avoir à peu près la même variance résiduelle ; les variances qui doivent être comparées pour examiner la validité de cette hypothèse sont donc celles qui figurent sur une même **colonne** du schéma suivant ; on doit répéter 5 fois cette comparaison, une fois pour chaque catégorie socio-professionnelle. Pour examiner l'hypothèse inverse (effet du type de famille et absence d'effet de la C.S.P.), on doit évidemment comparer les variances qui figurent sur une même **ligne** du schéma, et cette comparaison doit être répétée 6 fois, une fois pour chaque type de famille.

SCHÉMAS

**pour les tests d'homogénéité des variances résiduelles
d'une dépense quelconque selon la catégorie sociale et le type de famille**

Ty \ CS	1	2	3	4	5
1					
4					
6					

Pour tester l'homogénéité des C.S.
à l'intérieur du type 4
comparer les variances résiduelles
des cinq sous groupes de cette ligne

Ty \ CS	1	3	5
1			
2			
3			
4			
5			
6			

Pour tester l'homogénéité des Ty
à l'intérieur de la C.S. 3
comparer les variances résiduelles
des 6 sous groupes de cette colonne

Pour effectuer les comparaisons, il faut tenir compte du caractère aléatoire des estimations de la variance résiduelle, et donc recourir à un test statistique. Nous en exposons rapidement le principe en supposant que nous voulons savoir si la variance résiduelle de la dépense courante peut être considérée comme identique dans les six groupes de ménages ouvriers. A l'aide des six estimations de la variance résiduelle dont nous disposons (une pour chaque type de famille), nous désirons choisir entre deux hypothèses : l'hypothèse H_1 , selon laquelle les échantillons sont issus de populations dont la variance résiduelle est quelconque, et l'hypothèse alternative et plus restrictive H_0 , pour laquelle ces variances sont identiques. Pour faire ce choix on compare la vraisemblance d'obtenir les six estimations que nous avons si H_1 est vraie, avec la vraisemblance des mêmes six valeurs dans le cas où ce n'est plus H_1 mais H_0 qui est vraie. Les six estimations étant affectées d'erreurs d'échantillonnage, le rapport des deux vraisemblances est lui aussi aléatoire ; mais on sait que si H_0 est vraie, le logarithme du rapport des deux vraisemblances est, à une constante près, distribué approximativement comme la variable χ^2 avec $m - 1$ degrés de liberté, m étant le nombre des échantillons considérés (dans notre

exemple $m = 6$) ⁽¹⁾. Si la valeur que nous obtenons pour ce rapport peut être considérée comme appartenant à la distribution de χ^2 ainsi définie, l'hypothèse H_0 est admissible.

Le test que nous venons de décrire n'est applicable en toute rigueur que pour comparer les variances de variables gaussiennes. Les variables en cause sont ici les écarts :

$$u_{ik} = X_{ik} - \hat{\alpha}_k Y_{ik} - \hat{\beta}_k$$

La relation qui les définit montre que nous avons le droit de considérer que ces écarts sont gaussiens si X et Y le sont. Cette condition est à peu près satisfaite pour les écarts relatifs à la dépense alimentaire et à la dépense courante. En effet comme nous l'avons dit plus haut, le revenu (Y) et ces deux dépenses obéissent à des lois qui sont proches de la loi de Laplace-Gauss ⁽²⁾. Mais il en va différemment pour les dépenses permanentes et les dépenses exceptionnelles : la dissymétrie de leurs distributions nous interdit de supposer que leurs écarts U_{ik} sont gaussiens. Nous ne renonçons pas pourtant à appliquer le test aux variances résiduelles de ces deux dépenses puisqu'on ne dispose pas de méthode simple pour procéder autrement ; mais les indications que ce test peut nous fournir sont moins décisives ; l'interprétation de ses résultats comportera donc une part d'arbitraire.

Les valeurs extrêmes

Nous devons faire une dernière considération de méthode. Elle tient à la présence de valeurs extrêmes dans quelques sous-groupes. Celles-ci affectent la comparaison des variances. Dans l'estimation de la variance en effet l'écart entre les observations et la moyenne intervenant au carré, une seule valeur extrême est susceptible de doubler ou de tripler le résultat.

(1) La quantité distribuée comme χ^2 avec $m - 1$ degrés de liberté est :

$$L = \frac{1}{C} \log_e \left\{ \frac{(S^2)^v}{(s_1^2)^{v_1} + \dots + (s_k^2)^{v_k} + \dots + (s_m^2)^{v_m}} \right\}$$

avec

$$C = 1 + \frac{1}{3(m-1)} \left\{ \sum_{k=1}^m \frac{1}{v_k} - \frac{1}{v} \right\}$$

v_k = effectif du groupe k , moins un

$$v = \sum_k v_k \quad \text{et} \quad S^2 = \frac{1}{v} \sum_k v_k s_k^2$$

Pour un exposé plus complet de ce test, voir : E. MORICE et F. CHARTIER, *Méthode Statistique*, Paris, 1956. Deuxième partie, p. 236.

(2) Voir plus haut, p. 39.

TABLEAU 3

**Nombre des observations extrêmes
dans chaque catégorie socio-professionnelle**

Catégorie socio-professionnelle	Dépenses courantes	Dépenses permanentes	Dépenses exceptionnelles
Gros indépendants.....	4	2	3
Petits indépendants.....	5	3	1
Cadres moyens.....	14	1	3
Ouvriers.....	9	—	1
Autres actifs.....	—	—	—
Ensemble.....	32	6	8

Par convention, nous considérons comme extrêmes toutes les observations qui présentent à la fois les deux caractéristiques suivantes :

- la dépense étudiée est une valeur nettement extérieure à la distribution observée dans le même sous-groupe ;
- et cette dépense dépasse à elle seule le revenu déclaré.

Aucun ménage ne répond à ces deux critères en ce qui concerne la dépense alimentaire. Par contre, le procédé conduit à exclure 32 ménages pour les calculs relatifs aux variances résiduelles des dépenses courantes, 6 pour les charges permanentes et 8 pour les dépenses exceptionnelles. Le tableau 3 indique comment ces ménages se répartissent dans les catégories socio-professionnelles. La présence de ces valeurs extrêmes ne met pas forcément en cause la qualité des observations effectuées lors de l'enquête. Il y a à cela au moins deux raisons. D'une part des valeurs extrêmes existent dans toute distribution et rien ne s'oppose à ce que les ménages financent parfois une dépense en faisant appel à leurs capitaux. D'autre part, un assez grand nombre de dépenses ont été relevées dans l'enquête sur une période inférieure à l'année et multipliée ensuite par un coefficient convenable pour obtenir une estimation de la dépense annuelle⁽¹⁾. Il n'y a rien d'étonnant à ce qu'une dépense relevée sur un mois ou un trimestre et multipliée ensuite par 12 ou par 4 dépasse le revenu déclaré sur un an. Mais, qu'elle que soit l'interprétation que l'on donne de ces valeurs extrêmes, nous pensons qu'il est utile d'apprécier leur incidence sur les résultats de l'analyse. Le tableau 4 fournit par conséquent chaque fois qu'il y a lieu une double estimation de la variance résiduelle.

(1) Cette procédure affecte naturellement les estimations de la variance des dépenses. Nous y reviendrons dans la conclusion.

TABLEAU 4

Estimations des variances résiduelles des dépenses en (100 NF)²

SOUS-GROUPE		NOMBRE D'OBSERVATIONS	DÉPENSE TOTALE	DÉPENSES ALIMENTAIRES	DÉPENSES COURANTES *	DÉPENSES PERMANENTES *	DÉPENSES EXCEPTIONNELLES *
C.S.P.	TY						
1	1	12	10 819	54	212 (2 266)	117	586 (5 074)
	2	14	1 750	167	964 (946)	194 (179)	123 (162)
	3	8	4 647	80	1 079	152	1 536
	4	9	730	171	364	21	157
	5	10	1 405	101	395	296	400
	6	26	3 201	236	724	283 (677)	448
2	1	11	1 008	52	320	250 (707)	123
	2	28	1 718	217	407 (431)	158	126
	3	10	787	138	142 (146)	11 (12)	542 (478)
	4	12	1 278	107	252	35	315
	5	14	1 825	88	202 (206)	88 (653)	171
	6	15	7 669	290	1 196 (3 805)	373	425
3	1	44	1 700	158	160 (655)	48	255
	2	64	1 792	95	319 (414)	40	354 (770)
	3	26	564	74	219 (275)	9	156
	4	33	1 796	220	404 (947)	55	88 (276)
	5	29	2 907	99	332 (964)	85	520 (610)
	6	47	1 644	168	386 (513)	94 (899)	122
4	1	99	895	70	141 (692)	25	112
	2	99	634	132	117 (205)	14	54 (70)
	3	59	773	113	198 (301)	31	54
	4	49	387	86	156	14	24
	5	59	666	112	267 (333)	34	56
	6	157	618	224	192	42	37
5	1	7	156	63	54	24	10
	2	11	1 231	167	309	36	152
	3	5	42	54	17	26	33
	4	9	514	125	240	13	101
	5	7	67	20	118	58	43
	6	10	233	98	169	43	112

* On a indiqué entre parenthèses la valeur de l'estimation de la variance résiduelle avant exclusion des valeurs extrêmes. Cette indication fait naturellement défaut chaque fois qu'il n'y a pas de valeur extrême.

4. RÉSULTATS : LES FACTEURS QUI INFLUENT SUR LA DISPERSION DES DÉPENSES

Sur le tableau 5 figurent les résultats du test présenté plus haut : avant exclusion des valeurs extrêmes, dans sa partie supérieure ; après exclusion, dans sa partie inférieure. Pour chaque catégorie de dépenses, on trouve dans ce tableau trois indications : la valeur du rapport des deux vraisemblances (notée L) estimée d'après l'échantillon ; le résultat du test ; enfin l'estimation de la variance résiduelle commune (notée S^2) qui doit être retenue si le test conclut à l'homogénéité. Le test a été appliqué pour apprécier l'homogénéité d'abord entre les types de famille (5 premières lignes), ensuite entre les catégories socio-professionnelles (6 dernières lignes).

Nous examinons d'abord les résultats relatifs à la dépense alimentaire pour laquelle aucune observation n'a dû être exclue.

La dispersion de la dépense alimentaire et l'effet du type de famille

L'hypothèse d'homogénéité est acceptable entre les catégories socio-professionnelles d'un même type de famille (5 fois au seuil de 5 % et une fois au seuil de 1 %), par contre elle doit être rejetée entre les types de familles ⁽¹⁾. **La variance résiduelle de la dépense alimentaire diffère donc selon les types de familles mais elle ne dépend pas de la catégorie socio-professionnelle.** En outre on ne peut pas considérer que l'influence du type de famille sur ces variances soit le reflet des différences de revenus moyens qui existent entre ces divers types. En effet s'il y avait une action du revenu elle rendrait les variances plus sensibles aux catégories socio-professionnelles qu'aux types de familles, les écarts entre les revenus moyens étant beaucoup plus importants dans la première classification que dans la seconde ⁽²⁾. Pour chaque type de famille nous retenons donc les valeurs S^2 (colonne 5 et six dernières lignes du tableau 5) comme estimations de la variance résiduelle de la dépense alimentaire.

D'après ces estimations, la variance résiduelle diffère beaucoup entre les types de familles ; la plus faible est égale à 91, la plus forte à 214. On se rappelle que toutes les familles de l'échantillon sont composées de deux adultes, avec un ou plusieurs enfants ; les traits qui distinguent les types de familles sont donc le nombre et l'âge des enfants. De façon générale les estimations de la variance résiduelle tendent à croître avec ces deux caractères ; il semble en outre que la dispersion de l'âge des enfants (et pas seulement l'âge de chacun) augmente la dispersion résiduelle des dépenses alimentaires. Ces constatations rejoignent le sens commun. On conçoit en effet que la diversité des conduites à l'égard des dépenses alimentaires soit plus grande entre les familles nombreuses qu'entre les ménages de deux personnes avec un seul enfant. Ce phénomène n'est du reste pas indépendant du niveau des dépenses alimentaires par ménage, naturellement plus élevé dans le premier cas que dans le second. En outre si les dispersions à revenu constant grandissent avec la taille de la famille pour l'alimentation, on s'attend à ce que le mouvement inverse se produise

(1) Ce rejet apparaît déjà dans les 5 premières lignes du tableau 5 (voir surtout 5^e ligne, C.S.P. ouvriers). Au reste, accepter à la fois l'homogénéité entre les types de familles et entre les C.S.P. reviendrait à accepter l'homogénéité des 30 sous-groupes à la fois. Nous avons vérifié que cette hypothèse est inadmissible ; le rapport L obtenu sur les 30 sous-groupes à la fois dépasse très largement la valeur du χ^2 correspondant au seuil de confiance de 1 %.

(2) Pour ce dernier point nous renvoyons le lecteur au tableau 2 dans lequel sont indiqués les revenus moyens des ménages classés selon la C.S.P. et le type de famille.

en d'autres domaines de dépenses, les familles nombreuses ayant alors moins de marge de liberté que les autres. C'est en effet ce que suggérera plus loin l'examen des variances résiduelles des dépenses exceptionnelles.

La dispersion dans les trois autres catégories de dépenses et l'effet du revenu moyen

Nous examinons maintenant les résultats du test appliqué aux trois autres catégories de dépenses, avant et après exclusion des valeurs extrêmes. **Avant les exclusions**, aucune espèce d'homogénéité n'apparaît pour aucune des trois dépenses : courantes, permanentes et exceptionnelles. Pour ces deux dernières bien sûr, l'interprétation des résultats doit tenir compte du fait que leurs distributions ne sont pas gaussiennes et que la valeur du rapport L peut être majorée uniquement pour cette raison ; mais si l'on compare les rapports L observés à la valeur de χ^2 correspondant au seuil de confiance de 1 %, on s'aperçoit qu'ils atteignent entre deux et quatre fois cette valeur ; même dans la situation particulière où nous nous trouvons, nous jugeons que ce dépassement est trop important pour accepter l'hypothèse d'homogénéité.

La partie inférieure du tableau 5 présente les résultats des tests d'homogénéité effectués sur les variances résiduelles **après exclusion** des valeurs extrêmes. Dans presque tous les cas le rapport L est beaucoup plus faible qu'avant cette élimination. Pour les dépenses courantes, l'homogénéité est maintenant acceptable entre les types de familles de trois catégories socio-professionnelles (C.S.P. 1, 3 et 5). Dans les deux autres catégories cependant (C.S.P. 2 et 4 : petits indépendants et ouvriers), le rapport L dépasse très légèrement le seuil de 5 % ; mais nous pensons que ce fait ne suffit pas pour justifier un traitement spécial de ces deux catégories socio-professionnelles. Nous acceptons l'hypothèse d'une distinction des variances résiduelles des dépenses courantes selon la catégorie socio-professionnelle et nous notons seulement qu'un autre effet intervient peut-être, dont la liaison avec le type de famille est assez ténue puisqu'il ne se manifeste que dans deux catégories socio-professionnelles sur cinq.

Pour les charges permanentes et les dépenses exceptionnelles le test (après exclusion) conclut plus nettement à l'hétérogénéité entre les catégories socio-professionnelles (six dernières lignes) qu'entre les types de familles (cinq premières lignes). Le rapport L est en effet généralement plus grand dans le premier cas que dans le second ; entre les types de famille, L ne dépasse nettement le seuil de 1 % que pour les cadres moyens et les ouvriers. Il y a donc peut être une influence des types de familles sur les variances résiduelles des charges permanentes et des dépenses exceptionnelles, mais cette influence est moins nette que celle que le test attribue aux catégories socio-professionnelles. Pour ces deux dépenses comme pour la dépense courante, le facteur « catégorie socio-professionnelle » semble ainsi jouer un rôle important.

Si l'on examine les estimations des variances résiduelles de ces trois dépenses dans chacune des cinq catégories socio-professionnelles (ce sont les quantités notées S^2 dans les colonnes 8, 11 et 14 du tableau 5, cinq premières lignes), on remarque qu'elles croissent dans le même sens que les revenus moyens des catégories socio-professionnelles. Une hypothèse très simple consisterait donc à admettre que les variances résiduelles ne se distinguent pas selon la catégorie socio-professionnelle mais selon le niveau de revenu du groupe de ménages auquel elles se rapportent. Si cette hypothèse est vraie, on doit observer la même évolution des variances avec les revenus moyens des sous-groupes, que l'on considère des groupes de ménages classés selon leur catégorie socio-professionnelle comme jus-

TABLEAU 5

Tests d'homogénéité des variances résiduelles

(Résultats des tests : n.s. = non significatif ;

		DEGRÉS DE LIBERTÉ	NOMBRE D'OBSER- VATIONS	DÉPENSES ALIMENTAIRES		
				Rapport L	Résultat du test	S ² (100 NF) ²
		1	2	3	4	5
Avant exclusion des valeurs extrêmes	Homogénéité entre les types de famille d'une seule C.S.P.					
	Gros indépendants.	5	79	7,86	n.s.	159
	Petits indépendants.	5	90	10,22	n.s.	168
	Cadres moyens.	5	243	13,99	*	136
	Ouvriers.	5	522	45,90	**	140
	Autres actifs.	5	49	6,23	n.s.	101
	Homogénéité entre les C.S.P. d'un seul type de famille					
	Deux adultes avec :					
	1 enfant de moins de 6 ans	4	173	12,80	*	91
	1 enfant de 6 à 13 ans..	4	216	7,20	n.s.	135
	2 enfants de moins de 6 ans	4	108	2,24	n.s.	102
	1 enfant de moins de 6 ans et un de 6 à 13.	4	112	8,70	n.s.	137
	2 enfants de 6 à 13 ans..	4	119	4,37	n.s.	101
	3 enfants et plus.	4	255	4,00	n.s.	214
Après exclusion des valeurs extrêmes	Homogénéité entre les types de famille d'une seule C.S.P.					
	Gros indépendants.	5				
	Petits indépendants.	5				
	Cadres moyens.	5				
	Ouvriers.	5				
	Autres actifs.	5				
	Homogénéité entre les C.S.P. d'un seul type de famille					
	Deux adultes avec :					
	1 enfant de moins de 6 ans	4				
	1 enfant de 6 à 13 ans..	4				
	2 enfants de moins de 6 ans	4				
	1 enfant de moins 6 ans et un de 6 à 13 ans.	4				
	2 enfants de 6 à 13 ans..	4				
	3 enfants et plus.	4				
				Il n'y a pas de « valeurs extrêmes » parmi les observations individuelles des dépenses alimentaires.		

selon la catégorie socio-professionnelle et le type de famille

TABLEAU 5

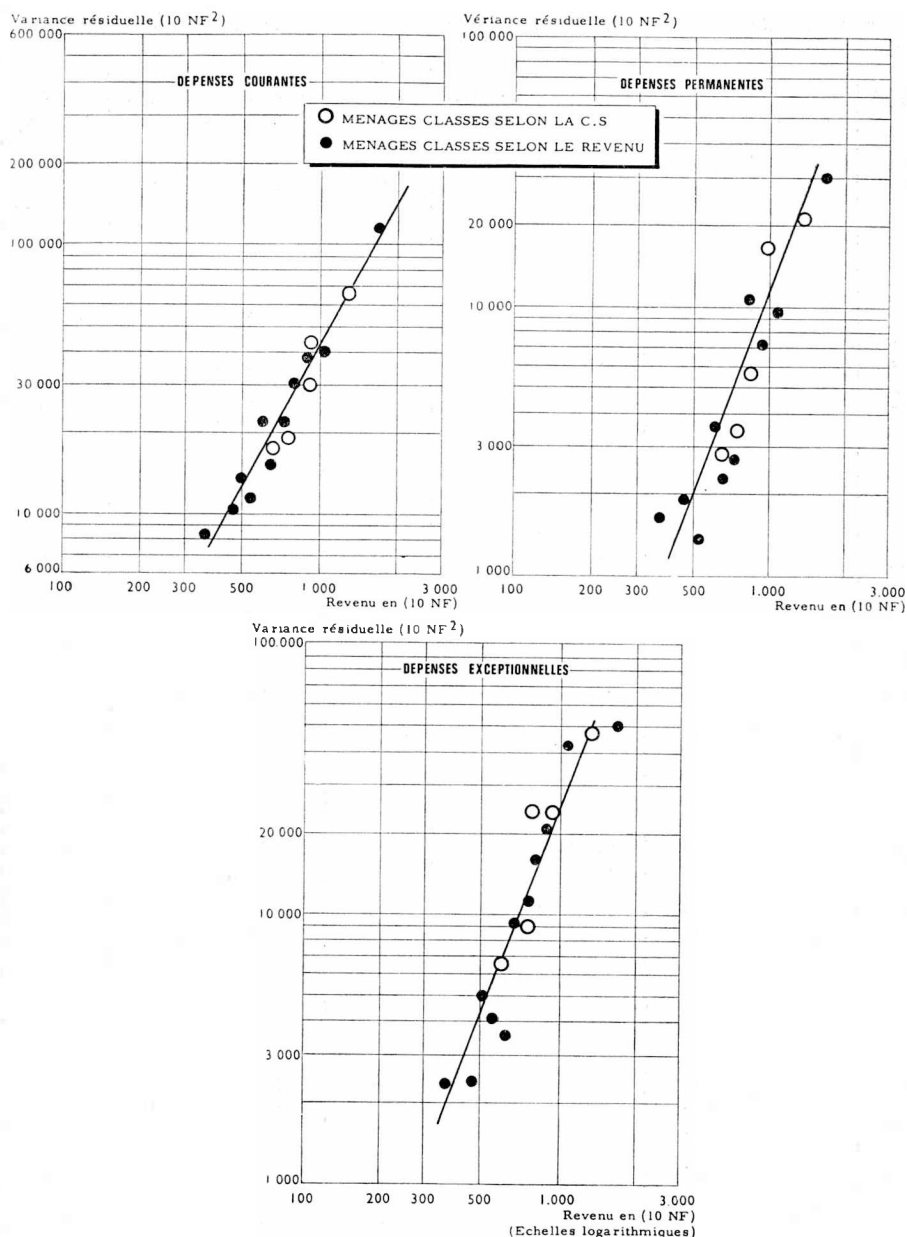
* = significatif à 5% ; * * = significatif à 1%)

DÉPENSES COURANTES			DÉPENSES PERMANENTES			DÉPENSES EXCEPTIONNELLES		
Rapport L	Résultat du test	S ² (100 NF) ²	Rapport L	Résultat du test	S ² (100 NF) ²	Rapport L	Résultat du test	S ² (100 NF) ²
6	7	8	9	10	11	12	13	14
10,33 55,16 17,36 72,81 8,68	n.s * * * * * * n.s	949 894 598 314 182	26,68 43,10 242,14 44,73 3,74	* * * * * * * * n.s	343 302 213 29 34	48,24 11,69 54,53 55,33 9,93	* * * * * * * n.s	1 149 243 401 61 90
20,13 21,31 14,54	* * * * * *	739 344 321	102,04 95,22 27,78	* * * * * *	74 50 31	159,47 114,01 64,60	* * * * * *	451 297 204
31,68 17,77 117,80	* * * * * *	426 470 494	19,46 67,39 225,10	* * * * * *	29 135 279	56,46 60,06 43,53	* * * * * *	143 230 116
7,10 14,21 9,95 15,64 8,68	n.s * n.s * * n.s	654 427 308 174 182	12,64 27,45 37,22 44,73 3,74	* * * * * * * n.s	204 165 56 29 34	15,79 12,39 36,18 53,92 9,93	* * * * * * * n.s	474 245 253 57 90
5,69 41,79 17,76	n.s * * * *	157 270 249	42,70 96,44 27,84	* * * * * *	47 51 31	29,58 65,89 65,61	* * * * * *	170 161 206
8,71 2,82 43,33	n.s n.s * *	258 279 321	19,46 25,85 80,50	* * * * * *	29 73 92	40,78 52,79 43,53	* * * * * *	86 205 116

qu'à présent, ou selon leurs revenus. Les graphiques II vérifient bien cette hypothèse pour les trois catégories de dépenses que nous examinons ⁽¹⁾.

GRAPHIQUES II

Variances résiduelles et revenus moyens



(1) Les sous-groupes de ménages classés dans l'ordre croissant des revenus sont au nombre de 11. Nous avons déjà utilisé cette division de l'échantillon. Mais dans le cas présent nous en avons exclu, avant de calculer les variances résiduelles, les mêmes observations extrêmes que dans les trente sous-groupes classés selon la C.S.P. et le type de famille : 32 observations pour les dépenses courantes, 6 pour les charges permanentes et 8 pour les dépenses exceptionnelles. Les corrélations présentées dans les graphiques II sont très mauvaises si l'on n'effectue pas ces exclusions.

Ces graphiques sont en coordonnées doublement logarithmiques ; parmi les différents ajustements que nous avons essayés celui-ci est le meilleur.

En outre, cette hypothèse est cohérente avec la légère hétérogénéité que les tests décelaient entre les sous-groupes d'une même catégorie socio-professionnelle, ces sous-groupes présentant des revenus moyens différents.

De l'examen des dépenses courantes, permanentes et exceptionnelles nous retenons par conséquent que la croissance de la dispersion résiduelle avec les revenus moyens est nette. Les graphiques II permettent d'apprécier l'intensité de ce phénomène. Ils montrent qu'elle est à peu près la même pour les charges permanentes et les dépenses exceptionnelles ; elle est plus faible pour les dépenses courantes. En utilisant les droites ajustées sur ces graphiques on voit que si les revenus

TABLEAU 6

**Effet du type de famille sur la variance résiduelle
des dépenses alimentaires et la variance résiduelle des dépenses exceptionnelles***

C.S.P.	Ty	DEUX ADULTES avec ... enfant(s)	NOMBRE D'OBSER- VATIONS	REVENU MOYEN (10 NF)	VARIANCES RÉSIDUELLES (100 NF) ²	
					Dépenses alimentaires	Dépenses exceptionnelles
3	1	Un de moins de 6 ans	44	781	154	249
	2	Un de 6 à 13 ans....	93	830	95	410
	5	Deux de 6 à 13 ans .				
	3	Deux de moins de 6 ans.....	26	620	71	150
	4	Un de moins de 6 ans et un de 6 à 13 ans.	80	860	190	105
	6	Trois et plus				
4	1	Un de moins de 6 ans	99	571	70	112
	2	Un de 6 à 13 ans....	217	636	110	53
	5	Deux de 6 à 13 ans .				
	3	Deux de moins de 6 ans.....				
	4	Un de moins de 6 ans et un de 6 à 13 ans.	49	603	84	34
	6	Trois et plus	157	724	223	

* Le tableau ne retient que l'estimation commune lorsque plusieurs sous-groupes ont des variances résiduelles non significativement différentes.

moyens de deux groupes de ménages diffèrent de 50 % par exemple, les variances résiduelles entre ces groupes diffèrent de 100 % pour les dépenses courantes et de plus de 130 % pour les dépenses permanentes et les dépenses exceptionnelles.

Les variances résiduelles des dépenses exceptionnelles présentent en outre une particularité ; un effet du type de familles paraît s'y superposer à celui du revenu. Alors qu'ils disposent de revenus supérieurs en moyenne, les groupes composés des familles les plus nombreuses ont souvent en effet des variances résiduelles plus faibles que les autres. Ceci apparaît du moins dans les types de familles des cadres moyens et des ouvriers (C.S.P. 3 et 4), peut-être parce que nous disposons pour les étudier d'observations plus nombreuses (1). Le tableau 6 fournit les éléments de cette comparaison. Comme nous le pensions plus haut, il y aurait ainsi pour les divers types de familles une compensation entre les dépenses alimentaires et les dépenses exceptionnelles : lorsque la variété des comportements est moins grande dans une de ces catégories de dépenses, elle serait plus accusée dans l'autre.

5. LA VARIANCE RÉSIDUELLE DE LA DÉPENSE TOTALE

Les résultats précédents peuvent être utilisés pour donner une interprétation de la variance résiduelle de la **dépense totale**, et surtout de ses variations entre les sous-groupes de ménages dont on peut voir l'ampleur en se reportant au tableau 4. Auparavant il faut examiner les covariances résiduelles entre les quatre catégories de dépenses puisqu'elles interviennent aussi dans la variance résiduelle de la dépense totale. Celle-ci en effet, selon une formule classique, se définit par :

$$\text{Var}(\Sigma U_i) = \Sigma \text{Var}(U_i) + 2\Sigma \text{Cov}(U_i, U_r)$$

Dans cette formule, U_r et U_i désignent les résidus de l'une des quatre catégories de dépenses (2). Les covariances entre ces résidus, au nombre de six, expriment la liaison entre les dépenses prises deux à deux, une fois l'influence du revenu éliminée. Nous les étudions par l'intermédiaire des coefficients de corrélation partielle qui se prêtent plus facilement à un examen statistique. Ces coefficients sont définis par

$$r_{U_i U_r} = \frac{\text{Cov}(U_i, U_r)}{\sqrt{\text{Var}(U_i) \cdot \text{Var}(U_r)}}$$

Nous disposons de trente estimations de chacun de ces six coefficients, une par sous-groupe de ménages. Les résultats des différents calculs

(1) Afin de simplifier le graphique II c, nous avons renoncé à distinguer les types de familles à l'intérieur des C.S.P. 3 et 4, comme les tests pourtant y engagent. La dépendance des variances résiduelles à l'égard des revenus moyens en est légèrement modifiée.

(2) Cette formule est analogue à celle qui définit la variance totale de la dépense totale. Il suffit d'y remplacer les résidus U_i et U_r par les dépenses elles-mêmes : X_i et X_r .

et tests que nous avons effectués pour parvenir à des estimations sont résumés dans le tableau 7 ⁽¹⁾.

TABLEAU 7

Estimation des coefficients de corrélation partielle $[r(U_i, U_r)]$
entre les quatre catégories de dépenses

$U_i \backslash U_r$	DÉPENSES ALIMENTAIRES	DÉPENSES COURANTES	DÉPENSES PERMANENTES	DÉPENSES EXCEPTIONNELLES
Dépenses alimentaires		0,17	0	ϵ
Dépenses courantes	0,17		0	0,20
Dépenses permanentes....	0	0		0
Dépenses exceptionnelles .	ϵ	0,20	0	

On voit que la corrélation partielle est nulle ou négligeable dans quatre cas. Dans les deux autres, elle est faible mais positive : 0,17 entre les dépenses alimentaires et les dépenses courantes ; 0,20 entre les dépenses courantes et les dépenses exceptionnelles. Ces résultats sont valables pour les trente sous-groupes ; ni la catégorie socio-professionnelle, ni le type de famille n'agissent sur la valeur des corrélations partielles.

Les covariances résiduelles sont donc nulles (ou peuvent être considérées comme telles) sauf deux :

Cov. résid. { Ali., Courant }

$$= 0,17 \sqrt{\text{Var. résid. \{ Ali. \} \cdot Var. résid. \{ Courant \}}}$$

Cov. résid. { Courant, Except. }

$$= 0,20 \sqrt{\text{Var. résid. \{ Courant \} \cdot Var. résid. \{ Except. \}}}$$

Si l'on se rappelle les caractéristiques des variances résiduelles qui figurent aux seconds membres, on voit que ces deux covariances dépendent du niveau des revenus moyens et aussi, surtout pour la première, des types de familles.

La variance résiduelle de la dépense totale est donc composée de termes qui, à l'exception d'un seul : la variance résiduelle de la dépense alimentaire, croissent tous avec les revenus moyens des groupes de ménages⁽²⁾.

(1) Les tests ont porté successivement sur l'hypothèse d'une valeur nulle pour les six corrélations partielles (prises séparément) ; puis, pour les corrélations significativement différentes de zéro, sur l'hypothèse qu'elles ont la même valeur dans les trente sous-groupes de ménages. Les résultats des tests sont les mêmes, que l'on exclue ou non les valeurs extrêmes. Mais la valeur des coefficients de corrélation partielle diffère dans ces deux cas ; elle est légèrement plus élevée après les exclusions. C'est cette seconde estimation que nous avons toujours retenue afin de rester cohérents avec le traitement appliqué aux variances résiduelles.

(2) Du moins peut-on affirmer ceci pour la division de la dépense totale en quatre catégories, utilisée dans cet article.

L'action de ce facteur est prépondérante, celle du type de famille au contraire est très faible. En effet comme nous l'avons souligné plus haut, avec la taille des familles la variance résiduelle tend à croître pour les dépenses alimentaires et à décroître pour les dépenses exceptionnelles ; les deux effets s'annulent à peu près au niveau de la dépense totale.

L'action des revenus moyens sur la variance résiduelle de la dépense totale peut naturellement être observée si l'on examine les estimations de cette variance dans les divers sous-groupes de ménages ⁽¹⁾. Il n'est pas

TABLEAU 8

Composantes de la variance résiduelle de la dépense totale
Valeurs ajustées pour des groupes de différents revenus moyens (*)

		REVENUS MOYENS ANNUELS					
		6 000 NF		9 000 NF		12 000 NF	
		Var ou Cov, en (100 NF) ²	En %	Var ou Cov, en (100 NF) ²	En %	Var ou Cov, en (100 NF) ²	En %
Variances résiduelles des ...	Dépenses ali- mentaires.....	137	27,8	137	15,0	137	9,2
	Dépenses cou- rantes	165	33,9	340	37,4	560	38,0
	Dépenses per- manentes ..	29	6,0	84	9,2	170	11,5
	Dépenses ex- ceptionnelles	65	13,4	180	19,8	350	23,7
Covariances résiduelles entre ... **	Alimentaires et courantes ..	51	10,5	72	7,8	83	5,6
	Courantes et exception. ...	41	8,4	99	10,8	177	12,0
Variance résiduelle de la dé- pense totale		488	100,0	912	100,0	1 477	100,0

* Pour l'explication des valeurs figurant dans ce tableau, se reporter au texte page 59, note 1.
** Les covariances résiduelles sont, dans ce tableau, multipliées par deux pour obtenir directement la valeur de leur contribution à la variance résiduelle de la dépense totale.

(1) Cet examen peut être effectué aussi à l'aide du graphique 5 de l'article précédent. Ce graphique relie le revenu moyen et la variance de l'écart Z entre le revenu et la dépense totale ; mais en raison de l'indépendance que J. CRAMER constate entre le revenu et l'écart Z, la variance de Z est une estimation de la variance résiduelle de la dépense totale. En sens inverse par conséquent, les remarques présentées dans cette section sur la variance résiduelle de la dépense totale s'appliquent aussi bien à la variance de l'écart Z. En particulier on peut juger par le tableau 8 du rôle joué par chaque catégorie de dépenses dans la dispersion de l'écart Z.

inutile cependant de reconstituer cette variance pour des groupes sociaux définis maintenant a priori. Ceci est possible à partir des termes qui définissent cette variance en tenant compte des résultats obtenus pour chacun d'entre eux après les exclusions des valeurs extrêmes. On peut ainsi apprécier de façon plus systématique quel est l'effet des différences de revenu moyen sur la dispersion de la dépense totale et distinguer ses origines selon les diverses dépenses. Le tableau 8 présente cette reconstitution pour trois groupes de ménages distingués par le revenu moyen ⁽¹⁾. On voit que la variance résiduelle de la dépense totale varie de un à trois entre un groupe dont le revenu moyen est faible (6 000 NF par an et par ménage) et un autre groupe dont le revenu moyen est le double du premier (12 000 NF). On constate en outre que, selon les niveaux de revenu, les quatre catégories de dépenses contribuent de façon fort différente à la dispersion résiduelle de la dépense totale. Moins les groupes sont riches, plus les dépenses alimentaires et les dépenses courantes tiennent une place importante dans une variance résiduelle totale qui est faible. Plus ils sont riches, plus large est la contribution des charges permanentes et des dépenses exceptionnelles. Ces « chassés-croisés » rappellent ceux des coefficients budgétaires des différentes dépenses.

6. REMARQUES ET CONCLUSIONS

Les résultats de cet article ont été obtenus pour la plupart grâce à l'exclusion de 32 observations sur les 983 que comprend l'échantillon. Chacun reste libre de penser que nous avons ainsi renoncé à une information qui, pour ne pas se prêter à l'analyse, n'en mérite pas moins de considération. Ce procédé revêt en effet une part d'arbitraire. Mais il nous a permis d'arriver à quelques conclusions cohérentes ; dans le cadre très empirique où nous nous sommes placés, ce fait suffit à le justifier. Au reste, exclure des observations ne préjuge pas forcément de leur validité. Certaines valeurs extrêmes peuvent être parfaitement exactes et dans ce cas les éliminer revient seulement à considérer qu'elles font partie d'un ensemble particulier qu'une information trop étroite empêche d'examiner.

L'étude de différents groupes de ménages a permis d'illustrer la croissance de la dispersion résiduelle des dépenses, commandée surtout par le revenu moyen des groupes. Or on sait que les estimations classiques des paramètres d'un modèle ne sont pas « efficaces » si on considère comme un seul ensemble des groupes de ménages à l'intérieur desquels la dispersion des résidus prend des valeurs différentes ⁽²⁾. En manifestant que ces différences peuvent être très importantes, l'étude précédente souligne par conséquent l'intérêt qu'il y a à se prémunir contre elles pour améliorer les estimations. Les précautions courantes consistent soit à transformer les variables en leur logarithme si d'autres raisons militent en ce sens, soit à faire toutes les estimations sur des ensembles de ménages assez étroitement définis, à l'intérieur desquels la variance résiduelle puisse être à peu près constante. La première méthode atténue les écarts de la dispersion résiduelle, la seconde peut être beaucoup plus intéressante. Dans l'échantillon que nous avons considéré par exemple les résultats montrent qu'il

(1) La valeur de la variance résiduelle que nous avons utilisée dans le tableau 8 pour la dépense alimentaire est celle qui a été obtenue plus haut pour les sous-groupes de ménages composés de deux adultes avec un enfant de moins de 6 ans et un enfant de 6 à 13 ans (voir tableau 5, 9^e ligne, colonne 5). Pour les dépenses exceptionnelles nous aurions dû introduire cette même spécification du type de famille mais les données trop imprécises sur ce point nous obligent à ne tenir compte que du revenu moyen ; étant donné le type de famille choisi, cela n'aurait du reste pas modifié beaucoup la valeur que nous avons lue sur la droite ajustée du graphique II c. Pour les dépenses courantes et les charges permanentes, ce sont aussi les droites ajustées des graphiques II que nous avons utilisées.

(2) Cette difficulté est exposée dans les traités d'économétrie. On peut consulter G. ROTTIER, *Initiation à l'économétrie*, p. 93.

ne serait pas avantageux de considérer ensemble les 983 ménages pour estimer les paramètres de comportement ; pour les dépenses alimentaires cependant le fait d'explicitier dans le modèle le facteur type de famille-égaliserait à peu près les variances résiduelles de tous les sous-groupes et permettrait ainsi de négliger toute division des ménages.

Ces remarques de méthode ne doivent pas faire oublier qu'il peut être utile d'étudier les distributions résiduelles des dépenses pour la contribution qu'elles apportent à notre connaissance du comportement des ménages. Cet article s'est limité à une étude descriptive des variances de ces distributions. Mais ses résultats encouragent à aller plus loin. Il est intéressant en effet de constater que l'action des revenus moyens et des types de familles sur les variances résiduelles n'est pas désordonnée et qu'elle s'exprime à travers des régularités qui paraissent non dépourvues de sens. D'autres analyses de ces phénomènes seraient évidemment nécessaires si l'on voulait s'appuyer sur eux pour étudier l'évolution de la distribution probable des dépenses à l'intérieur de groupes sociaux donnés. Si avec l'augmentation des revenus la dispersion de cette distribution devait s'accroître encore dans l'avenir, son étude deviendrait très utile. Elle présenterait toutefois des exigences nouvelles. Les premières exigences dont nous voulons parler sont relatives à la procédure des enquêtes. Il faut remarquer en effet que l'enquête de 1956, comme la plupart des autres enquêtes sur les budgets familiaux, ne répond pas très bien aux nécessités d'une étude de la dispersion des dépenses entre les ménages ; cela tient au fait déjà mentionné plus haut qu'un grand nombre de consommations y sont observées sur des périodes inférieures à l'année : une semaine pour les dépenses alimentaires, un trimestre pour les achats de vêtements, etc... Pour les consommations relevées sur une semaine par exemple, les variances que l'on peut calculer sont affectées par la dispersion des dépenses de chaque ménage entre toutes les semaines de l'année et cet effet, qui peut être très important, n'est pas susceptible d'être isolé. Les variances obtenues dans ces conditions surestiment la dispersion des dépenses entre les ménages ; plus la dispersion des dépenses entre les périodes est forte, plus ce biais est grand. Il faut tout d'abord souligner son incidence sur les résultats présentés ici : on peut concevoir en effet que la dispersion des dépenses entre les périodes croît avec le revenu. En particulier les ménages les plus riches accèdent de plus en plus nombreux à des achats qui ne sont pas de première nécessité et qui, en raison de leur poids dans le budget ou de leur caractère indivisible, font l'objet d'arbitrages très personnels quant à leur répartition dans le temps. Lorsqu'on travaille sur des grands groupes de produits, la dispersion des dépenses dans le temps risque par conséquent d'être plus forte pour les ménages riches que pour ceux qui disposent d'un revenu modeste. Ce phénomène peut expliquer que nous ayons observé une aussi forte sensibilité des variances résiduelles aux revenus moyens. Il est probable que cette sensibilité serait moins accusée si nous n'avions affaire qu'à la dispersion entre ménages. On voit maintenant que pour tenter de résoudre les difficultés de cet ordre, il faudrait dans les prochaines enquêtes, choisir des périodes d'observation qui tiennent compte le mieux possible de la fréquence réelle des différents achats et concevoir une exploitation des données qui permettent d'étudier les variations des dépenses dans le temps.

En dehors de l'amélioration des procédures d'enquêtes, une autre exigence devrait être satisfaite si l'on voulait approfondir l'étude de la dispersion des dépenses. Elle touche au schéma utilisé pour l'analyse de ces dispersions. Pour franchir le stade de la description, il faudrait en particulier examiner les processus qui peuvent rendre compte de la genèse des lois de distribution des dépenses dans un groupe social et les interpréter dans le cadre d'une théorie cohérente du consommateur.

ANNEXE

(Commune aux deux articles précédents)

I. DESCRIPTION DU SOUS-ÉCHANTILLON ÉTUDIÉ

a) Toutes les observations qui le constituent sont indépendantes

Dans l'enquête de 1956 environ 16 000 ménages ont été interrogés. Pour que les données de l'enquête soient à peu près représentatives de la population française, on a procédé à la duplication d'environ 5 000 ménages⁽¹⁾. Les résultats publiés jusqu'à ce jour proviennent donc d'un ensemble de 21 000 budgets. Mais l'échantillon initial de 16 000 ménages est le seul qui soit constitué d'observations indépendantes. Cette qualité présente un grand avantage pour l'analyse statistique à condition que l'on renonce à toute fidélité dans le caractère représentatif des résultats, comme nous avons décidé de le faire ici. Le sous-échantillon sur lequel nous travaillons a donc été extrait de l'échantillon initial de 16 000 ménages composé d'observations indépendantes, et non du lot des 21 000 budgets.

b) On a retenu d'abord toutes les observations relatives à un sous-ensemble de l'échantillon initial

Parmi ces 16 000 ménages, on a éliminé d'abord tous ceux qui n'habitent pas des villes de province ou qui ne sont pas composés de deux adultes (avec un ou plusieurs enfants de moins de 14 ans). La première restriction exclut d'une part la population des communes rurales et toute celle des villes de moins de 10 000 habitants ; d'autre part la population parisienne pour laquelle la qualité de l'enquête est moins bonne, ne serait-ce qu'en raison du taux de refus élevé. En raison de la seconde restriction la plupart des ménages retenus sont sans doute composés d'un couple avec des enfants. Le tableau suivant montre que le sous-ensemble ainsi défini représente un peu moins de 7 % de l'échantillon initial de 16 000 ménages. Ce sous-ensemble comprend environ 1 200 ménages.

c) On a enfin éliminé quelques catégories particulières de ménages

Nous avons enfin éliminé de ce sous-ensemble les catégories suivantes :

- ménages ayant refusé de déclarer leurs revenus ;
- ménages dont le chef est agriculteur, personnel de service ou inactif ;
- ménages dont les cartes perforées accusent des erreurs manifestes.
- enfin tous les ménages dont l'autoconsommation ou l'autofourniture dépasse 200 NF par an.

1) Voir **Consommation**, n° 2 de 1960, p. 19-20.

Répartition des 16 000 ménages interrogés en 1956 (en %)

(Entouré d'un trait fort : le sous-ensemble utilisé)

COMPOSITION DU MÉNAGE	CATÉGORIE DE COMMUNE			
	Communes rurales et communes urbaines de moins de 10 000 habitants	Communes urbaines de 10 000 habitants et plus en province	Agglomération parisienne	Ensemble
1 adulte avec ou sans enfant, 2 adultes sans enfant	22,5	14,2	9,7	46,4
2 adultes avec un ou plusieurs enfants	10,8	6,7	3,7	21,2
3 adultes et plus avec ou sans enfant	18,8	9,2	4,4	32,4
ENSEMBLE	52,1	30,1	17,8	100,0

Par la dernière restriction nous avons voulu éviter que l'autofourniture et l'autoconsommation qui ont été mal saisies et qui sont difficiles à évaluer, jouent un rôle trop important parmi les dépenses. A elle seule, cette restriction a réduit les effectifs du sous-ensemble d'environ 15 %. Les autres limites que nous avons énumérées excluent chacune une dizaine de ménages.

Finalement le sous-ensemble défini plus haut ne comprend plus que 983 ménages. Ce sont eux qui constituent le sous-échantillon que nous avons étudié.

d) Ce sous-échantillon a été divisé selon le type de famille et la catégorie socio-professionnelle du chef de ménage

Les 983 ménages ont été répartis selon 6 types de familles et 5 catégories socio-professionnelles. Les 30 sous-groupes ainsi constitués sont définis et leurs effectifs publiés dans le tableau I de l'article de Jan Cramer (p. 7). Les 5 catégories socio-professionnelles sont un regroupement de la nomenclature en 15 classes utilisée dans les exploitations antérieures ⁽¹⁾ :

(1) Voir par exemple : **Consommation**, n° 2 de 1960, p. 83-85. Le code qui figure ici en marge est celui qui est utilisé dans les deux articles de J. Cramer et M. Guillot.

C.S.P. 1 : gros indépendants : industriels, gros commerçants, professions libérales et cadres supérieurs ;

C.S.P. 2 : petits indépendants : artisans et petits commerçants ;

C.S.P. 3 : cadres moyens : cadres moyens et employés ;

C.S.P. 4 : ouvriers : contremaîtres, ouvriers, manœuvres ;

C.S.P. 5 : autres actifs.

Les cinq autres catégories socio-professionnelles ont été exclues du sous-échantillon : exploitants et salariés agricoles, gens de maison et autre personnel de service et inactifs.

2. DIVISION DE LA DÉPENSE TOTALE EN QUATRE GROUPES DE BIENS ET SERVICES

La dépense totale a été divisée en quatre catégories définies en fonction de la nomenclature de biens et services employée dans le compte rendu de l'enquête (1). Ce sont les numéros de cette nomenclature qui figurent dans les énumérations suivantes.

a) Consommation alimentaire

Ce poste comprend tous les achats alimentaires (postes 111-195), les boissons, les consommations hors du domicile, l'autoconsommation (dont l'importance est faible en raison de la façon dont a été choisi l'échantillon) et la glace à rafraîchir (regroupée avec eau, gaz et électricité dans la nomenclature sommaire de l'enquête).

b) Dépenses non-alimentaires courantes

221-225 Vêtements et chaussures.

232 Services domestiques.

241 Une partie seulement : tissus d'ameublement, linge de maison et couvertures.

243 Petits appareils ; réparations de biens durables.

251 Eau, gaz, électricité, à l'exception de la glace à rafraîchir.

252 Combustibles.

253 Produits d'entretien.

261 Hygiène.

262 Une partie seulement : les dépenses de pharmacie pour des produits d'usage courant achetés sans ordonnance.

273 Utilisation des véhicules, **moins** les dépenses remboursées par l'employeur (dans la nomenclature sommaire cette déduction n'a pas été faite).

274 Transports collectifs et P.T.T., à l'exception des frais de déménagement.

275 Vacances.

281 Tabac, allumettes.

282 Une partie seulement : les livres, disques, articles de sport, optique et photographie.

283 Autres dépenses de culture et de loisirs, à l'exception des frais d'éducation.

29 Une partie seulement de ces autres dépenses : billets de loterie, argent de poche, dons et quêtes, autres dépenses diverses, abonnements et cotisations et « autres dépenses ».

(1) Voir **Consommation**, n° 2 de 1960, pages 48 et suivantes.

c) Charges permanentes

- 231 Loyer et charges, à l'exception des réparations de l'intérieur et de l'extérieur.
- 233 Les frais des résidences secondaires.
- 283 En partie seulement : les frais d'éducation.
- 29 Une partie des autres dépenses : les pensions.
- 30 Les impôts.
- 241 } De tous les achats de biens durables compris dans ces postes,
- 242 } on retient seulement les mensualités des achats à crédit ainsi
- 271 } que les mensualités versées pour des achats antérieurs de plus
- 272 } d'une année (ce poste n'a pas été compris dans la nomenclature sommaire).
- 282 }

d) Dépenses exceptionnelles

- 231 Les frais de réparation seulement.
- 241 }
- 242 } De tous les achats de biens durables compris dans ces postes,
- 271 } les achats au comptant seulement, **moins** le produit des biens
- 272 } durables vendus, compris dans les dépenses diverses dans la
- 282 } nomenclature sommaire.
- 29 }
- 262 } Les dépenses de santé nettes, à l'exception des produits de pharmacie d'usage courant.
- 263 }
- 274 Les frais de déménagement seulement.
- 29 Les achats de bijoux et les amendes.