

Le Centre de Recherches et de Documentation sur la Consommation, association à but non lucratif régie par la loi de 1901, est un organisme scientifique fonctionnant sous la tutelle du Commissariat Général du Plan d'Équipement et de la Productivité. Son Conseil d'Administration est présidé par M. Claude GRUSON, Directeur Général de l'Institut National de la Statistique et des Études Économiques. Ses travaux sont orientés par un Conseil Scientifique que dirige le président de son Conseil d'Administration.

Les travaux du C.R.E.D.O.C. se développent dans les cinq lignes suivantes :

- Étude de l'évolution de la consommation globale par produit et par groupe socio-professionnel.
- Analyse du comportement du consommateur et économétrie de la demande.
- Établissement de perspectives de consommation à moyen terme.
- Méthodologie de l'étude de marché des biens de consommation.
- Étude du développement urbain.

Les résultats de ces travaux sont en général publiés dans la revue trimestrielle « Consommation ».

Exceptionnellement, ils peuvent paraître sous forme d'articles dans d'autres revues françaises ou étrangères ou bien faire l'objet de publications séparées, lorsque leur volume dépasse celui d'un article de revue.

Le Centre de Recherches et de Documentation sur la Consommation peut en outre exécuter des études particulières à la demande d'organismes publics ou privés. Ces études ne font qu'exceptionnellement l'objet de publication et seulement avec l'accord de l'organisme qui en a demandé l'exécution.

MEMBRES DU BUREAU

Président : Claude GRUSON, Directeur Général de l'Institut National de la Statistique et des Études Économiques.

Vice-Présidents :

Francis-Louis CLOSON : Président Directeur-Général de la Société « Information et Publicité ».
Erwin GULDNER, Directeur des Affaires Commerciales au Secrétariat d'État au Commerce Intérieur.

Trésorier : Paul LEMERLE, Inspecteur des Finances, Chargé de Mission au Commissariat Général du Plan d'Équipement et de la Productivité.

Secrétaire : Edgar POESSEL, Inspecteur Central du Trésor, Chef du Service Administratif et Financier du C.R.E.D.O.C.

Directeur : G. ROTTIER

Directeur adjoint : E.-A. LISLE

*Toute reproduction de textes ou graphiques est autorisée
sous réserve de l'indication de la source.*

IX^e année

N° 4 - octobre-décembre 1962

Consommation

COMMISSARIAT GÉNÉRAL DU PLAN D'ÉQUIPEMENT ET DE LA PRODUCTIVITÉ
INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

CENTRE DE RECHERCHES
ET DE DOCUMENTATION SUR LA CONSOMMATION
30, rue d'Astorg — PARIS 8^e

DUNOD

É D I T E U R

92, rue Bonaparte PARIS-6^e

Téléphone : DANton 99-15 C. C. P. PARIS 75-45

France : 42 F - Étranger : 50 F - Le numéro : 12 F

sommaire

ÉTUDES

Jan CRAMER

Dépenses et revenus des ménages d'après l'enquête
de 1956

Analyse d'un sous-échantillon restreint 5

Michel GUILLOT

La dispersion des dépenses et le revenu des ménages
Suite de l'analyse d'un sous-échantillon de l'enquête de 1956

37

Agnès PITROU

Croissance industrielle et développement du commerce
dans le Bassin de Berre

65

NOTES ET CHRONIQUES

Liliane KARAIMSKY

Les achats de produits alimentaires par les collecti-
vités

87

BIBLIOGRAPHIE

L'ÉCART ENTRE REVENUS ET DÉPENSES ET LA DISPERSION DES DÉPENSES DANS UNE ENQUÊTE SUR LES BUDGETS FAMILIAUX

Les analyses de budgets familiaux habituellement publiées ne considèrent pas directement les comptes individuels des ménages. Les familles sont d'abord classées en un certain nombre de groupes définis à partir de considérations économiques (comme les revenus) ou sociologiques (comme les structures familiales ou les catégories sociales) et l'étude porte sur les dépenses moyennes de chacun de ces groupes. Les calculs sont ainsi considérablement allégés sans que l'estimation des paramètres des lois de comportement en soit notablement affectée.

L'étude des données individuelles présente évidemment plus de souplesse et permet une critique plus sévère des données. Elle autorise aussi une analyse plus fine des comportements. Malgré son coût plus élevé, il paraît donc utile d'y recourir fréquemment, au moins pour des travaux expérimentaux.

Un autre avantage de cette technique est de rendre possible l'examen de problèmes méthodologiques particuliers, inabordables tant qu'on travaille sur données groupées. C'est dans cette voie que se situent l'étude de J. Cramer et celle de M. Guillot publiées dans ce numéro et consacrées à l'examen de quelques problèmes posés par l'observation des revenus et des dépenses dans l'enquête sur les budgets familiaux de 1956. Ces études sont réalisées à partir d'un sous-échantillon de cette enquête, constitué d'un millier de ménages urbains de province, comprenant tous 2 adultes, avec un ou plusieurs enfants. On a ainsi volontairement choisi une population dont les caractères soient assez homogènes. Les auteurs se sont heurtés à l'incertitude qu'introduisent les erreurs d'observation et les erreurs aléatoires qui sont peu ou prou le lot de toute enquête par sondage mais que les habituelles analyses sur données groupées masquent partiellement. Malgré cette difficulté quelques conclusions utiles peuvent être dégagées de ces deux articles.

1. Dans toutes les enquêtes sur les budgets familiaux, l'écart entre le revenu et la dépense totale de consommation est anormalement élevé.

J. Cramer observe empiriquement que cet écart est indépendant du niveau de revenu. Cette constatation, jointe à des hypothèses qui paraissent admissibles sur les relations entre dépense totale, revenu et « niveau de vie », permet de donner un fondement à des positions retenues intuitivement depuis assez longtemps et justifiées sur des bases théoriques dans quelques travaux récents :

— le revenu déclaré dans l'enquête est un meilleur indicateur du « niveau de vie » que la dépense totale habituellement prise comme variable explicative.

— l'utilisation du revenu déclaré dans l'analyse des courbes d'Engel, malgré sa grave sous-estimation, ne biaise pas l'estimation des élasticités par rapport au « niveau de vie » comme c'est le cas, on le sait, pour la dépense totale.

Le premier résultat justifie en particulier que, dans l'analyse de données groupées, les ménages ne soient pas classés selon leurs dépenses totales mais selon leurs revenus. C'est ce second mode de classement qui a été utilisé lors de l'exploitation générale de l'enquête de 1956.

2. La dispersion des dépenses individuelles, dans des groupes de ménages de revenus et de caractéristiques sociales voisins, est une des constatations les plus frappantes des enquêtes sur les budgets familiaux. M. Guillot s'intéresse à un aspect de ce phénomène : la dépendance entre cette dispersion et les caractéristiques d'un groupe de ménages bien défini. Son analyse confirme l'idée selon laquelle la dispersion des consommations individuelles croît systématiquement sous l'influence du revenu, et, bien que plus faiblement, de la composition de la famille. Elle met en lumière la régularité de ces liaisons : pour certaines catégories de dépenses, à une différence de 50 % entre les revenus moyens de deux groupes de ménages correspond une différence au moins égale entre les dispersions (mesurées par l'écart-type autour des droites d'Engel).

On a là une première idée de la façon dont s'élargit la variété des comportements lorsque le niveau de vie s'élève. Corrélativement, on voit quelle part de ce comportement les analyses classiques de courbes d'Engel risquent de laisser inexplicées. L'interprétation des résultats obtenus dans cet article est cependant limitée par le fait qu'une partie de la dispersion, non isolable en raison des conditions de l'enquête de 1956, est due à la variation des dépenses de chaque ménage entre les différentes périodes de l'année alors qu'ils n'ont été invités à relever leurs dépenses que sur une période assez courte. Des travaux engagés sur d'autres enquêtes où la période de relevé est plus longue, permettront bientôt de reprendre l'étude de ce problème.

G.R.

DÉPENSES ET REVENUS DES MÉNAGES

D'APRÈS L'ENQUÊTE DE 1956

(Analyse d'un sous-échantillon restreint)

par

Jan CRAMER *

I. INTRODUCTION

L'une des difficultés de l'économétrie des budgets familiaux tient à ce qu'elle cherche à décrire le comportement du consommateur en fonction d'une grandeur qui n'est pas directement observable : le niveau de vie. Les enquêtes auprès des ménages peuvent cependant fournir deux mesures approximatives de cette grandeur : les ressources déclarées par le ménage et l'ensemble de ses dépenses. Notre intention dans le présent article est d'étudier de façon empirique si l'on peut avoir des raisons, dans l'ajustement des courbes d'Engel, de choisir entre ces deux indicateurs.

D'après la définition des ressources et des dépenses, leur écart devrait être faible pour un ménage donné ; le choix serait donc indifférent. Or dans l'enquête de 1956 les dépenses des ménages dépassent largement et d'une façon générale les ressources qu'ils déclarent avoir. Ce phénomène est commun à la plupart des enquêtes qui permettent de rapprocher ces deux éléments. En 1857 déjà, Engel rejeta une enquête que lui présentait Rau et dans laquelle les dépenses dépassaient les revenus ; il ne savait accepter, écrit-il, l'explication donnée par ce dernier qui prétendait que l'écart correspondait au vol, à la quête et à l'endettement abusif de la classe ouvrière (1). Moins rigoureux qu'Engel, la plupart de ses successeurs qui constataient de tels écarts ont le plus souvent invoqué une sous-estimation probable des revenus sans pousser plus loin l'analyse. Seuls Cole et Utting ont entrepris en 1956 une étude approfondie de la question (2). L'enquête dont ils se servent a l'avantage de réunir des estimations de toutes les dépenses particulières, des impôts, de l'épargne et des revenus, permettant une reconstitution complète des comptes des ménages ; en revanche ces estimations n'ont pas trait aux mêmes ménages, de sorte qu'on ne peut comparer dépenses et ressources que pour des groupes de ménages et non pas pour des ménages individuels. L'étude détaillée de Cole et Utting

* Chargé de mission au C.R.E.D.O.C. de 1959 à 1961, Jan CRAMER est actuellement professeur à l'Université d'Amsterdam.

(1) E. ENGEL, Die Productions und Konsumtionsverhältnisse des Königreichs Sachsen. Article de 1857, réédité dans le *Bulletin de l'Institut National de Statistique*, vol. 9 (1895).

(2) COLE, DOROTHY et UTTING, J.E.G. Estimating expenditure, savings and income from household budgets ; *Journal of the Royal Statistical Society*, Séries A, vol. 119 (1956). Voir en particulier l'intervention de Kemsley lors de la discussion.

conclut à l'existence indiscutable d'un écart important, sans toutefois en donner une explication (1).

On peut penser que le recours aux observations individuelles est de nature à faciliter l'examen de ces problèmes en évitant toute perte d'information et en permettant de repérer les quelques observations extrêmes qui ont une influence décisive sur l'issue des analyses. Il nous a été possible d'appliquer cette procédure à l'enquête en 1956. C'est pour cette raison que nos conclusions sont susceptibles de compléter les résultats acquis dans les travaux antérieurs.

Dans les sections 2 et 3 nous analysons de façon systématique la différence entre ressources et dépenses. La discussion conduit à plusieurs procédés pour l'estimation des courbes d'Engel ; nous comparons leurs résultats pour le seul cas de la consommation alimentaire dans la section 4.

L'option en faveur des données individuelles alourdit sensiblement les calculs, et il est impossible pour le moment d'étudier l'ensemble des 20 000 ménages de l'enquête selon cette procédure. L'étude que nous entreprenons n'a de toute façon qu'un intérêt méthodologique ; il n'y a donc aucun inconvénient à se borner à l'examen d'un échantillon restreint.

Parmi tous les ménages observés par l'enquête, nous n'avons retenu que les ménages composés de deux adultes, avec un ou plusieurs enfants de moins de 14 ans, qui habitent les villes de province. En outre, nous avons exclu les ménages dont le chef est agriculteur ou inactif, et tous ceux dont l'autoconsommation ou l'autofourniture dépasse 200 NF par an. Nous sommes parvenus ainsi à un échantillon de 983 ménages ; c'est cet échantillon que nous allons examiner dans cet article (2).

Nous avons réparti ces 983 ménages selon 6 types de famille et 5 catégories socio-professionnelles. Cette répartition qui est définie dans le tableau I est constamment utilisée par la suite.

Il convient encore de définir les principales variables que nous faisons intervenir. Les **revenus** ou **ressources** du ménage sont les revenus de toutes sortes déclarés par le ménage (3). Cette déclaration se rapporte dans la plupart des cas aux revenus de la dernière année de calendrier (4). La **dépense totale** de chaque ménage a été estimée à partir des dépenses déclarées par le ménage pour des périodes différentes selon les produits : une semaine pour l'alimentation, un mois pour les biens durables peu coûteux, trois mois pour les vêtements, et douze mois pour d'autres biens durables et les vacances (5). Toutes ces dépenses ont été additionnées après avoir été ramenées à la base annuelle. Cette dépense totale comprend toutes les dépenses de consommation, y compris l'autoconsommation et l'autofourniture s'il y a lieu, mais elle ne comprend ni impôts, ni épargne sous quelque forme que ce soit.

(1) Sur l'écart entre les ressources et les dépenses observées dans les enquêtes auprès des ménages on peut aussi consulter le compte rendu méthodologique d'une enquête faite aux États-Unis en 1950 : « Methodology of the survey of consumer expenditures in 1950 », H. HUMES LAMALE, *University of Pennsylvania*, 1959, p. 96. Les auteurs font remarquer que presque dans tous les groupes de ménages qu'ils ont eu à constituer la dépense est supérieure au revenu, cela provient à leur avis de ce que le revenu est sous-estimé et de ce que les dépenses le sont aussi, mais un peu moins.

(2) Pour une description plus détaillée de cet échantillon, on peut se reporter à l'annexe, p. 61.

(3) Ou, rarement, estimés par l'enquêteur.

(4) Voir *Consommation*, n° 1 de 1958, p. 7.

(5) Pour la liste complète, voir *Consommation*, n° 2 de 1960, p. 15.

TABLEAU I

Structure du sous-échantillon selon le type de famille et la catégorie socio-professionnelle
(Nombre de ménages)

TYPE DE FAMILLE		CATÉGORIE SOCIO-PROFESSIONNELLE					ENSEMBLE
		Gros indépendants	Petits indépendants	Cadres moyens	Ouvriers	Autres actifs	
2 adultes avec :		C.S.P. 1	C.S.P. 2	C.S.P. 3	C.S.P. 4	C.S.P. 5	
1 enfant de moins de 6 ans	Ty 1	12	11	44	99	7	173
1 enfant de 6 à 13 ans.....	Ty 2	14	28	64	99	11	216
2 enfants de moins de 6 ans.....	Ty 3	8	10	26	59	5	108
1 enfant de moins de 6 ans et 1 enfant de 6 à 13 ans. }	Ty 4	9	12	33	49	9	112
2 enfants de 6 à 13 ans	Ty 5	10	14	29	59	7	119
3 enfants ou plus	Ty 6	26	15	47	157	10	255
ENSEMBLE		79	90	243	522	49	983

Dans la section 4 nous considérons la **dépense alimentaire** des ménages. Cette variable comprend les achats alimentaires, y compris les boissons, ainsi que l'autoconsommation et l'autofourniture s'il y a lieu.

2. DÉPENSES ET REVENUS

L'écart dans l'enquête 1956

Avant d'étudier l'écart entre ressources et dépenses dans notre échantillon restreint, il est bon de rappeler avec plus de précision que nous ne l'avons fait jusqu'à présent comment cet écart se présente dans l'ensemble de l'enquête. Pour les 21 000 ménages de l'enquête, il y a simultanément sous-estimation des dépenses et excédent des dépenses par rapport aux revenus.

Les données de la comptabilité nationale font ressortir une sous-estimation moyenne des dépenses par ménage d'environ 1 200 NF, soit de 20 % (1). Quant à l'excédent des ressources par rapport aux dépenses, ces deux grandeurs étant mesurées par l'enquête, il est d'environ 1 100 NF par ménage (2).

L'écart entre dépenses et ressources observé dans l'enquête est trop grand pour être attribué aux aléas du sondage, et il ne peut pas non plus être imputé à quelques ménages extrêmes ou à un sous-groupe particulier : il s'agit d'un phénomène général qui ne varie guère selon la catégorie socio-professionnelle, la catégorie de commune ou la composition de la famille.

Nous retrouvons un écart semblable dans l'échantillon que nous étudions ; pour l'ensemble des 983 ménages les ressources s'élèvent en moyenne à 7 723 NF, tandis que les dépenses sont de 9 595 NF (3). Cet écart dont la moyenne est de 1 900 NF varie beaucoup d'un ménage à l'autre. Dans 232 cas sur 983, il est négatif ; c'est-à-dire qu'alors les ressources l'emportent sur les dépenses.

Mais pour aller plus avant dans l'analyse il est nécessaire de préciser le rôle que nous y ferons jouer à chaque observation individuelle ; la position que nous adoptons à cet égard est fondamentale pour la suite de ce travail.

Le rôle des observations individuelles

Ce qui nous préoccupe est le fait que les dépenses dépassent largement les ressources dans chaque groupe important de ménages, de sorte que ce phénomène ne peut être attribué au hasard de l'échantillonnage. Dès que nous considérons des ménages individuels nous trouvons d'une part de nombreux cas où le rapport des deux variables est correct, tandis que d'autre part il n'est pas d'observation aberrante, aussi extrême soit-elle,

1) La confrontation entre les données de l'enquête et celles de la comptabilité nationale ne peut être faite que pour les postes dont les définitions sont comparables (80 % de la dépense totale environ). Cette confrontation a donné les résultats suivants (en moyenne par ménage, par an et en NF) :

Dépenses des ménages d'après la comptabilité nationale, postes comparables avec l'enquête seulement	7 920
Dépenses des mêmes postes d'après l'enquête	6 883

Pour ces seuls postes, la sous-estimation de l'enquête est ainsi de 1 000 NF ; on n'a pas de raison de supposer que sa valeur relative soit plus faible pour les dépenses non comparables. D'où l'ordre de grandeur de 1 200 NF que nous retenons.

(2) Les moyennes par ménage et par an, en NF sont exactement :

Dépense totale, tous produits, d'après l'enquête	7 848
Dont autoconsommation et autofourniture	531
Ressources déclarées lors de l'enquête	6 369
Impôts directs déclarés lors de l'enquête	157

(3) Une comparaison entre ces chiffres et ceux que nous donnons dans la note précédente montre que le niveau des dépenses et des revenus (ainsi que leur écart) est sensiblement plus élevé dans l'échantillon que dans l'ensemble de l'enquête ; ceci s'explique en grande partie par le fait que l'échantillon ne comprend ni retraités, ni ménages agricoles.

qui ne puisse trouver une explication acceptable. Il y a, par exemple, des ménages où les dépenses dépassent trois fois leurs ressources ; or rien n'empêche qu'au moment de l'enquête leur situation n'ait été exactement telle : il peut s'agir de lourdes dépenses extraordinaires, d'achats d'équipement, d'une baisse temporaire des ressources, etc... Cependant ces observations, que nous ne savons pas rejeter quand elles sont prises une par une, conduisent dans l'ensemble à un résultat inacceptable. Dans ces conditions, le seul procédé fécond est de **considérer les observations individuelles comme des variables aléatoires qui ne nous intéressent que dans la mesure où elles nous apportent des renseignements sur la distribution de probabilité à laquelle elles appartiennent.**

Notre premier souci est de ne pas formuler au départ des hypothèses trop optimistes quant à l'homogénéité des ménages ; nous considérons donc d'abord un par un les plus petits sous-groupes que nous avons distingués, c'est-à-dire les 30 sous-groupes du sous-échantillon que définit la double classification selon la catégorie socio-professionnelle et le type de famille (voir tableau I). Pour chacun d'entre eux, nous examinons d'abord la distribution des ressources et des dépenses des ménages.

L'écart et le niveau des ressources

A titre d'exemple, les graphiques 1, 2 et 3 fournissent ces distributions pour trois sous-groupes de ménages. On y voit comment sont reliées les deux variables étudiées. Chaque point est représentatif d'un ménage. Les échelles sont arithmétiques et les mêmes dans les deux directions, de sorte que les points situés au-dessous et à droite de la diagonale tracée représentent des ménages pour lesquels les revenus dépassent les dépenses ; pour la plupart des ménages c'est le cas inverse qui se produit.

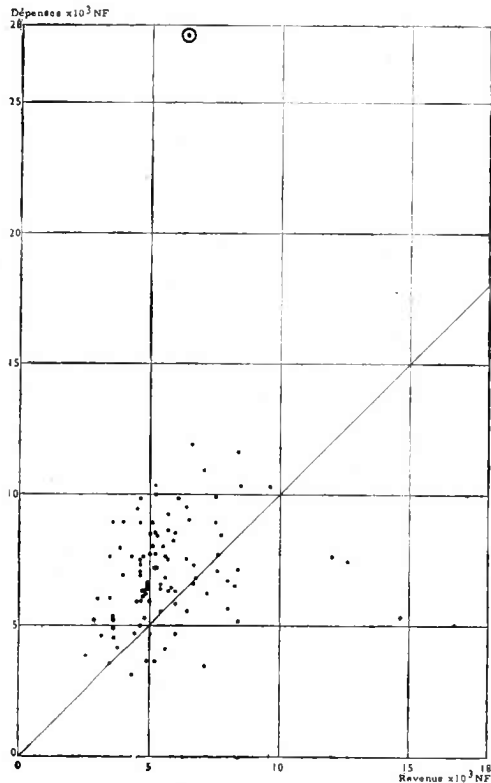
Ces nuages de points permettent trois conclusions. D'abord les points sont fortement dispersés, ce qui veut dire qu'il n'y a qu'une faible relation entre dépenses et revenus. Cette dispersion paraît augmenter dans la mesure où les points s'éloignent de l'origine ; si nous considérons les écarts individuels (la distance verticale de chaque point à la diagonale), leur dispersion s'accroît sensiblement vers la droite des graphiques. La troisième constatation est la plus importante : autant qu'on puisse en juger la moyenne de ces écarts individuels reste constante. L'examen des graphiques non publiés, relatifs aux autres sous-groupes de ménages, ne semble pas contredire cette remarque. Si nous désignons la dépense totale par X, les revenus par Y et l'écart entre ces deux variables par Z,

$$Z = X - Y$$

l'absence de variation systématique de l'écart Z signifierait, puisqu'il s'agit d'une variable aléatoire, que cette variable est indépendante de l'une ou de l'autre des deux variables qui la définissent (1).

Cette indépendance correspondrait à une description tellement simple des observations que nous avons intérêt à en examiner la possibilité. Nous étudierons d'abord l'hypothèse de l'indépendance entre l'écart (Z) et les ressources (Y).

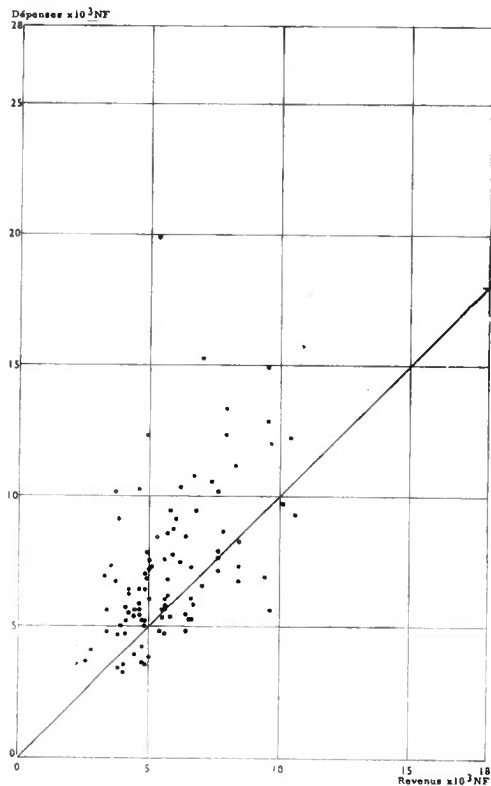
(1) Évidemment l'écart ne peut être indépendant des deux variables à la fois : l'identité qui le définit intervient. Dans le cas où Z est indépendant de Y, les dépenses X peuvent être considérées comme la somme de deux variables indépendantes ($X = Y + Z$) et un simple calcul de covariance montre que X ne peut être indépendant ni de Y, ni de Z. Le même raisonnement s'applique dans le cas où Z est indépendant de X.



GRAPHIQUE 1

Ressources et dépenses

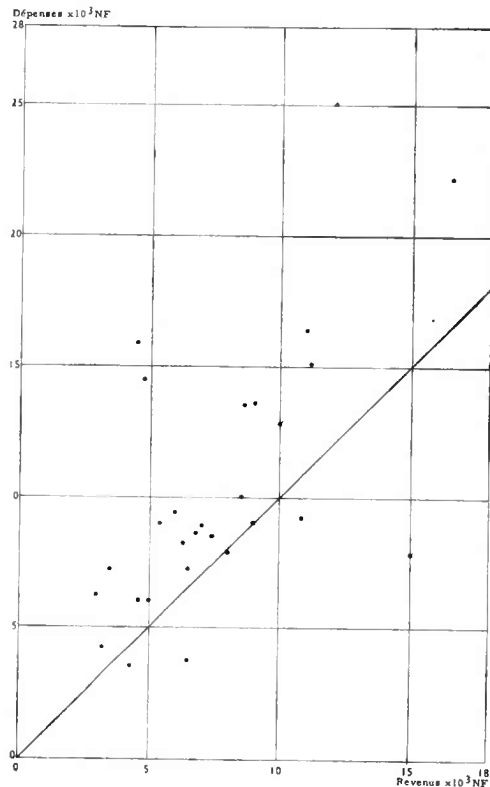
(99 ménages d'ouvriers composés de deux adultes et un enfant de moins de 6 ans)



GRAPHIQUE 2

Ressources et dépenses

(99 ménages d'ouvriers composés de deux adultes et un enfant de 6 à 13 ans)



GRAPHIQUE 3

Ressources et dépenses

(28 ménages de petits indépendants composés de deux adultes et un enfant de 6 à 13 ans)

Si cette hypothèse est vraie, la dépense X est la somme de deux variables aléatoires indépendantes, Z et Y , et dans ce cas le coefficient angulaire de la régression de X sur Y est égal à un. Sous cette forme, le modèle que nous venons d'énoncer est susceptible d'un test statistique. Il suffit de calculer la régression qui caractérise le nuage de points pour chacun des trente sous-groupes de ménages (et non seulement pour les trois cas illustrés par les graphiques), et de tester l'hypothèse d'un coefficient angulaire égal à un pour chacun d'entre eux (1).

Nous présentons brièvement le principe de ce test, dont les résultats sont indiqués dans le tableau 2 ; chaque ligne de ce tableau correspond à un sous-groupe, les codes de la catégorie socio-professionnelle et du type de famille correspondant à la classification du tableau 1. Nous avons ajusté dans chaque sous-groupe la régression de X sur Y , soit

$$X = aY + b$$

et nous avons ainsi obtenu des estimations \hat{a} (colonne 4). Chaque ajustement conduit en outre à un coefficient de corrélation R entre X et Y , indiqué en colonne 5. Pour tester si le coefficient a est ou non significativement différent de un nous considérons le rapport des deux quantités S_1 et S_2 , définies par

$$S_1 = \frac{1}{N-2} \sum_N (x - \hat{a}y)^2$$

et

$$S_2 = (\hat{a} - 1) \sum_N y^2$$

Dans ces expressions N est le nombre d'observations du sous-groupe étudié, auxquelles la sommation s'étend, et y et x sont les variables Y et X centrées autour de leurs moyennes respectives pour le sous-groupe. La première quantité S_1 est une estimation de la variance résiduelle de la régression que nous avons ajustée, et la deuxième somme S_2 correspond à l'apport à la variance résiduelle qui serait dû à l'imposition d'une valeur du coefficient a égale à un.

Nous considérons le rapport de ces deux quantités, soit

$$F = \frac{S_2}{S_1}$$

et nous affirmons que si notre hypothèse est vraie cette quantité est distribuée comme le rapport de deux variances F avec $(1, N - 2)$ degrés de liberté. Cette distribution est connue, et nous pouvons par conséquent vérifier si la valeur obtenue pour F est cohérente avec l'hypothèse ou si, au contraire, elle est tellement exceptionnelle que nous préférons abandonner l'hypothèse. Dans ce jugement nous appliquons les seuils habituels de 5% et 1% (2).

(1) On peut évidemment aussi bien tester l'hypothèse que Z soit indépendant de Y par l'examen de la corrélation entre ces deux variables. Ce test conduit aux mêmes résultats que celui que nous présentons ici.

(2) Pour un exposé plus complet de ce test voir N. TABARD, *Consommation et niveau de vie de quelques groupes sociaux*, *Consommation*, n° 2 de 1961, notamment p. 10-12.

TABLEAU 2 — Tests de l'hypothèse $\alpha = 1$ dans la régression $X = \alpha Y + b$
 (n.s. = non significatif ; * = significatif à 5% ; ** = significatif à 1%)

12

SOUS-GROUPE		NOMBRE D'OBSERVATIONS N	ESTIMATION $\hat{\alpha}$	CORRÉLATION r	RAPPORT F	DEGRÉS DE LIBERTÉ	RÉSULTAT DU TEST
Code C.S.P.	Code Ty						
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	1	12	0,861	0,293	0,02	1,10	n.s.
	2	14	0,843	0,957	4,53	1,12	n.s.
	2	13	0,604	0,777	4,02	1,11	n.s.
	3	8	0,684	0,869	2,86	1,6	n.s.
	4	9	0,664	0,705	1,77	1,7	n.s.
	5	10	0,580	0,730	4,78	1,8	n.s.
2	6	26	0,343	0,328	10,60	1,24	* *
	1	11	0,339	0,309	3,61	1,9	n.s.
	2	28	0,935	0,618	0,08	1,26	n.s.
	3	10	0,130	0,747	455,74	1,8	* *
	3	9	0,014	0,009	2,87	1,7	n.s.
	4	12	0,952	0,620	0,02	1,10	n.s.
3	5	14	0,491	0,810	24,66	1,12	* *
	6	15	0,448	0,325	2,34	1,13	n.s.
	1	44	0,822	0,538	0,81	1,42	n.s.
	2	64	0,718	0,733	11,31	1,62	* *
	3	26	0,665	0,392	1,10	1,24	n.s.
	4	33	0,738	0,471	1,11	1,31	n.s.
4	5	29	1,089	0,529	0,07	1,27	n.s.
	6	47	0,939	0,574	0,09	1,45	n.s.
	1	99	0,358	0,229	17,34	1,97	* *
	2	99	0,812	0,519	1,91	1,97	n.s.
	3	59	1,031	0,553	0,02	1,57	n.s.
	4	49	0,930	0,563	0,12	1,47	n.s.
5	5	59	0,843	0,586	1,03	1,57	n.s.
	6	157	0,825	0,536	2,82	1,155	n.s.
	1	7	0,388	0,256	0,85	1,5	n.s.
	2	11	0,954	0,581	0,01	1,9	n.s.
	3	5	1,065	0,921	0,06	1,3	n.s.
	4	9	0,502	0,239	0,40	1,7	n.s.
6	5	7	0,067	0,093	8,51	1,5	*
	6	10	0,170	0,281	16,40	1,8	* *

. Une observation extrême enlevée ; voir texte.

Le tableau 2 se termine donc par la valeur de F , les degrés de liberté et le résultat du test (colonnes 6, 7 et 8).

On notera que dans deux cas nous avons enlevé des observations extrêmes d'un sous-groupe. Il s'agit dans le premier cas (code CSP 1, Ty 2) d'un ménage avec 66 500 NF de revenus et 62 000 NF de dépenses, dans le deuxième cas (code CSP 2, Ty 3) d'un ménage avec 80 000 NF de revenus et 18 500 NF de dépenses. Ces ménages ont pu être repérés avant que l'analyse ne soit faite ; ils se situent nettement en dehors du sous-groupe auquel ils appartiennent ; nous avons fait l'analyse avec et sans ces observations. Dans le second sous-groupe où nous avons enlevé une observation le test n'est nettement significatif qu'en raison de cette seule observation. Il reste évidemment d'autres cas, par exemple dans le sous-groupe du graphique 2 (code CSP 4, Ty 1) où le résultat significatif du test est probablement dû à l'observation encerclée sur le graphique. Nous n'avons cependant pas pu repérer ce ménage avant que l'analyse soit faite, et il serait erroné de l'enlever une fois les résultats connus ; ce procédé ôterait toute validité aux tests. D'ailleurs même les observations aberrantes repérées d'avance ne peuvent être éliminées de l'échantillon ; ces observations sont aussi valables que les autres et il n'est pas légitime de les exclure. Il est pourtant intéressant de voir à quel point une seule observation peut déterminer le résultat du test. Cette constatation nous amène à garder une certaine marge d'appréciation des tests ; en effet, chaque fois qu'on examine avec soin les données individuelles, on est frappé par l'étendue et les discontinuités du champ de variation des ressources et des dépenses. L'hypothèse d'une distribution gaussienne des résidus à la régression (ou des écarts Z), qui est à la base de notre analyse, n'est donc probablement pas strictement valable, les observations extrêmes étant trop fréquentes. Ceci nous conduit à considérer les résultats des tests dans leur ensemble et à attribuer les quelques exceptions que nous y trouvons à des observations extrêmes.

Interprétés de cette façon, les résultats du tableau 4 sont une indication que l'hypothèse de l'indépendance de Z et Y peut être retenue. La conclusion est nette dans 24 sur 30 des sous-groupes ; parmi les six autres il y en a trois à faibles effectifs (10 ménages au moins), et un pour lequel le graphique 2 révèle le rôle d'une observation aberrante. Nous retenons donc l'hypothèse d'indépendance entre l'écart Z et le niveau des ressources. Par voie de conséquence, nous savons que les dépenses et l'écart ne sont pas indépendants ; leur liaison ne faisant que nous renvoyer à la définition de l'écart ne demande pas d'interprétation particulière (1).

Le type de famille et la catégorie socio-professionnelle

Avant de commenter cette conclusion d'indépendance entre l'écart Z et le niveau des ressources Y , nous poussons l'analyse un peu plus loin. Si l'écart a la même espérance mathématique, indépendante du revenu, **à l'intérieur** de chaque sous-groupe, il y a de fortes chances qu'il en soit de même **entre** les différents sous-groupes que nous avons distingués. Nous avons donc fait une analyse de variance des écarts Z qui porte en premier lieu sur la variation entre les 6 types de famille à l'intérieur de chacune des 5 catégories professionnelles. Pour ce test nous évaluons de nouveau le rapport de deux variances,

$$F = \frac{S_3}{S_4}$$

(1) Voir plus haut, p. 9, note 1.

où :

$$S_4 = \frac{\sum_{ij} (Z_{ij} - \bar{Z}_j)^2}{\sum_j N_j - K}$$

et :

$$S_3 = \frac{\sum_j N_j (\bar{Z}_j - \bar{Z})^2}{K - 1}$$

Dans ces expressions l'indice j se rapporte aux N_j ménages du type de famille j ($j = 1, \dots, 6$) d'écart moyen \bar{Z}_j ; l'écart moyen d'ensemble (c'est-à-dire tous types de famille combinés) étant \bar{Z} ; ceci pour chaque catégorie socio-professionnelle. Le rapport suit de nouveau la distribution connue de F avec $(K - 1, \sum_j N_j - K)$ degrés de liberté, ce qui nous permet de décider si les écarts ont la même moyenne indépendamment du type de famille.

TABLEAU 3

Analyse de variance de l'écart Z par sous-groupe

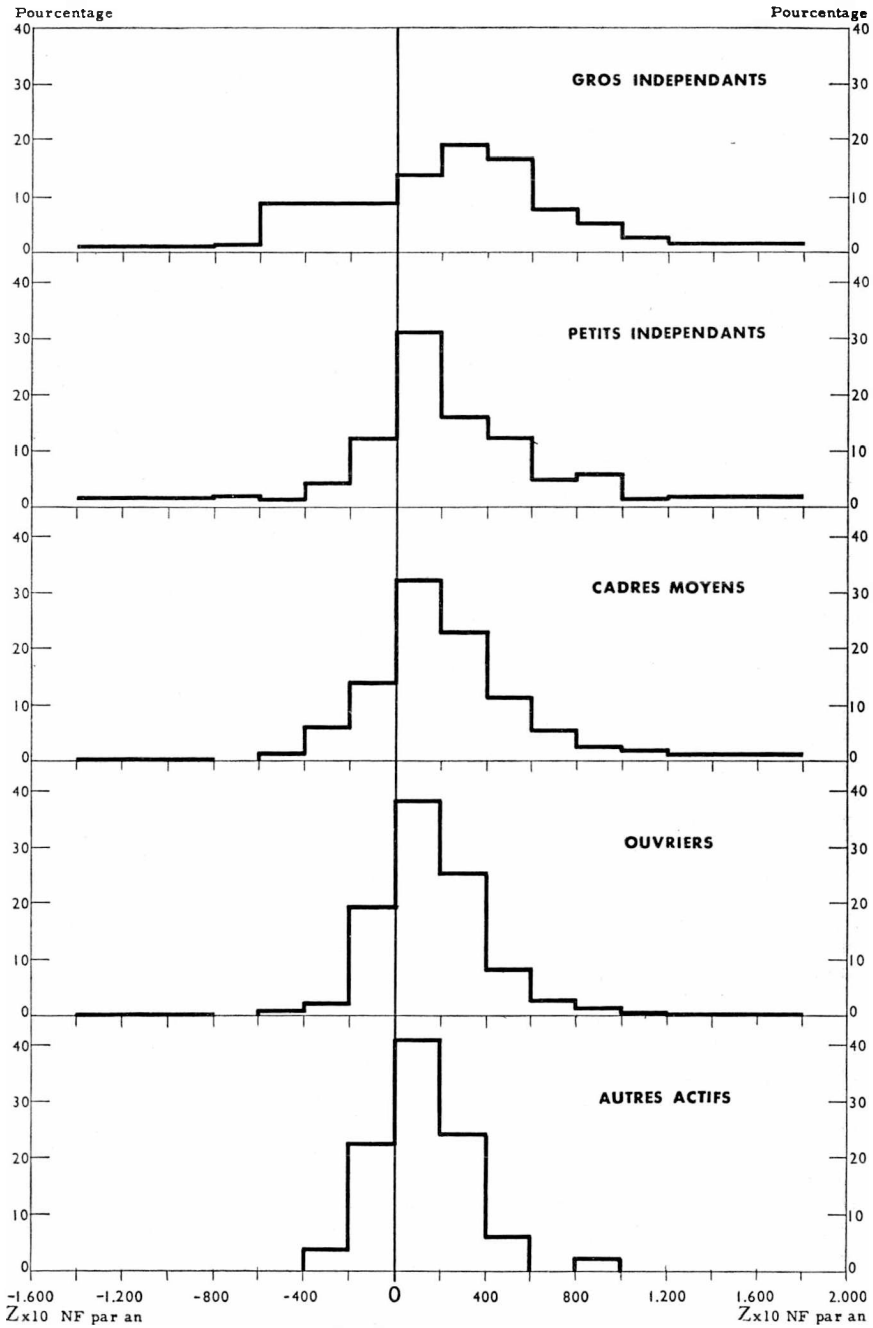
(n.s = non significatif ; * = significatif à 5% ; * * = significatif à 1%)

	NOMBRE D'OBSERVATIONS N	ÉCART MOYEN Z	RAPPORT F	DEGRÉS DE LIBERTÉ	RÉSULTAT DU TEST
	(1)	(2)	(3)	(4)	(5)
Variations entre types de famille d'une seule catégorie socio-professionnelle					
C.S.P. 1	79	227	0,67	5,73	n.s.
C.S.P. 1 .	78	235	0,65	5,72	n.s.
C.S.P. 2	90	159	1,66	5,84	n.s.
C.S.P. 2 .	89	230	1,05	5,83	n.s.
C.S.P. 3	243	250	0,10	5,237	n.s.
C.S.P. 4	522	161	0,54	5,516	n.s.
C.S.P. 5	49	144	1,29	5,43	n.s.
Variation entre catégories socio-professionnelles					
Ensemble	983	187	1,99	29,953	* *
Ensemble .	981	194	1,46	29,951	n.s.

. Une observation extrême (2 pour l'ensemble) enlevée.

GRAPHIQUE 4

Distribution de l'écart Z dans les différentes catégories socio-professionnelles



Ces tests sont répétés cinq fois, une fois pour chaque catégorie socio-professionnelle ; les résultats que l'on trouve au tableau 3 permettent de retenir l'hypothèse dans tous les cas. Ensuite nous avons appliqué le même test pour décider si l'écart varie selon la catégorie socio-professionnelle ; si on exclut les deux ménages aberrants, l'hypothèse peut encore une fois être retenue. Ces calculs sont également présentés dans le tableau 3.

La dispersion de l'écart

Nous concluons de ces analyses que l'écart Z est indépendant du niveau des revenus, du type de famille et de la catégorie socio-professionnelle. Cela veut dire que la moyenne des écarts observés ne varie pas sensiblement selon ces critères ; leur distribution n'est pourtant pas forcément la même d'un groupe de ménages à un autre : la variance et la forme de la distribution peuvent très bien différer. Le graphique 4, qui reproduit la distribution de Z de chacune des cinq catégories socio-professionnelles, montre en effet que si les moyennes sont assez semblables la dispersion varie d'une façon régulière : la variance diminue systématiquement quand on passe des gros indépendants aux « autres actifs ». Notons qu'en outre les distributions paraissent à peu près symétriques ; elles ne sont certainement pas gaussiennes, nous le disions plus haut, mais elles ne sont cependant pas très éloignées de cette forme.

La variation de la dispersion selon les catégories socio-professionnelles pourrait refléter l'hétéroscédasticité de Z par rapport aux revenus Y que suggéraient déjà les graphiques 1 à 3 ; c'est-à-dire qu'il peut y avoir une relation entre le niveau des ressources et la variance de l'écart entre dépenses et ressources. Dans ce cas les distributions du graphique 4 se distingueraient en raison des différents niveaux de ressources des catégories socio-professionnelles et non pas à cause d'autres facteurs propres à ces catégories. Pour vérifier cette possibilité nous avons réparti les 983 ménages du sous-échantillon en 11 groupes de 90 ménages chacun (1) après les avoir rangés selon le revenu, sans nous occuper ni de la catégorie socio-professionnelle, ni du type de famille. Ce regroupement permet un examen valable de la dispersion de Z , estimée à partir de groupes à effectif égal, selon le niveau de revenu.

Le graphique 5 montre que l'écart-type de Z croît régulièrement avec les ressources, et que sa variation se confond avec celle que l'on observe entre les groupes socio-professionnels. L'hétéroscédasticité de Z selon le niveau des revenus est nette.

Le même regroupement des ménages selon leurs revenus permet de vérifier encore le résultat déjà obtenu selon lequel la moyenne de Z est indépendante du revenu ainsi que de tout autre facteur. Le graphique 6 montre que cette moyenne ne varie en effet guère avec le revenu, à l'exception toutefois du groupe 11 qui comprend les ménages les plus riches du sous-échantillon. Cette exception n'est pas trop inquiétante : le sous-échantillon comprend sans doute quelques cas de surestimation erronée des ressources et ces ménages se trouvent forcément tous dans le groupe 11 et accusent en même temps une valeur de Z beaucoup trop petite ou même négative.

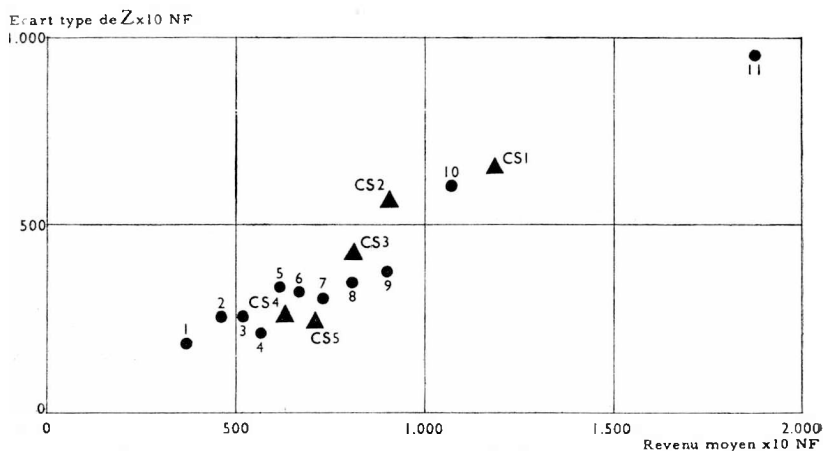
(1) Plus précisément 10 groupes de 90 ménages et un groupe de 83 ménages.

GRAPHIQUE 5

Écart-type de Z et ressources moyennes

Points de I à II : groupement des ménages selon le revenu.

CS I à CS 5 : groupement des ménages selon la catégorie socio-professionnelle.

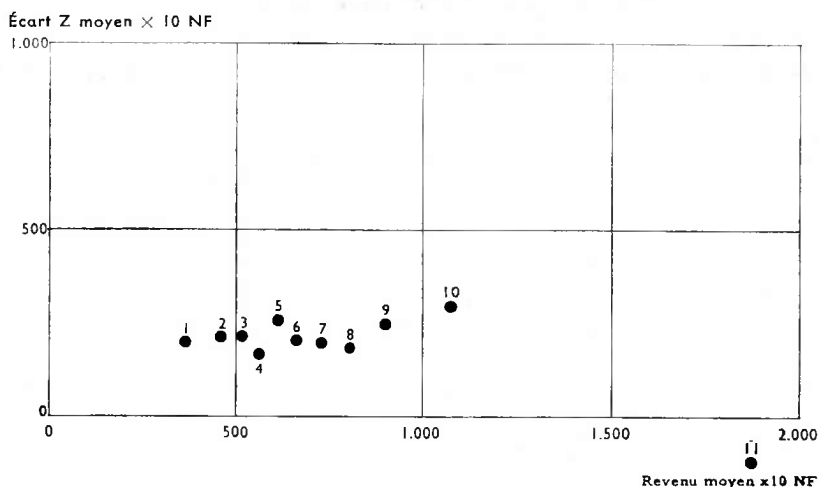


L'écart Z est donc indépendant des ressources et des autres critères retenus ; il est en outre hétéroscédastique avec le niveau des ressources. Nous pouvons désormais considérer que ce résultat est solide. Il correspond aux conclusions des travaux de D. Cole et J. E. G. Utting (1). Dans leur étude la constance de l'écart entre ressources brutes et dépenses est cependant passée presque inaperçue et nous devons reconnaître qu'il s'agit dans un cas comme dans l'autre d'un phénomène qu'on ne saurait **expliquer**. Nous verrons de plus près dans la section suivante que l'écart N est composé de plusieurs éléments fort différents et il n'y a aucune raison de croire que leur somme soit en moyenne constante ; il s'agit d'une constatation de fait, d'une description simple de l'état de choses que nous observons. Ajoutons encore que si les observations ne contredisent pas l'hypothèse simple que nous avons choisie, cela est dû surtout à la dispersion extrême des Z, et non pas à leur uniformité, de sorte qu'il est probable que d'autres hypothèses (par exemple d'un écart en moyenne proportionnelle aux ressources, c'est-à-dire constant en pourcentage de ces dernières) auraient également pu être acceptées.

(1) Voir article cité plus haut.

GRAPHIQUE 6

Ecart Z moyen et ressources moyennes (groupement des ménages selon le revenu)



Les conclusions immédiates qu'on peut tirer de nos résultats sont peu nombreuses. La suite de l'analyse est cependant considérablement simplifiée du fait que la différence entre les dépenses et les ressources ne varie ni avec le niveau des ressources, ni avec la catégorie socio-professionnelle, ni enfin avec le type de famille. Quant aux dépenses elles-mêmes, elles peuvent être considérées comme la somme de deux variables aléatoires et indépendantes (les ressources et l'écart) ; la variance des dépenses est donc toujours plus forte que celle des ressources. Toute autre conclusion issue de nos résultats dépend étroitement des hypothèses supplémentaires que nous introduisons. Nous en donnons un exemple dans la section suivante.

3. LE NIVEAU DE VIE ET SA MESURE (1)

Dans l'analyse des budgets de famille nous admettons que chaque ménage a un certain **niveau de vie** qui résume ses possibilités réelles de dépense et qui, par conséquent, détermine en grande partie le niveau et la composition de sa consommation. Le niveau de vie diffère des ressources du ménage dans la mesure où celui-ci veut en consacrer une partie à l'épargne ou, au contraire, est prêt à y suppléer, soit par l'emploi de son

(1) L'analyse de cette section aboutit à une répartition de Z en plusieurs éléments selon une distinction théorique d'après leur origine. L'article de Michel GUILLOT dans le même numéro donne une analyse complémentaire de la variance de Z selon les catégories de dépenses qui constituent la dépense totale.

patrimoine, soit en augmentant son endettement. Le niveau de vie correspond ainsi au niveau de dépenses que le ménage se fixe pour objectif, mais il s'agit d'un niveau d'équilibre, ou de long terme, plutôt que des dépenses effectives pendant une période déterminée (1).

Le niveau de vie ainsi conçu est le déterminant principal du comportement du consommateur ; mais il n'est pas susceptible d'observation directe, quelles que soient les conditions techniques d'une enquête. Les ressources et la dépense totale diffèrent toutes les deux du niveau de vie. Définissons rapidement ces différences en nous référant à l'enquête de 1956.

Définitions

DÉPENSES ET NIVEAU DE VIE

La dépense totale observée lors de l'enquête (X) diffère du niveau de vie (Q) par un écart (A) qui comprend, pour chaque ménage, deux éléments distincts. Les dépenses effectives pendant la période de l'enquête accusent, pour chaque produit, des variations aléatoires autour de leur niveau d'équilibre ; ces variations traduisent et les aléas du comportement du ménage et les irrégularités dues à l'indivisibilité habituelle de certains achats. En deuxième lieu il y a des erreurs d'observation : des oublis et des déclarations erronées, que les ménages soient de bonne foi ou non.

Ces deux éléments aléatoires varient selon les ménages ; pour un ménage donné ils varient selon les produits, mais il n'y a pas de raison pour qu'ils se compensent entre les différents produits ; la dépense totale accuse donc par rapport au niveau de vie une variation aléatoire dont on ne saurait sous-estimer l'importance. Nous signalons à cet égard la courte durée de la période d'observation d'une part, les erreurs systématiques parfois très grandes qu'accusent les dépenses de produits distincts pour l'ensemble de l'enquête d'autre part.

RESSOURCES ET NIVEAU DE VIE

Les ressources déclarées lors de l'enquête (Y) diffèrent du niveau de vie (Q) par un écart (B) qui comprend au moins trois éléments. D'abord nous avons vu que dans la définition de Y aucun compte n'a été tenu des impôts et de l'épargne, éléments par lesquels on devrait corriger les ressources pour s'approcher du niveau de vie. Ensuite, il y a des variations aléatoires des revenus perçus autour de leur niveau d'équilibre qui correspondrait au niveau de vie (1) et enfin les ressources déclarées lors de l'enquête accusent un écart par rapport aux revenus effectifs. Chacun de ces trois éléments se comporte, d'un ménage à un autre, comme une variable aléatoire ; on peut supposer que l'épargne est en moyenne positive, et que les ressources déclarées sous-estiment systématiquement le revenu réel ; ce qui importe est cependant que chaque élément a une variance et que ces variances déterminent ensemble la dispersion de leur somme B .

(1) Le lecteur averti constatera la ressemblance du niveau de vie avec le **revenu permanent** de Friedman : *A theory of the consumption function* (Chicago, 1956). Nous avons en effet emprunté à cet auteur plusieurs éléments de sa description des conditions de vie des ménages sans toutefois accepter ses hypothèses précises sur leur comportement.

Trois hypothèses

Nous introduisons maintenant trois hypothèses **fondamentales** pour l'analyse qui suit.

a) A EST INDÉPENDANT DE Q

La quantité A traduit la variation aléatoire des dépenses déclarées autour du niveau de vie ; elle est composée de sous-déclarations et de surestimations des dépenses de différents produits ainsi que des aléas dans les différents achats. Bien qu'il soit probable que la **dispersion** augmente avec le niveau absolu des dépenses, et par conséquent avec le niveau de vie, il paraît plausible qu'il n'existe aucune relation systématique entre le **montant moyen** de ces éléments et le niveau de vie.

b) B EST INDÉPENDANT DE A

Dans la mesure où les deux écarts traduisent des erreurs de déclaration ou l'effet de l'échantillonnage (courte durée de l'enquête sur les dépenses) leur indépendance semble acquise ; et nous admettons que ces éléments dominent. Quant aux variations aléatoires des ressources et des dépenses effectives, leur indépendance a été admise par Friedmann, mais cette hypothèse fait l'objet d'une vive controverse ; quoi qu'il en soit, nous tenons ces éléments pour négligeables en comparaison des autres composantes des deux écarts.

c) LA VARIANCE DE A DÉPASSE DE LOIN LA VARIANCE DE B

Nous ne pouvons pas vérifier de façon statistique cette hypothèse ; il s'agit d'une appréciation intuitive. Nous avons l'impression que dans l'écart B — différence des ressources déclarées par rapport au niveau de vie — l'élément systématique de la sous-déclaration lors de l'enquête prédomine, de sorte que le niveau moyen de B est assez élevé mais sa dispersion est peu importante. L'écart A par contre traduit les variations aléatoires des dépenses de différents produits dont la valeur et le sens sont très différents.

Le rapport entre la dépense observée et la dépense évaluée par la comptabilité nationale, calculé pour un grand nombre de produits, fait apparaître très souvent une sous-estimation ou une surestimation de plus de 10%. Si ces erreurs opposées se compensent dans une certaine mesure pour l'ensemble de l'enquête, rien ne permet de croire qu'il en est de même pour les ménages individuels. Un ménage qui sous-estime ses achats d'habillement ne surestime pas de ce fait ses dépenses d'huile ou de sucre.

(1) Friedmann estime que ces variations imprévues sont absorbées par des ajustements de l'épargne. Pour l'instant nous nous bornons à énumérer des éléments distincts sans prendre position sur la question de savoir si les variations aléatoires des revenus se traduisent par des achats « en supplément » de ceux qui sont déterminés par le niveau de vie, ou par des mouvements de trésorerie.

Le choix d'un indice du niveau de vie

Nous résumons maintenant nos définitions des différents écarts entre niveau de vie (Q), dépenses (X) et ressources (Y) :

$$(1) \quad X = Q - A \qquad (3) \quad Z = X - Y$$

$$(2) \quad Y = Q - B \qquad (4) \quad Z = B - A$$

Le résultat de la section précédente — indépendance de Z et Y — s'écrit :

$$(5) \qquad \text{Cov}(Z, Y) = 0$$

et les deux premières hypothèses de ce paragraphe sont

$$(6) \qquad \text{Cov}(A, Q) = 0$$

et

$$(7) \qquad \text{Cov}(A, B) = 0$$

Ces expressions permettent, par le calcul des variances et des covariances, de simplifier considérablement les relations entre les variables Q , Y et X . La substitution de (2) et (4) dans l'expression (5), compte tenu de (6) et (7), conduit à la formule (1)

$$(8) \qquad \text{Cov}(B, Q) = \text{Var}(B)$$

Pour les variances de X et Y nous avons en outre par la substitution de (1), (6) et (7)

$$(9) \qquad \text{Var}(X) = \text{Var}(Q) + \text{Var}(A)$$

et par la substitution de (2) et (8)

$$(10) \qquad \text{Var}(Y) = \text{Var}(Q) - \text{Var}(B)$$

Comme toute variance est positive, ce résultat indique de nouveau que la variance des dépenses dépasse celle des revenus. Le tableau 4 montre qu'en effet le rapport des deux variances tend vers une valeur d'environ 2,5 (2).

Ces résultats permettent encore de simplifier beaucoup l'expression de la corrélation des dépenses et des revenus avec le niveau de vie. Nous avons, par définition du carré du coefficient de corrélation des dépenses X et du niveau de vie Q :

$$(11) \qquad R^2_{X,Q} = \frac{\text{Cov}(X, Q)^2}{\text{Var}(X) \cdot \text{Var}(Q)}$$

(1) Substituons d'abord (2) et (4) en (5) :

$$\text{Cov}\{(B - A), (Q - B)\} = \text{Cov}(B, Q) - \text{Var}(B) + \text{Cov}(A, Q) + \text{Cov}(A, B) = 0$$

Par (6) et (7) les deux derniers termes sont nuls, donc :

$$\text{Cov}(B, Q) - \text{Var}(B) = 0$$

d'où l'expression (8).

(2) Ce rapport dépasse celui des moyennes : même si nous considérons les coefficients de variation, la dispersion des dépenses est plus grande que celle des revenus. Le même phénomène s'observe dans l'enquête effectuée aux U.S.A. en 1950. Voir ouvrage cité plus haut, p. 86-91.

Nous substituons (1) dans le numérateur et, après développement, (6); nous simplifions de dénominateur par (9), ce qui conduit à

$$(12) \quad R^2_{X,Q} = \frac{\text{Var}(Q)}{\text{Var}(Q) + \text{Var}(A)}$$

De même la corrélation entre ressources déclarées et niveau de vie s'écrit

$$(13) \quad R^2_{Y,Q} = \frac{\text{Cov}(Y,Q)^2}{\text{Var}(Y) \cdot \text{Var}(Q)}$$

Le numérateur est simplifié par la substitution successive de (2) et (8), le dénominateur par (10), de sorte qu'on obtient

$$(14) \quad R^2_{Y,Q} = \frac{\text{Var}(Q) - (\text{Var} B)}{\text{Var}(Q)}$$

Une comparaison de (12) et (14) montre déjà que notre troisième hypothèse $\text{Var}(A)$ plus grande que $\text{Var}(B)$ implique que $R^2_{Y,Q}$ est probablement supérieur à $R^2_{X,Q}$. Plus précisément, le rapport des deux coefficients de corrélation est, par l'emploi de (9) et (10),

$$(15) \quad \frac{R^2_{Y,Q}}{R^2_{X,Q}} = \frac{\text{Var}(X) \cdot \text{Var}(Y)}{\text{Var}(Q)^2}$$

et si nous admettons que $\text{Var}(X) \approx 2,5 \text{Var}(Y)$ il suffit que

$$(16) \quad \frac{\text{Var}(Y)}{\text{Var}(Q)} \geq \sqrt{\frac{1}{2,5}}$$

pour que la corrélation des ressources avec le niveau de vie dépasse celle des dépenses. Cette dernière condition est encore équivalente à (1)

$$(17) \quad \frac{\text{Var}(A)}{\text{Var}(B)} > 1,55$$

Nous n'hésitons pas à admettre que cette condition est satisfaite.

Dans ces conditions le revenu déclaré est plus étroitement lié au niveau de vie que les dépenses, c'est-à-dire que pour des ménages individuels le revenu est un meilleur indice du niveau de vie. Cette conclusion ne contredit pas le fait qu'en moyenne le revenu soit plus éloigné du niveau de vie que la dépense ; c'est seulement parce que la sous-estimation du revenu est relativement **uniforme** qu'il est une meilleure indication que les dépenses, plus fortement dispersées. Cette conclusion n'est donc qu'en apparence paradoxale ; elle justifie notamment la procédure employée intuitivement lors de l'exploitation de l'enquête qui consiste à grouper les ménages selon

(1) Pour le montrer on utilise les expressions (9) et (10) et l'approximation $\text{Var}(X) \approx 2,5 \text{Var}(Y)$ d'abord ; ensuite on introduit l'inégalité (16).

leurs revenus et à estimer ensuite leur niveau de vie d'après leur dépense totale (1).

TABLEAU 4

Rapport des variances des dépenses et des ressources dans 23 sous-groupes (1)

Valeur de Var (X)/Var (Y)	Nombre de sous-groupes	Nombre de ménages dans ces sous-groupes
0,3—0,59	1	10
0,6—1,19	4	113
1,2—2,39	7	326
2,4—4,79	9	437
4,8—9,59	2	41

(1) Dans deux des 30 sous-groupes du tableau 2 nous avons enlevé une observation aberrante avant de calculer les variances et leur rapport ; ensuite nous n'avons pris en considération que les sous-groupes de 10 ménages ou plus, des effectifs plus faibles ne permettant pas l'estimation des variances. La dernière colonne montre d'ailleurs que ce sont encore les sous-groupes de faibles effectifs qui se trouvent aux extrémités de la distribution du rapport des variances.

Les résultats de l'analyse dépendent des trois hypothèses que nous avons introduites. Ces hypothèses ne sont pas pleinement justifiées — aucune hypothèse ne l'est — et on peut nous reprocher leur choix arbitraire. Il est certain qu'il s'agit de raccourcis simplificateurs de la description, c'est-à-dire d'approximations, plutôt que d'hypothèses au sens propre du terme. La supposition que A et B sont indépendants ne signifie pas que dans la population la covariance de A et B soit rigoureusement nulle ; mais elle signifie que nous jugeons dans le cas présent la relation entre A et B assez faible pour être négligée. Il en est de même pour les autres hypothèses. On pourrait évidemment en proposer de différentes qui conduiraient à une autre interprétation ; rien ne s'y oppose. Nous croyons cependant que les hypothèses retenues ici ont l'avantage de la simplicité, et qu'elles facilitent l'interprétation des données dont nous disposons sans trop déformer leur caractéristiques.

4. CONSOMMATION ALIMENTAIRE ET NIVEAU DE VIE

A titre d'exemple nous appliquons les résultats précédents à une analyse du comportement des consommateurs qui porte sur la dépense alimentaire. Dans les analyses de ce genre on suppose qu'il existe des *courbes d'Engel*, c'est-à-dire des relations entre la consommation des différents produits et le niveau de vie des ménages. La dépense ou la consommation d'un produit i , C_i , est donc déterminée par l'équation

$$(18) \quad C_i = \lambda_i Q + x_i + U_i$$

où λ_i et x_i sont des paramètres qui caractérisent le produit i et U_i est une variable aléatoire qui traduit les erreurs d'observations ainsi que les aléas

(1) Voir G. ROTTIER, Niveau de vie et consommation de la population non agricole, *Consommation*, n° 3, 1959.

réels de la consommation du produit i . Nous supposons que cette erreur est indépendante de la variable explicative,

$$(19) \quad \text{Cov}(U_i, Q) = 0$$

et nous cherchons en tenant compte de cette hypothèse classique à obtenir des estimations de λ et α qui décrivent la loi de la consommation du produit i (1).

Hypothèses complémentaires

Nous supposons qu'il existe une courbe d'Engel (18) pour chacun des t produits ($i = 1 \dots t$) qui figurent dans les budgets des familles. La somme des dépenses ainsi définies correspond évidemment à la dépense totale X ,

$$(20) \quad X = \sum_i c_i = Q \sum_i \lambda_i + \sum_i \alpha_i + \sum_i U_i$$

D'autre part, nous avons défini le niveau de vie comme le niveau de dépense d'équilibre voulu par le ménage, ce qui nous permet d'identifier cette variable avec la somme des parties systématiques de chaque dépense observée, ou

$$(21) \quad Q = Q \sum_i \lambda_i + \sum_i \alpha_i \quad (2)$$

Les deux équations (20) et (21) conduisent immédiatement à

$$(22) \quad \sum_i U_i = -A$$

La somme des aléas des dépenses individuelles est égale à l'écart entre dépense totale et niveau de vie.

Nous complétons les relations entre dépense d'un produit et dépense totale par l'hypothèse que, tout au moins pour le produit i que nous étudions, c'est-à-dire la consommation alimentaire, l'erreur U_i est indépendante de l'ensemble des erreurs des autres dépenses, c'est-à-dire :

$$(23) \quad \text{Cov}(U_i, -A - U_i) = 0$$

d'où, par (22),

$$(24) \quad \text{Cov}(U_i, A) = -\text{Var}(U_i)$$

(1) La forme linéaire de la courbe d'Engel — équation (18) — facilite non pas seulement l'exposé de la méthode mais simplifie encore considérablement les calculs : nous y resterons fidèle dans la suite. En fait nous ignorons quelle est la forme des courbes d'Engel qui correspondrait le mieux aux observations individuelles. Il s'agit donc d'une approximation d'autant plus justifiée que nous ajustons des courbes distinctes à chacun des sous-groupes définis dans le tableau 1. Dans ces sous-groupes le coefficient de variation des revenus et des dépenses est de l'ordre de 0,3 à 0,5, ce qui correspond — compte tenu des effectifs très faibles de ces groupes — à un champ de variation de un à trois environ. Les données groupées de toute la population non agricole qui ont servi dans l'étude antérieure de G. Rottier à l'ajustement de courbes d'Engel accusaient des variations de un à dix et davantage. (Voir *Consommation*, n° 3, 1959.)

(2) Ceci suppose évidemment $\sum \lambda_i = 1$ et $\sum \alpha_i = 0$. Voir D. G. CHAMPERNOWNE et G. D. N. WORSWICK, Notes on the adding-up criterion, *Review of Economic Studies*, vol. 22 (1954), p. 57.

Si nous introduisons cette hypothèse pour chacun des t produits, elle serait équivalente à

$$(25) \quad \text{Cov}(U_i, U_j) = 0 \quad \text{pour tout} \quad i \neq j (i, j = 1 \dots t)$$

et sa validité dépendrait alors étroitement de la définition des produits distincts. Si les produits sont définis par une classification très détaillée qui distingue par exemple entre cigarettes et tabac, entre viande de bœuf et viande de veau, il est probable que les U ne sont pas tous indépendants, au moins dans la mesure où ils traduisent des aléas réels et non pas des erreurs d'observations. Ces inter-relations disparaissent si nous groupons substitués et compléments ensemble dans des groupes de produits quasi indépendants les uns des autres.

Une fois que nous avons décomposé l'écart A en une somme des U_i , il est normal de compléter l'indépendance de A et de B , l'écart entre ressources et niveau de vie, par

$$(26) \quad \text{Cov}(U_i, B) = 0$$

Logiquement c'est cette dernière hypothèse qui devait précéder celle de l'indépendance de A et B que nous avons introduite plus haut. La justification est la même : nous n'invoquons pas pour autant l'indépendance entre les variations des dépenses et celles des revenus comme le fait Friedmann, mais le rôle dominant des erreurs d'observation. Il faut souligner qu'ici encore il s'agit dans l'équation (26) d'une approximation qui nous amène à considérer la covariance comme négligeable, plutôt que d'une hypothèse rigoureuse.

Trois estimations du paramètre λ

Il existe trois estimations simples du paramètre λ de l'équation (18) : estimations par les moindres carrés, avec soit la dépense (X), soit les ressources observées (Y) comme variable explicative tenant lieu du niveau de vie, et la méthode des variables instrumentales proposée récemment par Liviatan (1). Nous considérons les caractéristiques de ces estimateurs dans le cadre très particulier de l'analyse précédente ; encore, nous bornons-nous à la considération de leur limite en probabilité (notée : $Plim$). L'utilité de cette notion est contestable, compte tenu des effectifs très faibles des sous-groupes sur lesquels les ajustements seront faits ; malheureusement la formulation de la distribution des estimateurs pour des petits échantillons présente des difficultés insurmontables. Nous avons renoncé pour la même raison à l'évaluation de la variance des estimateurs.

AJUSTEMENT I

Le premier ajustement, classique, consiste à utiliser la dépense comme indice du niveau de vie et à estimer λ par les moindres carrés. L'estimation de $\lambda, \hat{\lambda}_1$, est donc définie par

$$(27) \quad \hat{\lambda}_1 = \frac{\sum cx}{\sum x^2}$$

(1) Voir N. LIVIATAN, Errors in Variables and Engel Curve Analysis, *Econometrica*, vol. 29 (1961), p. 336.

où les c et x sont les écarts de C et X à leurs moyennes. La limite en probabilité de cette expression est

$$(28) \quad \text{Plim}(\hat{\lambda}_1) = \frac{\text{Cov}(C, X)}{\text{Var}(X)}$$

Dans le numérateur, nous substituons à C , d'après (18), (1)

$$(29) \quad C = \lambda Q + U$$

et à X , d'après (1), nous substituons $X = Q - A$; ensuite nous utilisons (19) et (24) pour simplifier le résultat. Le dénominateur est déjà donné en équation (9). Nous obtenons ainsi

$$(30) \quad \text{Plim}(\hat{\lambda}_1) = \frac{\lambda \text{Var}(Q) + \text{Var}(U)}{\text{Var}(Q) + \text{Var}(A)}$$

L'estimateur $\hat{\lambda}_1$ est donc biaisé dans la mesure où $\text{Var}(U) \neq \lambda \text{Var}(A)$, ce qui dépend et de l'importance du produit parmi les composantes aléatoires de la dépense [$\text{Var}(U)/\text{Var}(A)$], et de la propension marginale à la consommation du produit (λ). Comme ces deux caractéristiques tendent à varier dans le même sens il est difficile de deviner dans quel sens le biais opère.

AJUSTEMENT 2

Le deuxième ajustement est l'estimation par les moindres carrés avec le revenu comme indice du niveau de vie. L'estimation est donc

$$(31) \quad \hat{\lambda}_2 = \frac{\sum cy}{\sum y^2}$$

d'où

$$(32) \quad \text{Plim}(\hat{\lambda}_2) = \frac{\text{Cov}(C, Y)}{\text{Var}(Y)}$$

Nous simplifions le numérateur de nouveau par la substitution de (29) et (2), et par l'emploi de (8) et (26); pour le dénominateur nous employons (10). On obtient :

$$(33) \quad \text{Plim}(\hat{\lambda}_2) = \frac{\lambda \{ \text{Var}(Q) - \text{Var}(B) \}}{\text{Var}(Q) - \text{Var}(B)} = \lambda$$

de sorte que cet estimateur n'est pas biaisé.

AJUSTEMENT 3

Le biais de l'estimateur du premier ajustement est un phénomène très général, quoique sa forme varie évidemment selon les hypothèses particulières qu'on impose; en dernière analyse il est dû au fait que les erreurs U de la dépense du produit que l'on étudie sont reportées sur la

(1) Ici comme dans la suite nous enlevons l'indice i indicatif du produit puisque toute la discussion porte sur un produit donné.

dépense totale. Pour remédier à ce biais, que Summers a signalé le premier⁽¹⁾, Liviatan a proposé d'employer le revenu comme variable instrumentale dans l'estimation de la relation entre dépense du produit étudié et dépense totale. L'estimateur est dans ces conditions

$$(34) \quad \hat{\lambda}_3 = \frac{\sum cy}{\sum xy}$$

d'où

$$(35) \quad \text{Plim}(\hat{\lambda}_3) = \frac{\text{Cov}(C, Y)}{\text{Cov}(X, Y)}$$

Le numérateur de cette expression est le même qu'en (32), pour le dénominateur nous substituons (1) et (2) et utilisons ensuite (6), (7) et (8), ce qui conduit à

$$(36) \quad \text{Plim}(\hat{\lambda}_3) = \frac{\lambda \{ \text{Var}(Q) - \text{Var}(B) \}}{\text{Var}(Q) - \text{Var}(B)} = \lambda$$

de sorte que $\hat{\lambda}_3$ n'est pas biaisé.

Nous disposons ainsi de deux estimateurs non-biaisés ($\hat{\lambda}_2$ et $\hat{\lambda}_3$) et d'un estimateur ($\hat{\lambda}_1$) qui est biaisé dans la mesure où $\text{Var}(U) \neq \lambda \text{Var}(A)$. Dans l'application à la consommation alimentaire, il se peut encore fort bien que l'égalité $\text{Var}(U) = \lambda \text{Var}(A)$ soit approximativement satisfaite, ce qui éliminerait le biais ; celui-ci est de toute façon assez vague pour que nous hésitions à nous prononcer sur sa direction. Pour choisir entre les trois estimateurs on préférerait comparer leurs dispersions, mais l'évaluation de leurs variances, telles qu'elles sont définies dans le système que nous considérons, exige et des approximations et des hypothèses supplémentaires avant que n'apparaisse une différence nette en faveur de l'une ou de l'autre.

Application à la dépense alimentaire

Nous pouvons toutefois considérer les résultats des trois ajustements pour la consommation alimentaire⁽²⁾. Ces ajustements ont été faits sur chacun des sous-groupes de la classification du tableau que nous avons déjà employés auparavant. Cette classification permet éventuellement à l'hétérogénéité du comportement selon la catégorie sociale ou le type de famille d'apparaître dans les résultats. Nous avons cependant omis les deux ménages aberrants d'une part et éliminé d'autre part tout sous-groupe de moins de 10 ménages. Il reste 23 groupes⁽³⁾ pour lesquels le tableau 5 indique les trois estimations dans les colonnes (6), (8) et (10). Nous avons ajouté les corrélations entre C et X ou Y dans les deux premiers ajustements ainsi que le point moyen — consommation alimentaire et dépense totale, les deux en moyenne — auquel il convient d'évaluer l'élasticité de la consommation alimentaire, définie par

(1) Voir R. SUMMERS, A note on least squares bias in household expenditure analysis, *Econometrica*, vol. 27 (1959), p. 268.

(2) Dépense alimentaire y compris boissons et autoconsommation ou autofournitures'il y a lieu (voir p. 63).

(3) Ce sont les mêmes sous-groupes qui figurent dans le tableau 4, p. 23

TABLEAU 5. — Estimations de la relation entre consommation alimentaire et niveau de vie dans 23 sous-groupes

28

SOUS-GROUPE CODE		NOMBRE DE MÉNAGES N	POINT MOYEN X (10 NF par an)		AJUSTEMENT 1		AJUSTEMENT 2		AJUSTEMENT 3
C.S.	Ty		\bar{C}	\bar{X}	$\hat{\lambda}_1$	R_{cx}	$\hat{\lambda}_2$	R_{cy}	$\hat{\lambda}_3$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	1	12	3 774	14 678	0,035	0,293	0,213	0,730	0,247
	2.	13	4 546	13 989	0,124	0,496	0,056	0,250	0,093
	3	8
	4	9
	5	10	4 606	14 468	0,218	0,730	0,154	0,726	0,265
	6	26	5 598	16 744	0,141	0,328	0,049	0,180	0,144
2	1	11	3 257	10 386	— 0,037	0,309	0,104	0,401	0,306
	2	28	4 192	10 341	0,316	0,618	0,383	0,671	0,410
	3.	9
	4	12	3 968	10 114	0,131	0,620	0,165	0,427	0,173
	5	14	4 277	12 555	0,023	0,810	0,003	0,041	0,007
	6	15	5 247	15 219	0,114	0,325	0,085	0,317	0,189
3	1	44	3 603	10 044	0,194	0,538	0,085	0,212	0,104
	2	64	3 776	10 528	0,099	0,733	0,093	0,516	0,129
	3	26	3 579	8 393	0,056	0,392	0,130	0,224	0,195
	4	33	4 353	10 699	0,193	0,471	0,264	0,479	0,358
	5	29	4 223	11 918	0,144	0,529	0,216	0,558	0,199
	6	47	4 732	11 474	0,115	0,574	0,236	0,483	0,251
4	1	99	3 160	7 401	0,082	0,229	0,049	0,114	0,136
	2	99	3 571	7 159	0,330	0,519	0,226	0,348	0,278
	3	59	3 630	7 863	0,216	0,553	0,144	0,235	0,140
	4	49	3 921	7 643	0,320	0,563	0,293	0,415	0,315
	5	59	4 025	8 673	0,184	0,586	0,155	0,308	0,183
	6	157	4 448	8 690	0,391	0,536	0,291	0,348	0,352
5	1	7
	2	11	3 873	9 428	0,247	0,581	0,063	0,128	0,066
	3	5
	4	9
	5	7
	6	10	4 817	10 712	0,432	0,281	0,306	0,631	1,802

(37)

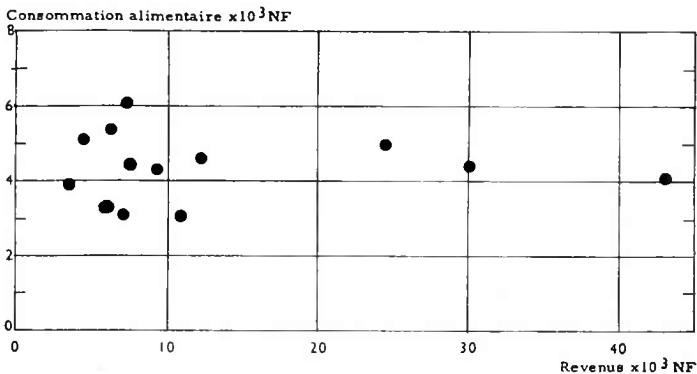
$$\hat{\eta} = \frac{\bar{X}}{C} \hat{\lambda}$$

Dans la mesure où \bar{X} sous-estime \bar{Q} , valeur inconnue, cette transformation tend vers une sous-estimation de η ; mais nous ne pouvons faire mieux.

Avant d'aborder l'examen des estimations, nous notons que la qualité des ajustements, telle que les coefficients de corrélation la traduisent, est beaucoup plus faible que dans les analyses habituelles. Cela est dû en grande partie au fait que nous considérons des observations individuelles et non des données groupées. Ces observations individuelles sont très fortement dispersées, comme en témoignent les graphiques 7 et 8 qui résument

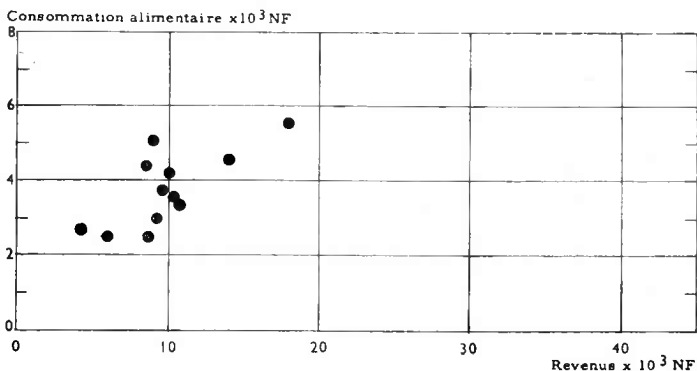
GRAPHIQUE 7

Ressources et consommation alimentaire
(14 ménages de gros indépendants, composés de deux adultes et deux enfants de 6 à 13 ans)



GRAPHIQUE 8

Ressources et consommation alimentaire
(12 ménages de gros indépendants, composés de deux adultes et un enfant de moins de 6 ans)



deux cas extrêmes — le sous-groupe à corrélation la plus faible (graphique 7) et celui à corrélation la plus forte (graphique 8) entre consom-

mation alimentaire et revenus. Ces nuages de points invitent, par leur extrême dispersion, à beaucoup de prudence dans l'interprétation des résultats de l'analyse statistique ; on peut s'étonner que ces résultats soient tout de même assez cohérents.

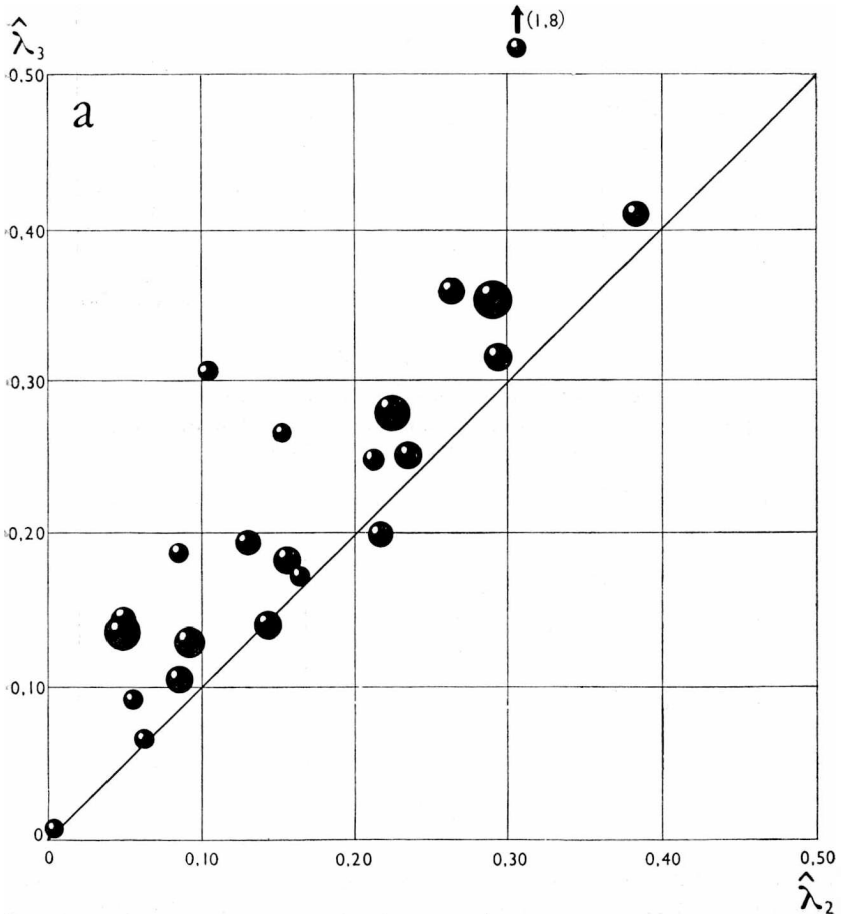
Nous examinons ensuite les rapports entre les trois estimations de λ pour chacun des sous-groupes. Les graphiques 9a-9c montrent leurs interrelations. Bien que nous ayons renoncé à l'évaluation des écarts-types, on peut préjuger qu'ils varient d'un sous-groupe à l'autre de façon inversement proportionnelle à la racine carrée du nombre d'observations. Pour illustrer les différences qui existent dans la précision des estimations entre les sous-groupes, chaque estimation est représentée par un point dont la surface est proportionnelle à \sqrt{N} .

Les résultats sont inattendus. Pour deux estimateurs prétendus non-biaisés, le graphique 9 (a) montre que les estimations sont étroitement

GRAPHIQUES 9

Rapports entre les trois estimations de λ

Graphique 9 a — $\hat{\lambda}_2$ et $\hat{\lambda}_3$



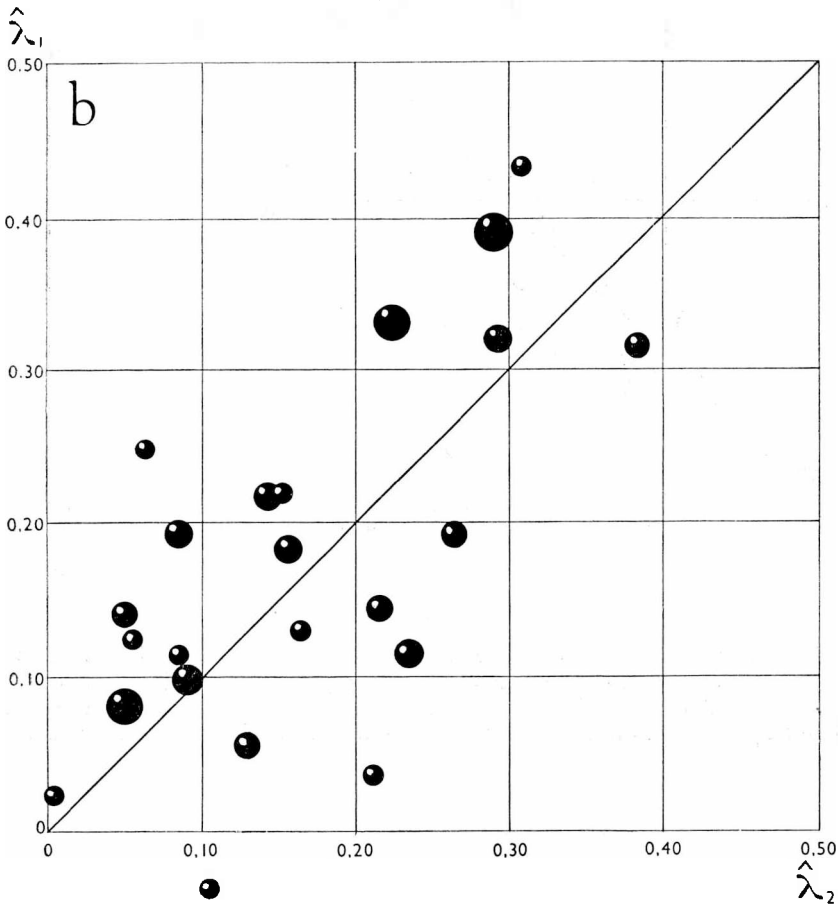
liées mais qu'il y a une tendance assez nette pour que $\hat{\lambda}_3$ dépasse $\hat{\lambda}_2$. Une comparaison des définitions des deux estimateurs, équations (31) et (34), montre que cette tendance traduit le fait que pour la plupart des sous-groupes Σy^2 dépasse Σxy or le rapport de ces deux quantités détermine l'estimation \hat{a} ,

$$\hat{a} = \frac{\Sigma xy}{\Sigma y^2}$$

que nous avons considérée en section 2 et dont les valeurs sont indiquées en colonne 4 du tableau 2. Cette estimation est en effet généralement inférieure à 1 bien que nous ayons conclu de son analyse statistique qu'une valeur $a=1$ pouvait être retenue. La concentration des valeurs observées en-dessous de un est troublante et mérite un examen plus poussé ; pour l'instant nous nous bornons cependant à la signaler sans remettre en cause l'analyse précédente.

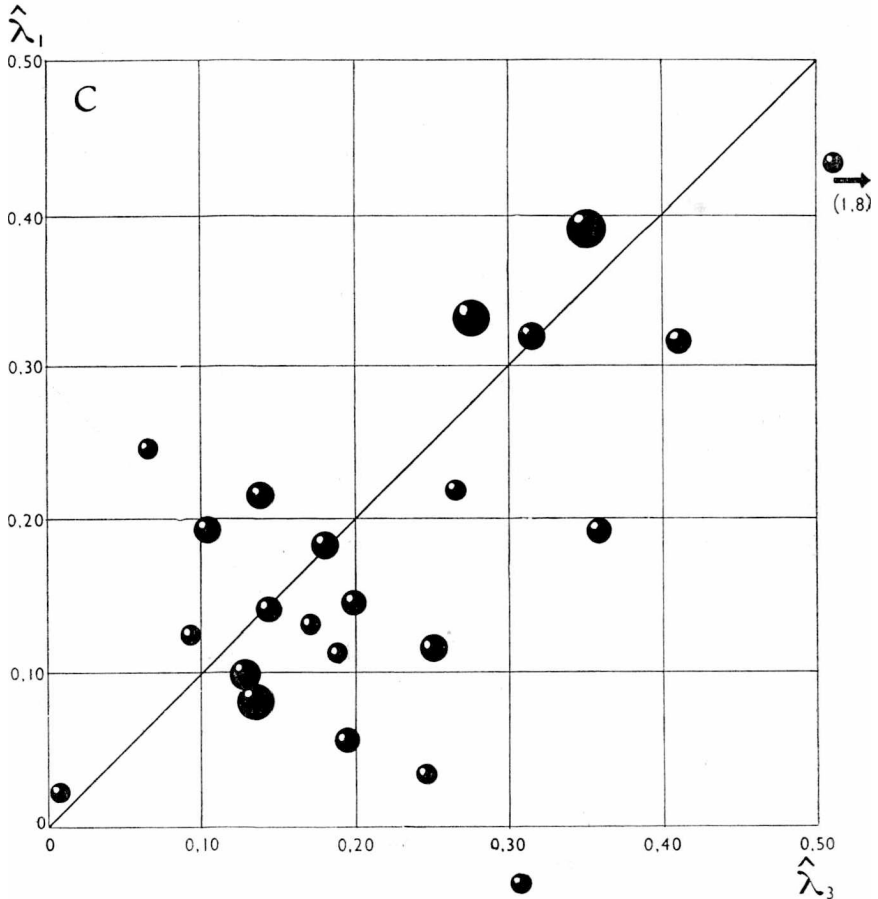
Les graphiques 9 (b) et 9 (c) représentent à leur tour les rapports de $\hat{\lambda}_1$,

Graphique 9 b — $\hat{\lambda}_1$ et $\hat{\lambda}_2$



estimateur peut-être biaisé, avec $\hat{\lambda}_2$ et $\hat{\lambda}_3$. Si la dispersion est plus grande qu'en 9(a), aucune tendance systématique en faveur de l'une ou de l'autre estimation n'apparaît. Nous concluons donc que, dans le cas présent, le biais de $\hat{\lambda}_1$ est négligeable, ce qui veut dire que le rapport des variances

Graphique 9 c — $\hat{\lambda}_1$ et $\hat{\lambda}_3$



$\text{Var}(U)/\text{Var}(A)$ est approximativement égal au coefficient λ dont $\hat{\lambda}_2$ et $\hat{\lambda}_3$ donnent des estimations de l'ordre de 0,15 à 0,25 environ. Il est clair que cette conclusion est seulement valable pour la consommation alimentaire et ne s'applique pas nécessairement aux courbes d'Engel d'autres produits.

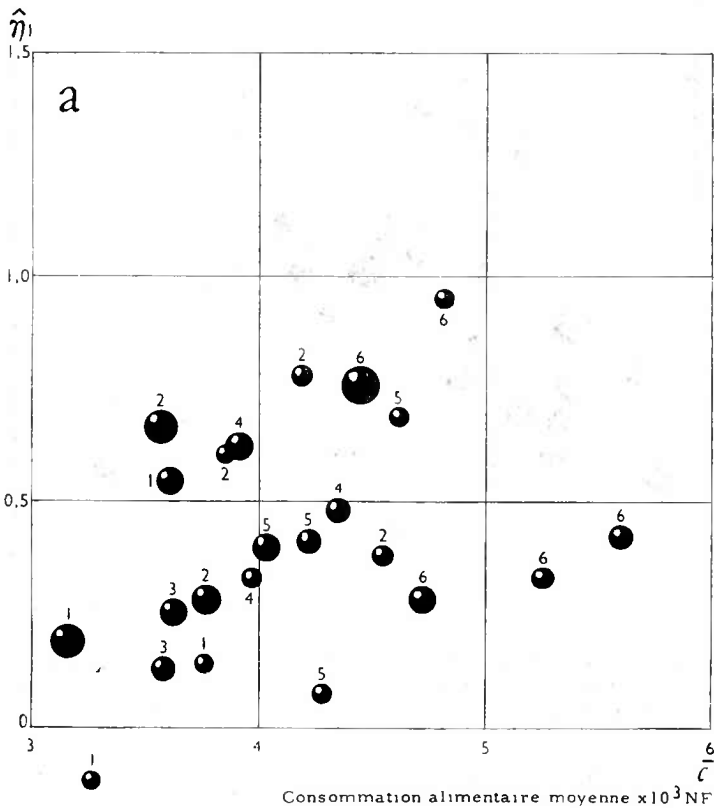
Les variations de l'élasticité

L'examen des trois estimations de λ nous permet ainsi de tenir le biais de λ_1 pour négligeable, mais il ne permet pas de choisir entre elles et nous devons pour l'instant admettre que nous disposons de trois estimations valables de λ pour chaque sous-groupe. La dispersion de ces estimations est grande, aussi bien entre sous-groupes qu'entre estimateurs ; du point de vue statistique cette dispersion est décevante, mais du point de vue économique elle peut avoir un grand intérêt dans la mesure où elle traduit une variation systématique du comportement des consommateurs suivant les caractéristiques des ménages. Ainsi doit-on s'attendre à des variations dans les courbes d'Engel selon les catégories socio-professionnelles — à une influence de la composition de la famille — enfin, à un infléchissement de la courbe d'Engel tel que l'élasticité de la consommation alimentaire décroisse avec le niveau de cette consommation. Afin de faciliter l'inter-

GRAPHIQUES 10

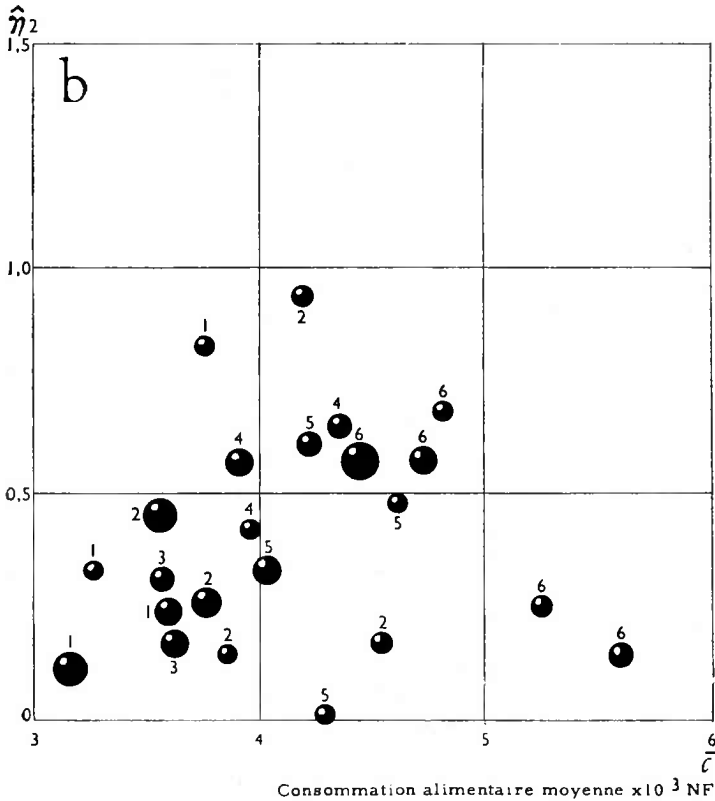
Estimation de l'élasticité η et consommation alimentaire moyenne
(23 sous-groupes de ménages)

Graphique 10-a



prétation économique de nos résultats, nous avons transformé les estimations de λ en trois estimations de l'élasticité-revenu de la consommation alimentaire, définies par l'équation (37) (1). L'examen de ces élasticités ne confirme aucune des hypothèses économiques que nous venons d'énumérer. Il existe bien une variation des élasticités selon les catégories socio-professionnelles, puisque les élasticités sont très fortement dispersées, mais cette variation n'est pas systématique. Il en est de même pour l'influence de la composition de la famille et pour la diminution de la valeur des élasticités avec le niveau croissant de la consommation alimentaire. Le lecteur lui-même fera ces constatations en examinant les graphiques 10. Dans ces graphiques sont mis en relation l'estimation de l'élasticité et le niveau moyen de la consommation alimentaire de chacun des trois estimateurs séparément ; la représentation des sous-groupes de différentes tailles est la même qu'au graphique 9 (2). La variation est grande mais nulle relation systématique n'apparaît. On peut croire que cela est dû au

Graphique 10-b



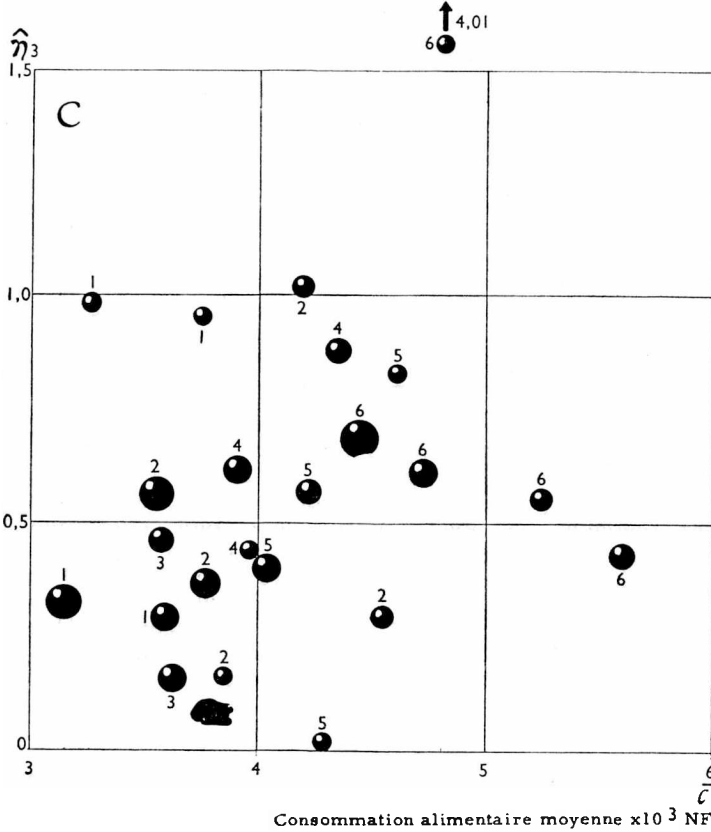
fait que nous avons mélangé tous les types de famille, sans tenir compte de ce facteur par l'introduction d'unités de consommation ou par un autre procédé du même genre. Or il suffit de relier les points du même type de

(1) Comme le rapport \bar{X}/\bar{C} varie assez peu d'un sous-groupe à l'autre, l'essentiel de la discussion suivante s'applique aussi bien aux estimations de λ qu'aux estimations de η .

(2) Voir plus haut p. 30.

famille pour constater que même parmi les groupes qu'ils représentent, la dispersion des élasticités reste entière et qu'aucune variation systématique selon le niveau moyen de la consommation n'apparaît. On notera que les familles nombreuses — types 4, 5 et 6 — se trouvent généralement vers la droite, c'est-à-dire que leur consommation alimentaire est plus élevée que pour les autres groupes ; cependant leurs dépenses totales (et leurs revenus) sont également plus élevées, et il ne s'agit guère d'un effet propre à la taille du ménage.

Graphique 10-c



Si par ailleurs on examine les moyennes de la dépense alimentaire et de la dépense totale dans les différents sous-groupes ⁽¹⁾, on constate qu'elles s'alignent indépendamment du type de famille autour d'une droite avec un coefficient angulaire d'environ 0,2 ce qui correspond pour l'ensemble à une élasticité d'environ 0,5 ; la seule exception — et encore elle est peu nette — concerne des sous-groupes du type 6, qui ont une consommation légèrement plus forte. A cette exception près, l'examen détaillé des

(1) Nous ne publions pas le graphique correspondant mais les données figurent dans les colonnes 4 et 5 du tableau 5.

courbes d'Engel pour la dépense alimentaire ne confirme pas les hypothèses de base introduites lors de l'exploitation antérieure de l'enquête, à savoir l'introduction d'une échelle d'unités de consommation et le choix d'une courbe semi-logarithmique. Cela ne veut pas dire que nos données condamnent ces hypothèses, qui correspondent après tout au bon sens ; rien ne s'oppose à ce qu'on les impose d'emblée à l'analyse, mais les effets réels auxquels elles correspondent n'apparaissent pas nettement dans les résultats que nous présentons ici. Il en irait peut être différemment si le champ de variation des types de famille était plus large et si les estimations étaient plus précises (1).

TABLEAU 6

Estimations de λ dans 23 sous-groupes (2)

ESTIMATION	NOMBRE DE SOUS-GROUPES	NOMBRE DE MÉNAGES dans ces sous-groupes
$\leq 0,099$	17	550
0,100—0,199	26	888
0,200—0,299	14	674
0,300—0,399	10	621
$> 0,400$	3	48

Quoi qu'il en soit, nous croyons qu'il est justifié de considérer toutes les estimations de λ dont nous disposons comme des estimations valables d'un seul paramètre (3), tout au moins pour le sous-échantillon étudié — déjà assez homogène — et pour la consommation alimentaire, seul produit étudié. Le tableau 6 montre une fois de plus que la dispersion est très grande ; la valeur centrale se situe aux environs de $\lambda = 0,175$, ce qui correspond pour l'ensemble des sous-groupes à une élasticité moyenne d'environ 0,45 ; cette élasticité est atteinte à un niveau de dépenses de 9 600 NF. Ce chiffre résume donc le résultat concret de l'analyse ; signalons qu'il correspond exactement à l'élasticité obtenue précédemment par de tout autres méthodes (4).

(1) Les groupes de ménages que nous considérons sont d'effectifs généralement faibles, d'où le manque de précision des estimations ; ils comprennent tous deux adultes avec un ou plusieurs enfants d'où la faible variation du nombre d'unités de consommation. Ces deux phénomènes rendent difficile l'appréciation des effets dont nous parlons.

(2) Chaque sous-groupe fournit trois estimations et a donc été compté trois fois dans chaque colonne de ce tableau.

(3) Voir note 1, p. 34.

(4) Voir **Consommation**, n° 3, 1959 ; élasticité déduite des deux lignes « Achats alimentaires » et « Autoconsommation et autofourniture » du tableau A, p. 38, cette élasticité est évaluée au point moyen : dépenses par U.C. = 3 300 NF. Le chiffre de 9 600 NF par ménage de notre sous-échantillon correspond à 3 500 NF par U.C.