

PRÉCISION DES ESTIMATIONS ET DURÉE DE L'ENREGISTREMENT DES COMPTES DANS LES ENQUÊTES SUR LES BUDGETS FAMILIAUX

par

Nicole TABARD

Un certain nombre d'études ont été faites concernant les estimations obtenues à partir de relevés de comptes portant sur des périodes variables. En particulier on s'est intéressé à la question de savoir si les dépenses enregistrées les premiers jours d'enquête étaient surestimées par rapport aux autres ou si un effet de lassitude entraînait une sous-estimation des dépenses au bout d'un certain temps. Le présent article ne s'intéresse pas aux moyennes mais à la dispersion des dépenses, la variance ou l'écart-type mesurant cette dispersion. La précision des estimations au niveau individuel est très importante pour une analyse en profondeur du comportement du consommateur. Notre but ici est de voir s'il y a une amélioration sensible de la mesure et comment elle évolue en fonction du nombre de semaines de relevés des comptes.

Du fait de leur coût élevé les enquêtes sur les budgets familiaux sont rarement limitées quant à leur objectif ; un même échantillon vaste et représentatif fournira des statistiques pour la comptabilité nationale et servira à l'élaboration de modèles économétriques de comportement des ménages. Ces observations sont obtenues par enregistrement des dépenses courantes pendant une période courte et par interview pour les dépenses moins fréquentes. Elles sont peu précises au niveau individuel. Si les erreurs de mesure peuvent être compensées par le nombre des observations pour le calcul de moyennes, elles risquent de nuire à une analyse reposant sur

l'étude des dispersions. La complexité des phénomènes qui déterminent les comportements impose de travailler sur une base statistique soigneusement construite. Pour obtenir des données plus précises et plus stables on a le plus souvent recours au groupement des observations où les erreurs s'estompent en même temps que les aléas. En fait, ces groupements entraînent la perte d'une information utile à l'explication et reposent sur une part d'arbitraire.

La question qui se pose est celle-ci : doit-on observer un petit nombre de ménages pendant une longue période ou un grand nombre pendant une période courte? La réponse à cette question dépend du problème que l'on veut résoudre.

— S'agit-il d'**estimer** les paramètres de lois connues, incluant des facteurs décidés à l'avance.

— Ou bien s'agit-il de **repérer** l'influence de facteurs généralement omis dans les analyses classiques, donc d'améliorer notre connaissance du comportement?

Dans le premier cas, où l'on s'intéresse aux paramètres et à leur précision on aura probablement intérêt à analyser des échantillons assez vastes observés sur période assez courte ; d'une part parce que la précision des estimateurs dépend du nombre d'observations, d'autre part parce que l'observation sur longue période demandant un effort assidu aux enquêtés, entraîne finalement une **sélection** des ménages d'où un biais possible des estimateurs.

Dans le second cas, c'est l'importance et la signification des variances à expliquer qui retient notre attention. Les corrélations observées dans les analyses classiques entre revenu et dépenses sont médiocres ; même si des analyses plus fines parviennent à améliorer l'explication nous devons nous attendre à ce que tous les facteurs qu'il nous est possible de repérer n'expliquent qu'une faible proportion des variances observées. De telles analyses exigent donc des observations très précises. La participation astreignante alors demandée aux enquêtés augmente les risques de refus. Cependant la méthode des enquêtes intensives paraît la seule qui nous permette de remplacer la description par l'analyse.

Nous étudions ici l'influence de la durée de l'enregistrement des comptes sur la précision des estimations obtenues. Cette étude repose sur une enquête pilote exécutée par le **Credoc** pour l'**Union nationale des Caisses d'Allocations Familiales**. La période d'une ou deux semaines d'enregistrement des dépenses fréquentes utilisée dans les enquêtes françaises semble brève si l'on veut obtenir des données détaillées valables au niveau individuel et le procédé de l'interview pour les dépenses plus rares conduit certainement à des oublis ou des erreurs, entre autre pour les sommes consacrées à l'utilisation des services. L'enquête dont les résultats sont analysés ci-après visait à étudier la possibilité d'un allongement de la période d'enregistrement des comptes. A cette fin les ménages ont été interrogés les uns pendant une semaine, d'autres pendant deux semaines, d'autres pendant quatre semaines et d'autres enfin pendant six semaines.

Les caractéristiques de l'échantillon analysé sont données dans la première partie de cet article, ainsi que l'évolution des taux de refus en fonction du nombre de semaines de relevé de comptes. La deuxième partie

s'intéresse à la dépense alimentaire dans son ensemble et compare la variance entre ménages à la variance entre les semaines pour chaque ménage ou variance intra-ménage. La troisième partie est consacrée à une analyse de la précision des estimations de chacun des produits alimentaires en fonction de sa fréquence d'achat. Les dépenses non-alimentaires ne sont pas étudiées ici ; la méthode d'enregistrement des comptes au jour le jour sur période relativement courte ne peut donner une bonne mesure des dépenses à fréquence rare. Ces dépenses ont cependant été enregistrées comme les autres, mais à seule fin d'éprouver la persévérance des enquêtés en vue de l'enquête définitive ultérieure.

I. POPULATION ET VARIABLES ÉTUDIÉES

L'enquête dont les résultats sont analysés ici a été réalisée en novembre et décembre 1961. Elle a porté sur 323 familles composées de 5 personnes : 2 adultes et 3 enfants bénéficiant des prestations familiales. Le chef de famille est salarié ⁽¹⁾. Ces familles résident dans cinq grandes villes : Paris et quatre villes de l'Ouest : Angers, Caen, Nantes et Rennes. Cet ensemble de ménages est donc assez homogène.

Ces familles avaient à enregistrer leurs dépenses au jour le jour pendant une, deux, quatre ou six semaines. Nous désignerons les quatre groupes de ménages auxquels ces périodes correspondent respectivement par A, B, C et D. Il n'y a pas eu d'interview sur les dépenses ; seules, les ressources ont été obtenues par cette méthode en fin d'enquête.

On avait remis un carnet à chaque personne du ménage susceptible d'effectuer des dépenses : toutes les mères de famille, 50 % des pères de famille et quelques grands enfants ont effectivement tenu un carnet de comptes. Ces dépenses ont été perforées sur cartes mécanographiques pour chaque personne et chaque jour d'enquête ; nous ne retenons ici que les **dépenses alimentaires hebdomadaires par ménage**.

Refus et abandons en cours d'enquête

Avant d'envisager d'allonger la durée des enquêtes, il fallait d'abord s'assurer que les taux de refus et d'abandons n'étaient pas trop importants ; c'est afin de comparer l'influence de la longueur des périodes d'enregistrement sur les taux de refus que nous avons choisi 4 périodes différentes. Pour obtenir 323 dossiers exploitables il a fallu contacter 490 ménages parmi lesquels 74 ne satisfaisaient pas aux conditions de l'enquête ⁽²⁾ et 17 ont été impossibles à joindre. Il reste donc 399 ménages qui se répartissent comme suit :

(1) Dans quelques cas les ressources du ménage comprennent des bénéfices commerciaux touchés par la mère de famille.

(2) Les ménages ont été tirés au sort dans les fichiers des Caisses d'Allocations Familiales des 5 villes citées. Ces fichiers tenus à jour tous les mois permettent de grouper les familles selon leur nombre d'enfants bénéficiant des allocations familiales, mais ne donnent pas de renseignement sur les autres personnes vivant au foyer. Or, on avait exclu les familles comprenant d'autres personnes que les 2 adultes ou les 3 enfants (par exemple un enfant aîné ne bénéficiant plus des prestations familiales ou une grand-mère). Ces familles n'ont pu être repérées qu'au cours de la première visite.

TABLEAU I

Refus et abandons en cours d'enquête

	ENSEMBLE	A		B		C		D	
		1 semaine		2 semaines		4 semaines		6 semaines	
		%		%		%		%	
Dossiers complets.....	286	67	83	61	79	75	68	83	64
Refus.....	76	14	17	12	16	21	19	29	22
Abandons après :									
1 semaine.....				4	5	7	6	4	3
2 semaines.....						5	5	6	5
3 semaines.....						3	2	5	4
4 semaines.....								3	2
5 semaines.....									
Total des abandons....	37			4	5	15	13	18	14
	399	81		77		111		130	

Le taux de refus croît assez peu avec la durée des relevés de comptes (1), de 17 à 22%. Il est très différent selon qu'il s'agit de familles résidant dans les villes de province ou à Paris. En province il se tient aux environs de 16% quel que soit le nombre de semaines requises, ce qui est assez surprenant. En revanche, à Paris, il croît régulièrement de 11% pour 1 semaine à 31% pour 6 semaines. Le nombre des abandons est plus faible que l'on ne prévoyait et surtout peu différent entre les échantillons C et D. Il faut cependant remarquer que les dossiers exploitables représentent 68% pour C et 64% pour D des ménages finalement contactés et répondant aux conditions de l'enquête. Cette distorsion incline à penser que les enquêtes exigeantes entraînent une **sélection** des ménages qui présente des inconvénients évidents si l'on s'intéresse à la représentativité des résultats. C'est sans doute le cas ici. Nous connaissons peu de chose sur les ménages ayant refusé leur collaboration à l'enquête, mais nous avons un certain nombre de renseignements intéressants sur ceux qui ont abandonné en cours de route. En particulier, lors de la première visite, des questions étaient posées sur la gestion du budget, la tenue des comptes, la prévision des dépenses et de l'épargne.

Sur les 18 ménages de l'échantillon D qui ont abandonné, 1 seul avait l'habitude de tenir ses comptes régulièrement (contre 23/83 parmi les autres) et 3 seulement ont déclaré mettre de côté à l'avance une partie de leurs ressources, en dehors d'une épargne contractuelle (contre 31/83 parmi les autres). Il semble donc que les ménages qui persistent aient, dans l'ensemble, un comportement plus réfléchi.

(1) Le nombre de semaines de comptes était fixé lors du tirage au sort ; les ménages en étaient informés dès la première visite.

Le seul renseignement que nous ayons sur les ménages ayant refusé est la profession. Les refus sont plus nombreux parmi les cadres supérieurs et moyens, ce qui entraîne comme corollaire que le pourcentage d'ouvriers et d'employés est plus élevé dans le groupe D. Les proportions d'ouvriers et d'employés sont de A à D, respectivement : 64 %, 66 %, 67 % et 79 %.

Revenus et dépenses des ménages enquêtés

L'analyse qui suit ne concerne que les ménages ayant tenu leurs comptes pendant toute la période fixée et ayant en outre déclaré la totalité de leurs ressources (1). Il reste en définitive 280 ménages sur lesquels portera notre analyse.

En opposition à ce qu'on observe d'ordinaire, les **revenus** semblent avoir été déclarés avec précision — la question n'a d'ailleurs pas offert de difficulté lors de l'enquête et la comparaison des ressources et des dépenses présentées ci-après est satisfaisante :

	Ressources mensuelles y compris prestations familiales	Dépenses totales mensuelles
(en francs par ménage)		
A	1 258	1 264
B	1 168	1 230
C	1 212	1 165
D	1 063	1 066 (2)

On peut évidemment objecter que la méthode d'enquête laissant une grande liberté aux familles (3) peut entraîner une sous-estimation des dépenses. Mais par ailleurs, compte tenu de la forte proportion d'ouvriers dans notre échantillon, les ressources déclarées, sont vraisemblables ; elles se rapportent au mois de novembre et ne comprennent donc pas les primes de fin d'année.

Les ressources moyennes présentées ci-après serviront à interpréter les différences des dépenses alimentaires spécialement entre les Parisiens et les provinciaux d'une part et entre les 4 groupes A, B, C, et D, d'autre part, et à se faire une idée de la dispersion des niveaux de vie des familles enquêtées. D'après le tableau II le niveau de vie des Parisiens dans cette enquête est de près de 50 % supérieur à celui des provinciaux. Malgré cela la plupart des calculs qui suivent porteront sur l'ensemble des 2 populations. Les différences entre échantillons sont assez importantes ; le taux de refus étant ordinairement plus grand dans les classes de revenu aisées on pouvait s'attendre d'une part à ce que les revenus décroissent de A à D et d'autre part à ce que le revenu des ménages ayant abandonné en cours de route soit plus fort que celui des autres ménages. En ce qui concerne le premier point, les revenus sont plus faibles pour le groupe D, mais peu différents pour les autres groupes (cependant la moyenne du groupe C

(1) Au total 8 ménages (dont 2 qui ont par la suite abandonné) n'ont pas déclaré leurs ressources ou n'en ont déclaré qu'une partie.

(2) 3 dépenses ont été éliminées pour le calcul de cette moyenne ; il s'agit d'achats immobiliers s'élevant respectivement à 1 000, 4 000 et 8 000 francs.

(3) Toutefois les carnets de comptes, hebdomadaires, étaient relevés toutes les semaines.

devient 1 160 si l'on enlève 2 ménages ayant des ressources supérieures à 40 000 F par an). Les revenus des ménages ayant abandonné ne sont pas plus élevés que ceux des autres.

TABLEAU II
Revenus mensuels des ménages enquêtés (1)

Différence entre Paris et la province :			
Paris	Province	Ensemble	
1 500 (45%)	1 070 (48%)	1 170 (46%)	
Différence entre les groupes de ménages :			
A	B	C	D
1 260 (47%)	1 170 (46%)	1 210 (50%)	1 060 (33%)
Ménages ayant abandonné :			
	B	C	D
	1 150	1 050	1 030
<p>(1) En francs par ménage. Ces revenus comprennent les prestations familiales, les heures supplémentaires, mais ne comprennent ni les primes de fin d'année, ni les avantages en nature. Entre parenthèses figure le coefficient de variation, rapport de l'écart-type à la moyenne.</p>			

2. PRÉCISION DE LA DÉPENSE ALIMENTAIRE SELON LA DURÉE DES COMPTES

Deux sortes de facteurs contribuent à la dispersion d'une dépense observée :

1. — Les facteurs de différences **systématiques** entre les ménages, facteurs classiques tels les revenus, le groupe social, la région, et facteurs dont l'influence est moins aisée à repérer, habitudes, mode de vie, goûts ou milieu environnant.

2. — Les facteurs entraînant des différences entre la dépense observée pour un ménage et sa dépense habituelle en période normale :

les **aléas** : présence d'invités ou absence d'une personne de la famille, proximité de la date de la paie, achats extraordinaires entraînant une modification de la structure du budget...;

les **erreurs de mesure**, erreurs de mémoire ou de transcription ;

les imprécisions dues à l'**indivisibilité** des achats; pour les achats peu fréquents en particulier, les dépenses observées peuvent différer beaucoup de la dépense réelle à long terme du ménage.

C'est, bien entendu, la part systématique de la variance qui nous intéresse pour une explication du comportement. En revanche la seconde partie nuit à l'analyse et doit être réduite.

Évolution de la dépense alimentaire

Le tableau III présente les moyennes et écart-type de l'ensemble de la dépense alimentaire pour chacune des semaines d'enquête prise séparément ; le tableau IV nous donne les mêmes mesures pour un nombre croissant de semaines. Ces deux tableaux concernent, pour les quatre groupes A, B, C et D, les villes de province seules et pour le groupe D l'ensemble de l'échantillon.

Le graphique I présente la distribution de la dépense alimentaire pour chacun des quatre groupes.

GRAPHIQUE I
Distribution de la dépense alimentaire

% de ménages ayant une dépense alimentaire inférieure à x
(échelle de Gauss)

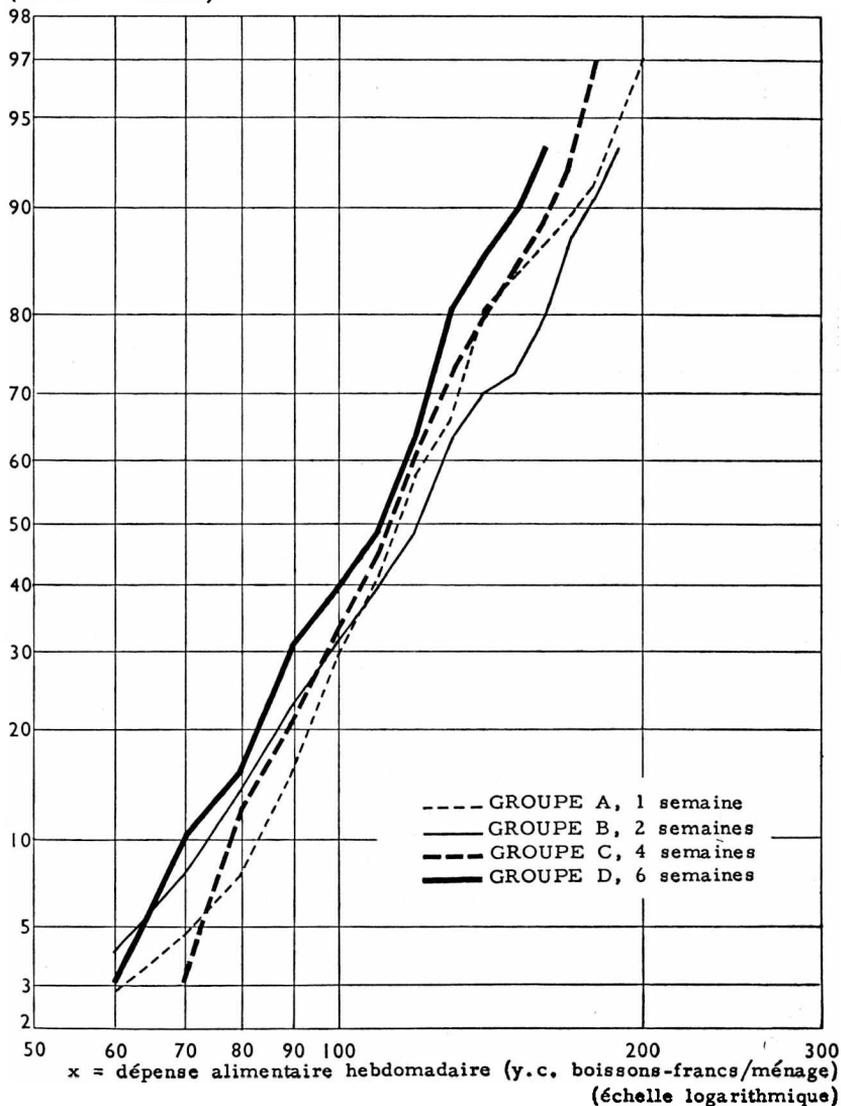


TABLEAU III

Dépense alimentaire pour chaque semaine séparément
(non compris les boissons, les dépenses de cantine et de restaurant)

(en francs par ménage)

Villes de province :		1 ^{re} semaine	2 ^e semaine	3 ^e semaine	4 ^e semaine	5 ^e semaine	6 ^e semaine	ENSEMBLE (1)
A. (51 ménages)	Moyenne	94						94
	Ecart-type	26						26
	C.V.	28%						28%
B. (42 ménages)	Moyenne	94	85					89
	Ecart-type	29	29					25
	C.V.	31%	34%					28%
C. (57 ménages)	Moyenne	91	88	88	86			88
	Ecart-type	30	24	23	23			21
	C.V.	33%	27%	26%	27%			23%
D. (64 ménages)	Moyenne	84	85	86	82	81	86	84
	Ecart-type	28	29	28	27	22	27	22
	C.V.	33%	34%	32%	33%	27%	31%	26%
Paris et villes de province :								
D. (80 ménages)	Moyenne	94	91	93	88	87	92	91
	Ecart-type	35	30	30	29	23	29	25
	C.V.	37%	33%	32%	32%	27%	31%	27%

C.V. : coefficient de variation : rapport de l'écart-type à la moyenne.

(1) Les moyennes calculées dans cette colonne sont hebdomadaires, pour permettre la comparaison avec les semaines séparées.

La dépense moyenne est un peu moins élevée en ce qui concerne les familles du groupe D qui ont un niveau de vie plus bas. Pour une raison analogue la moyenne du même groupe, Paris compris, est augmentée. L'annexe E donne la liste des dépenses des 80 ménages de ce groupe ; la différence entre Paris et la province y est évidente. Nous ne comparerons pas davantage les moyennes entre groupes, des causes d'hétérogénéité subsistent et surtout les effectifs sont très faibles, rendant les comparaisons fragiles.

Les moyennes concernant la première semaine sont plus élevées sauf pour les provinciaux du groupe D. En se reportant aux annexes B et C on constate que les dépenses **par produit** sont plus fortes la première semaine (pour 22 produits sur 44 en C et 31 en B). Il nous est difficile de conclure à une surestimation des dépenses la première semaine, ou à un effet de lassitude pour les autres ; pour des raisons matérielles l'enquête n'a pas commencé à la même date pour tous les ménages. Elle a commencé :

	Entre le 23 et le 29 octobre	Entre le 30 octobre et le 5 novembre	
pour	31 %	34 %	des familles A
	42 %	42 %	des familles B
	61 %	31 %	des familles C
	72 %	26 %	des familles D

Donc, l'enquête débutait plus tôt pour les familles pour lesquelles elle durait plus longtemps (1). Pour une forte proportion des D la première semaine de comptes aura coïncidé avec la dernière semaine du mois d'octobre, avant le paiement des salaires, ce qui peut expliquer une dépense plus faible. Tandis que pour les autres et surtout pour les C la première semaine inclut les premiers jours de novembre ce qui peut en revanche avoir entraîné un léger accroissement des dépenses (ou paiement de commandes).

Nous ne sommes donc pas dans les conditions requises pour juger d'une différence des moyennes observées selon la longueur de la période de relevé des comptes imposée aux ménages enquêtés. Toutefois, s'il existe une surestimation des dépenses enregistrées au cours de la première semaine lorsqu'il n'y a que deux semaines en tout (groupe B), cela ne semble pas se produire lorsque la période prescrite est plus longue.

Ce sont les dispersions qui nous intéressent plus spécialement. L'écart-type calculé pour l'ensemble des semaines est évidemment inférieur à celui qui se rapporte à chaque semaine prise séparément (2), quel que soit le

(1) Il y a eu encore quelques nouveaux contacts jusqu'au 19 novembre, les derniers contactés étant des ménages des groupes A et B.

(2) Il s'agit en effet de l'écart-type d'une somme de variables non indépendantes, qu'on peut écrire sous la forme :

$$\sigma^2 = (\sum \sigma_j)^2 - 2 \sum_{j \neq j'} (1 - \rho_{jj'}) \sigma_j \sigma_{j'}$$

où σ^2 désigne la variance de la somme des dépenses au cours des k semaines $\sigma^2 = \frac{k^2}{n-1} \sum_i (x_i - \bar{x})^2$,

σ_j^2 désigne la variance de la dépense de la semaine j $\sigma_j^2 = \frac{1}{n-1} \sum_i (x_{ij} - \bar{x}_j)^2$;

$\rho_{jj'}$ désigne le coefficient de corrélation entre les dépenses des semaines j et j'.

L'indice i se rapporte au ménage et varie de 1 à n, l'indice j à la semaine et varie de 1 à k

$$\bar{x} = \frac{1}{nk} \sum_{ij} x_{ij} = \frac{1}{n} \sum_i \bar{x}_i = \frac{1}{k} \sum_j \bar{x}_j.$$

groupe considéré. S'il en était autrement, la corrélation entre les semaines serait parfaite ; ce qui implique, puisqu'il s'agit de dépenses, que celles-ci soient égales quelle que soit la semaine considérée pour un même ménage. On ne se trouve évidemment pas dans ce cas où notre problème de durée d'observation ne se poserait plus puisqu'une seule semaine suffirait pour obtenir la meilleure estimation possible ; la variance ainsi calculée représenterait exactement la **variance entre ménage**. Plus la corrélation entre semaine est faible, plus l'écart-type calculé sur toute la période est petit par rapport à chacun des écarts-types hebdomadaires. Si cette corrélation était nulle il serait réduit dans la proportion de $1/\sqrt{k}$ (où k est le nombre de semaines comprises dans la période). Les valeurs négatives des corrélations, en particulier, entraîneraient un accroissement important du gain de précision des estimations, mais il n'est pas possible de trouver des corrélations négatives au niveau de la dépense totale alimentaire.

TABLEAU IV

Dépense alimentaire hebdomadaire, semaines cumulées
(non compris les boissons, les dépenses de cantine et de restaurant)
(en francs par ménage)

	1 semaine	2 semaines	3 semaines	4 semaines	5 semaines	6 semaines
Villes de province:						
A. Moyenne	94					
Ecart-type	26					
C.V.	28%					
B. Moyenne	94	89				
Ecart-type	29	25				
C.V.	31%	28%				
C. Moyenne	91	90	89	88		
Ecart-type	30	23	21	21		
C.V.	33%	26%	24%	23%		
D. Moyenne	84	84	85	84	84	84
Ecart-type	28	25	24	24	23	22
C.V.	33%	30%	28%	28%	27%	26%
Paris et villes de province:						
D. Moyenne	94	93	93	92	91	91
Ecart-type	35	30	28	27	26	25
C.V.	37%	32%	30%	29%	28%	27%
C.V. : coefficient de variation, rapport de l'écart-type à la moyenne.						

Si l'on prend comme mesure de l'accroissement de la précision des estimations l'expression :

$$g = \frac{\sum \sigma_j}{\sigma}$$

On obtient pour chaque échantillon les valeurs suivantes de g :

	2 semaines	3 semaines	4 semaines	5 semaines	6 semaines
Groupe B.	1,15	—	—	—	—
Groupe C.	1,16	1,20	1,20	—	—
Groupe D.	1,10	1,16	1,18	1,18	1,20

Cet accroissement de précision entraîne le même accroissement des coefficients de corrélation avec n'importe quelle variable. Supposons en effet que l'on calcule le coefficient de corrélation entre dépense alimentaire X et revenu R , soit $\rho(X, R)$, nous obtenons :

$$\rho(X, R) = \frac{\sigma^2_{X, R}}{\sigma_X \sigma_R} = \sum_j \rho(x_j, R) \frac{\sigma_j}{\sigma}$$

x_j désignant la dépense pour la semaine j , $X = \sum_j x_j$.

Si les corrélations sont approximativement égales, quelle que soit la semaine, leur valeur commune étant $\rho(x, R)$,

$$\rho(X, R) = \rho(x, R) \frac{\sum \sigma_j}{\sigma}$$

le coefficient obtenu à partir d'une estimation sur l'ensemble des semaines sera toujours supérieur au coefficient pour chacune des semaines prise séparément. Le tableau V présente les estimations obtenues pour les 4 échantillons et pour des durées variables.

TABLEAU V

Comparaison des corrélations ρ^2 entre dépense alimentaire et revenu calculées sur un nombre croissant de semaines

(Paris et villes de province)

	1 semaine	2 semaines	3 semaines	4 semaines	5 semaines	6 semaines
Groupe A.	0,15					
Groupe B.	0,33	0,36				
Groupe C.	0,41	0,42	0,44	0,40		
Groupe D.	0,22	0,25	0,30	0,30	0,30	0,32

Les corrélations calculées pour chaque semaine prise séparément sont inégales pour le groupe C, en particulier on obtient $\rho^2 = 0,13$ pour la quatrième semaine, ce qui explique la valeur plus faible 0,40 donnée plus haut pour les 4 semaines. En revanche les corrélations sont très voisines quelle que soit la semaine pour le groupe D ; elles varient entre $\rho^2 = 0,21$ et $\rho^2 = 0,24$. Il est donc normal d'observer un accroissement régulier de ρ^2 au fur et à mesure que les semaines s'ajoutent. Cet accroissement est de l'ordre de grandeur prévu, 1,40.

Variance intra-ménage et variance entre ménages

L'intérêt d'un allongement de la durée des relevés de comptes et l'amélioration de la précision qui en découle tiennent au fait que les dispersions observées pour un même ménage au cours des semaines ne sont pas négligeables, fréquence et montant des dépenses n'étant pas réguliers pour ce ménage. Considérant la semaine comme unité de période, on peut calculer une **variance intra-ménage**, mesurant la dispersion des dépenses hebdomadaires autour de leur moyenne pour toute la période, l'expression retenue étant :

$$V' = \frac{\sum_{ij} (x_{ij} - \bar{x}_i)^2}{n(k-1)}$$

La **variance entre ménages** est :

$$V = \frac{\sigma^2}{k^2} = \frac{\sum_i (\bar{x}_i - \bar{x})^2}{n-1}$$

Si les nk observations formées par les dépenses des n ménages au cours des k semaines pouvaient être considérées comme des estimations indépendantes d'une même variable, les expressions V' et kV seraient deux estimations indépendantes de la variance totale. On n'est évidemment pas dans ce cas, mais la décomposition classique nous a cependant semblé intéressante à présenter (tableau VI ci-après). Le calcul a été fait en rapportant chaque observation \bar{x}_i d'une part à la moyenne générale, et d'autre part à la droite de régression $x = aR + b$ où R désigne le revenu.

TABLEAU VI
Dépense alimentaire hebdomadaire (1)
Groupe D (80 ménages)

Source de variation	Somme des carrés	Degrés de liberté	Variances	(2)
Variance totale $\sum(x_{ij} - \bar{x})^2$	420 093	479	877	877
Variance intra-ménage : $\sum(x_{ij} - \bar{x}_i)^2$	120 958	400	302	302
Variance entre ménages : $\sum k(\bar{x}_i - \bar{x})^2$	299 135	79	3 786	631
Dans l'hypothèse d'une loi de la forme $x = aR + b$, décomposition de la variance entre ménages :				
Variance résiduelle : $k \sum_i [\bar{x}_i - \bar{x} - a(R_i - \bar{R})]^2$	203 308	78	2 606	434
Variance expliquée : $ka^2 \sum_i (R_i - \bar{R})^2$	95 827	1	95 827	

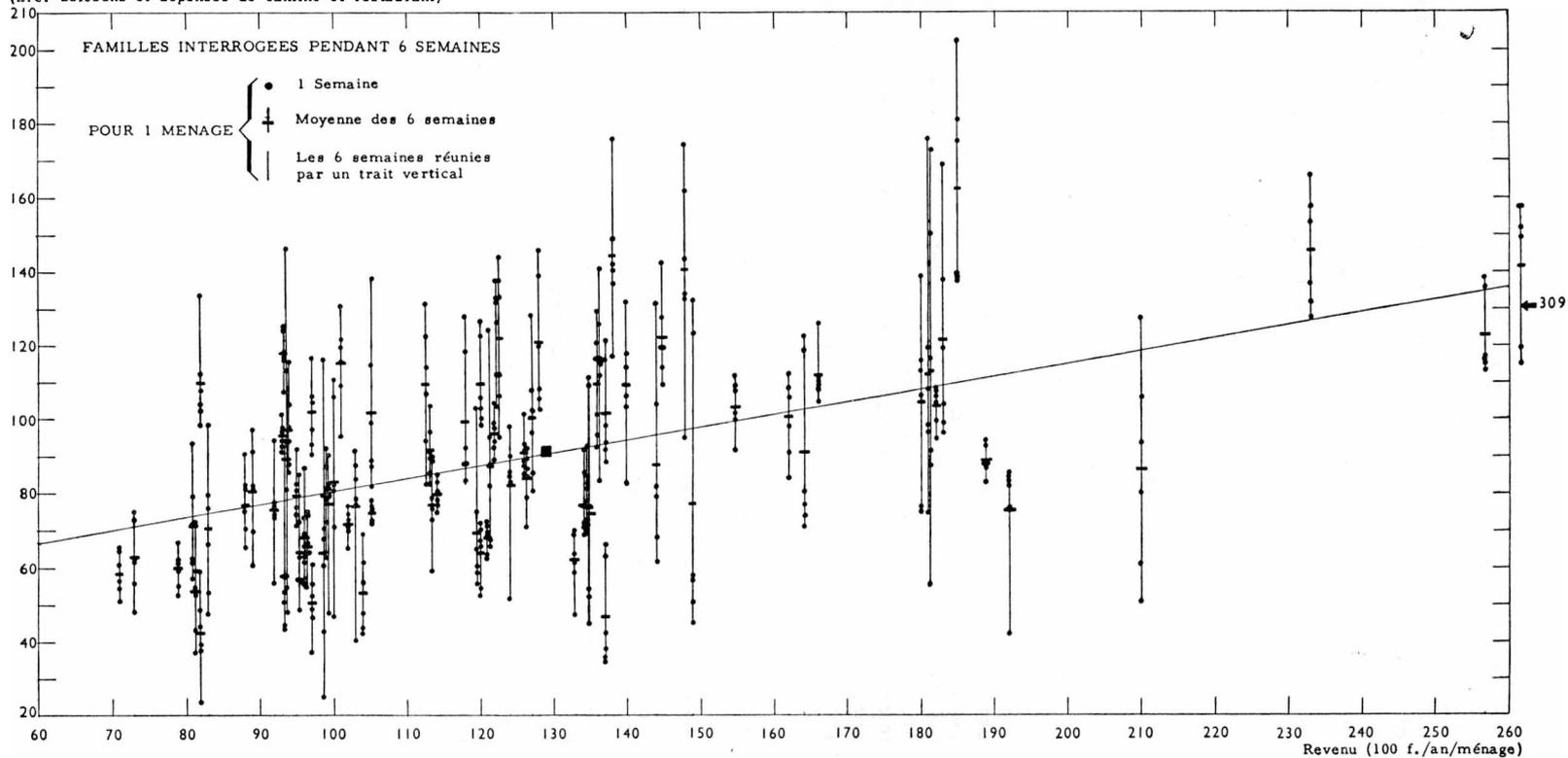
(1) La moyenne de cette dépense pour le groupe D est de 91 F par semaine et par ménage.
(2) Dans cette colonne figurent les variances divisées par k .

GRAPHIQUE II

Dépense alimentaire en fonction du revenu Groupe D.

Dépense alimentaire frs/semaine/ménage

(n.c. boissons et dépenses de cantine et restaurant)



En fait, la comparaison qui nous intéresse n'est pas celle de kV à V' , mais celle de V , la variance retenue pour l'explication, à V' mesurant les erreurs d'observation (1). La dernière colonne du tableau VI présente ces grandeurs. On notera l'importance de la variance intra-ménage, 302, qui représente seulement la moitié de la variance à expliquer, 631, et les 2/3 de cette variance une fois éliminée l'influence du revenu, 434. Le graphique II qui présente les dépenses alimentaires en fonction du revenu illustre bien les dispersions individuelles : on a relié par un trait vertical les 6 dépenses hebdomadaires d'un ménage. Il reste cependant bien des sources de variation : âge des enfants, villes de résidence, nombre de repas pris au foyer...

3. LA PRÉCISION DES ESTIMATIONS ET LA FRÉQUENCE D'ACHAT SELON LES PRODUITS

La dépense alimentaire considérée au paragraphe précédent est la somme d'un grand nombre de dépenses d'importance variable et surtout de fréquence variable. La corrélation entre les dépenses alimentaires totales se rapportant à différentes semaines ne peut être négative ; en revanche celle qui concerne des produits particuliers peut l'être. L'analyse de la précision des estimations prend donc un tout autre intérêt dans ces derniers cas. L'allongement de la durée des comptes permet d'une part d'atteindre un plus grand nombre de **consommateurs rares** et d'autre part d'obtenir, pour ces consommateurs effectifs une meilleure estimation de leur dépense habituelle. Avant d'analyser l'amélioration de la précision selon les produits, nous étudierons l'apparition des **nouveaux consommateurs** au cours des semaines et la **fréquence des achats**. Cette étude ne portera que sur le groupe des 80 ménages observés pendant 6 semaines (groupe D).

Apparition des consommateurs au cours des semaines et fréquence d'achat selon les produits

L'apparition des nouveaux consommateurs au cours des semaines dépend uniquement de leur rythme d'achat. Nous sommes habitués à analyser des résultats d'enquête portant sur période courte, il est donc intéressant d'examiner quelle proportion de consommateurs on saisit avec une ou deux semaines. Les graphiques III présentent ces résultats.

1) Remarquons que l'expression $g = \frac{\sum_j \sigma_j}{\sigma}$, est inversement proportionnelle au rapport de Fisher-

Snédécór $F = \frac{kV}{V'}$;

On déduit en effet de l'équation de l'analyse de la variance :

$$k \frac{\sum (x_{ij} - \bar{x}_j)^2}{k^2 \sum (\bar{x}_i - \bar{x})^2} = \frac{1}{F} \frac{n(k-1)}{n-1} + 1 - \frac{n \sum (\bar{x}_i - \bar{x})^2}{k \sum (\bar{x}_i - \bar{x})^2}$$

l'expression :

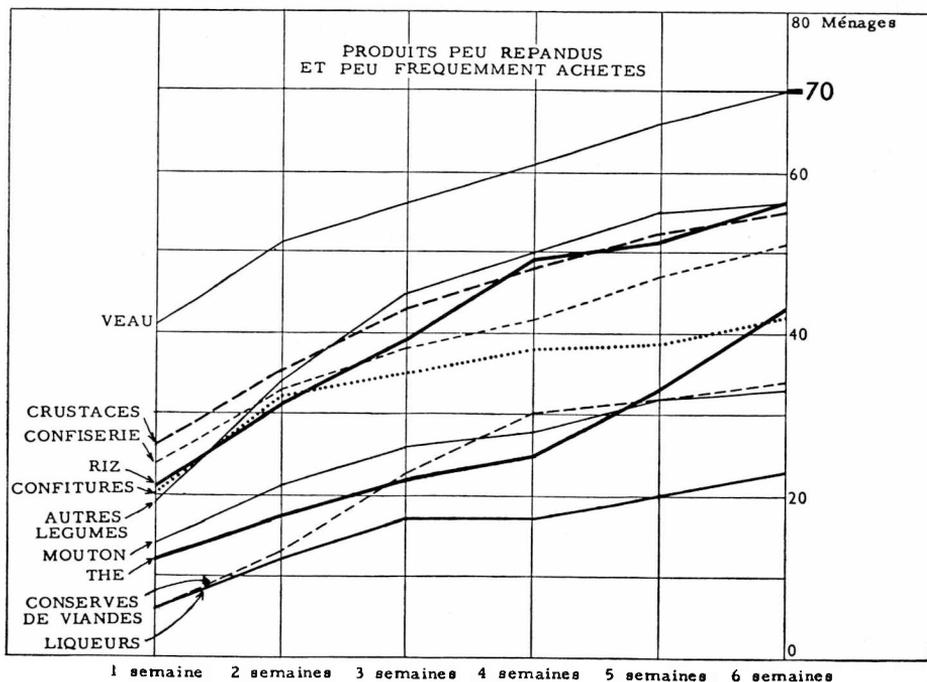
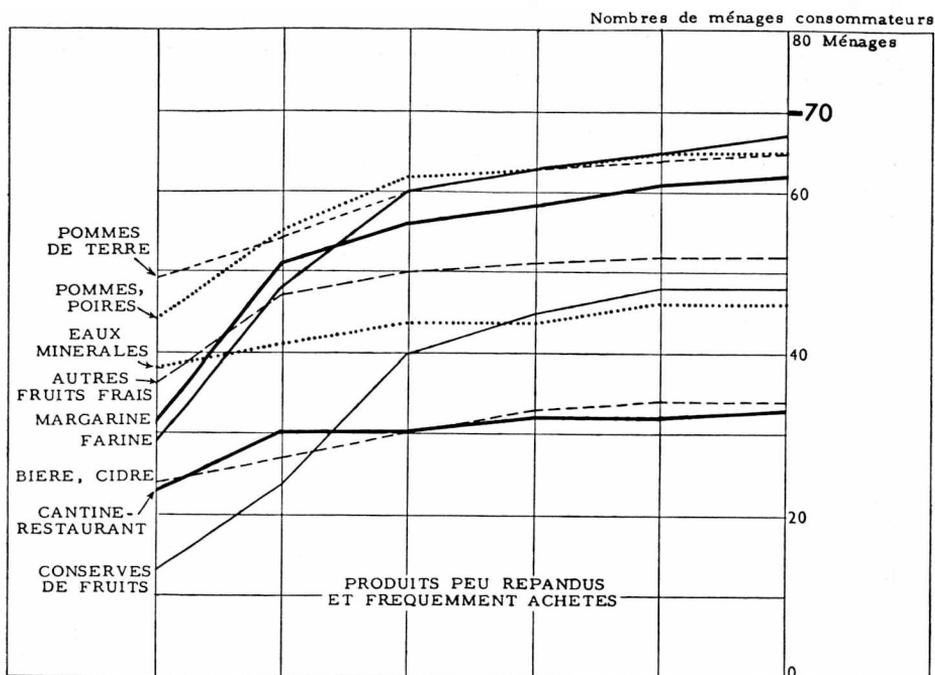
$$\frac{\sum (x_{ij} - \bar{x}_j)^2}{k^2 \sum (\bar{x}_i - \bar{x})^2} = \frac{\sum_j^2}{\sigma^2}$$

est approximativement égale à g^2/k .

Le dernier terme du second membre étant très petit, F est à peu près égal à $\frac{n}{n-1} \frac{k-1}{g^2-1}$.

GRAPHIQUE III (*)

Apparition des consommateurs au cours des semaines
Groupe D = 80 ménages



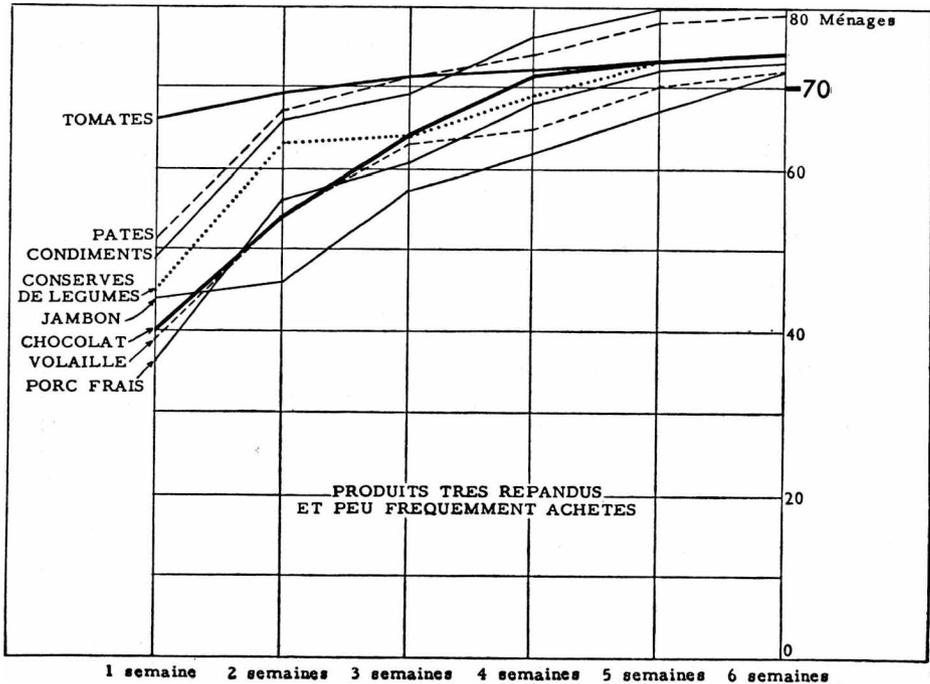
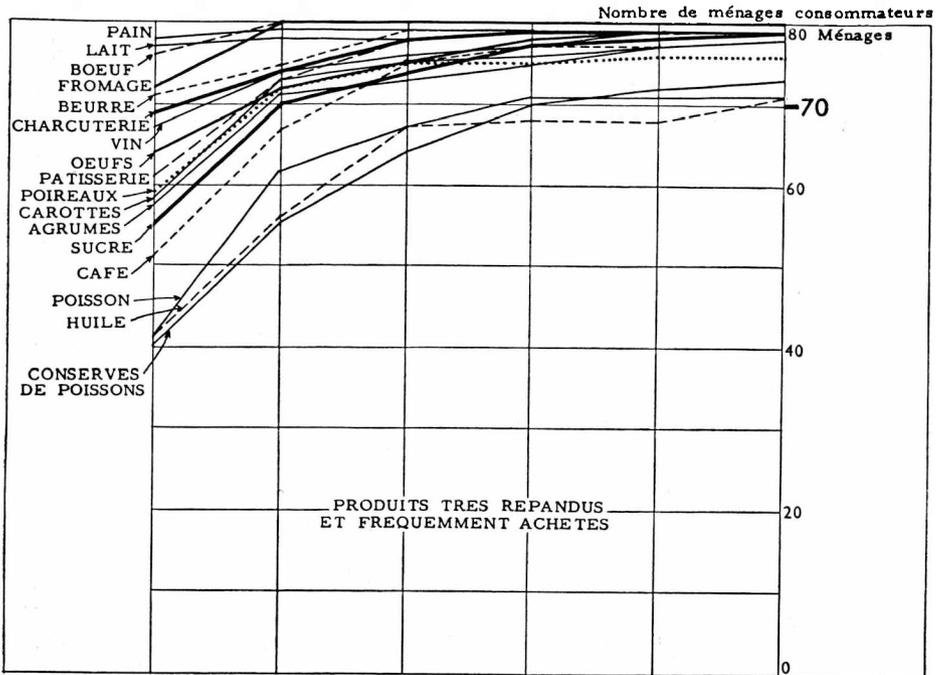
* Produits de grande consommation ou très répandus $n_6 > 70$.

Produits fréquemment achetés : $n_3 - n_1 = \frac{3}{4} (n_6 - n_1)$

n_i désignant le nombre de consommateurs appurés la $i^{\text{ème}}$ semaine.

GRAPHIQUE III (suite)

Apparition des consommateurs au cours des semaines



D'une semaine à l'autre le nombre de ménages croît évidemment, mais cet accroissement varie selon les produits et le nombre de nouveaux venus chaque semaine diminue évidemment (1). Pour certains produits, même très courants, la première semaine seule ne donne qu'une proportion médiocre des consommateurs : la moitié en ce qui concerne le porc et la volaille ; pour ces produits des consommateurs apparaissent encore la sixième semaine et pour beaucoup de produits, il demeure encore au bout de 6 semaines des ménages non-acheteurs.

En ce qui concerne les denrées périssables on peut considérer que ceux qui n'ont déclaré aucune dépense sur 6 semaines ne sont pas consommateurs ou sont des consommateurs très rares ; en ce qui concerne les produits stockables on peut aussi bien penser que les ménages en question font leur commande à un rythme supérieur à 6 semaines. Deux phénomènes interviennent par conséquent. D'une part la fréquence des achats, la pente des courbes relatives aux produits souvent achetés diminue rapidement (pour des produits tels que le pain, le bœuf, le lait, plus de 95% des consommateurs apparaissent dès la première semaine). D'autre part le caractère plus ou moins répandu de certaines denrées : thé, bière, apéritifs ou liqueurs, eau minérale. On a essayé de grouper les produits en tenant compte de ces deux caractéristiques.

Nous ne pouvons présenter ici la distribution des fréquences d'achats par les 44 produits; nous nous sommes donc contentés d'en donner une description par les quartiles présentés au tableau VII ci-après. Les données ayant été perforées pour chacun des jours d'enquête, deux achats d'une même denrée faits le même jour par la même personne du ménage comptent pour un seul. Pour chaque ménage le nombre d'achats par produit est généralement inférieur à 42 puisqu'il s'agit de 6 semaines de comptes (2). Les produits ont été rangés dans l'ordre décroissant du premier quartile, donc approximativement par ordre de fréquence décroissante. En fait les distributions ne sont pas toujours régulières, ce même tableau montre par exemple que pour l'eau minérale, la bière, le vin et les pommes de terre, il existe deux populations, celle des acheteurs au jour le jour et celle des ménages qui stockent. On repère aisément les produits à fréquence d'achat rare mais très répandus : porc, volaille, poisson, chocolat, café et les produits fréquemment achetés mais par un petit nombre de consommateurs seulement : eau minérale, bière.

Amélioration de la précision en fonction de la fréquence d'achat selon les produits

Pour mettre en évidence le rôle de la fréquence des achats dans l'amélioration de la précision (mesuré par l'expression g proposée plus haut), on peut répartir les ménages selon le nombre de semaines au cours desquelles ils ont effectué au moins un achat pour un produit donné. Si l'on exclut les ménages n'ayant effectué aucun achat, nous aurons ainsi 6 ensembles de ménages pour l'échantillon D . Désignons par p_k la proportion des ménages ayant acheté pendant k semaines, x_k la somme de

(1) Si pour simplifier on suppose que l'unité des rythmes d'achats est la semaine et si l'on désigne par α_{n-1} le nombre de consommateurs faisant un achat toutes les $n-1$ semaines, le nombre de nouveaux venus la $n^{\text{ième}}$ semaine est égal au nombre de nouveaux venus la $n-1^{\text{ième}}$ semaine diminué de $\frac{1}{n-1} \alpha_{n-1}$.

(2) Il dépasse 42 pour des denrées achetées simultanément par la mère de famille pour les besoins du ménage et par le père pour sa consommation personnelle. Ceci arrive très rarement et uniquement pour le vin et la bière.

TABLEAU VII

Fréquence des achats des ménages du groupe D (6 semaines)

PRODUITS	Nombre de non consommateurs	PARMI LES CONSOMMATEURS		
		25% ont fait un nombre d'achats inférieur ou égal à	50% ont fait un nombre d'achats inférieur ou égal à	75% ont fait un nombre d'achats inférieur ou égal à
Pain.....	1	30,3	34,2	36,9
Lait.....	1	29,3	36,3	39,8
Fromage.....	0	10,3	14,4	19,6
Bœuf.....	0	8,6	11,6	15,7
Charcuterie.....	1	7,8	11,6	15,6
Tomates.....	6	6,8	10,4	15,3
Beurre.....	1	6,2	10,1	12,8
Agrumes.....	2	5,2	9,7	13,6
Carottes.....	3	5,0	8,3	11,5
Poireaux.....	4	5,0	6,8	10,8
Vin.....	1	4,8	11,2	25,5
Pâtisserie.....	1	4,3	8,2	10,2
Œufs.....	1	3,8	5,5	8,8
Pâtes.....	1	3,1	4,3	7,1
Sucre.....	1	3,0	5,3	7,0
Condiments.....	0	2,8	5,2	7,4
Pommes, poires.....	15	2,6	6,3	10,8
Poisson.....	9	2,6	4,2	5,8
Eau minérale.....	34	2,5	8,0	27,5
Café.....	2	2,2	3,9	6,1
Bière, cidre.....	46	2,2	9,0	32,3
Jambon.....	8	2,2	4,2	7,7
Cantine, restaurant.....	47	2,1	4,8	8,4
Margarine.....	18	1,9	4,3	6,8
Conserves de légumes.....	6	1,8	3,2	5,1
Pommes de terre.....	15	1,7	6,6	11,9
Conserves de poisson.....	7	1,6	3,6	6,1
Volaille.....	8	1,6	2,9	4,2
Veau.....	10	1,4	3,5	5,8
Farine.....	12	1,3	2,5	3,9
Huile.....	10	1,2	1,9	3,2
Crustacés.....	25	1,1	2,4	5,1
Porc frais.....	7	1,0	2,6	4,4
Chocolat.....	6	1,0	2,2	4,9
Confiserie.....	29	0,9	2,0	3,8
Confiture.....	38	0,8	1,8	3,8
Conserves de fruits.....	32	0,8	1,7	3,3
Fruits frais métropolitains.....	28	0,7	1,5	3,0
Mouton.....	47	0,6	1,3	2,3
Riz.....	24	0,6	1,4	3,0
Conserves de viandes.....	46	0,5	1,0	2,3
Autres légumes.....	24	0,5	1,0	2,0
Liqueurs.....	57	0,4	0,8	1,7
Thé.....	37	0,3	0,7	1,1

leurs dépenses pour toute la période, et σ_k^2 la variance de cette somme, la variance calculée pour l'ensemble de l'échantillon, σ , figurant au dénominateur de g , s'exprime en fonction des 6 groupes définis ci-dessus par :

$$\sigma^2 = \sum_k p_k \sigma_k^2 + 2 \sum_{h \neq h'} p_h p_{h'} (x_h - x_{h'})^2$$

comme la moyenne pondérée des variances correspondantes σ_{jk}^2 ou σ_k^2 des groupes de ménages. La variance de la somme des dépenses pour le groupe k , σ_k^2 sera d'autant plus faible que les corrélations entre semaines pour ce groupe seront elles-mêmes plus faibles ; le bon sens nous fait supposer que les consommateurs achetant un produit toutes les semaines achètent pour un montant variant peu, la corrélation pour ce groupe sera positive. Si l'on considère, en revanche, le groupe des consommateurs rares où la fréquence des achats nuls est forte, la corrélation est négative. **Si la proportion relative des consommateurs rares est élevée les corrélations négatives l'emporteront sur les autres.** Le coefficient g peut ainsi atteindre et même dépasser \sqrt{k} . On retiendra pour mémoire cette valeur en observant toutefois que l'indépendance qu'elle implique entre les semaines est un peu artificielle ; elle tient en fait à l'indivisibilité des achats.

Quant à l'expression $\sum_{h \neq h'} p_h p_{h'} (x_h - x_{h'})^2$, elle est à peu près nulle pour

les produits dont le montant acheté varie en fonction inverse du nombre d'achats ; en même temps les corrélations négatives se rapportant à des montants élevés sont fortes en valeur absolue. On peut donc s'attendre à ce que le gain de précision soit plus important pour ces produits que les consommateurs stockent.

Les graphiques IV et V présentent les résultats obtenus pour les échantillons B, C, et D, Paris et les villes de province ensemble. En abscisse figure le nombre d'achats moyen par quinzaine, c'est-à-dire le rapport du nombre total d'achats pour le groupe considéré, divisé par le nombre des ménages ⁽¹⁾ ; en ordonnée figure le coefficient g . D'après ce qui précède, la fréquence des achats est en réalité proportionnelle à $1/g^2$. Nous n'avons fait aucune transformation de façon à ce qu'on puisse repérer sur les graphiques les produits pour lesquels le poids des corrélations négatives est le plus fort, produits se rapprochant de l'ordonnée \sqrt{k} ⁽²⁾.

Le graphique IV se rapporte à l'ensemble des ménages interrogés ; le graphique V aux seuls consommateurs, c'est-à-dire aux ménages ayant effectué au moins un achat au cours de la période d'enquête. Nous donnons les deux mesures car la seconde accentue la représentation du phénomène que nous voulons mettre en évidence ; nous savons en effet que, même pour le groupe D nous n'avons pas la totalité des consommateurs réels ⁽³⁾.

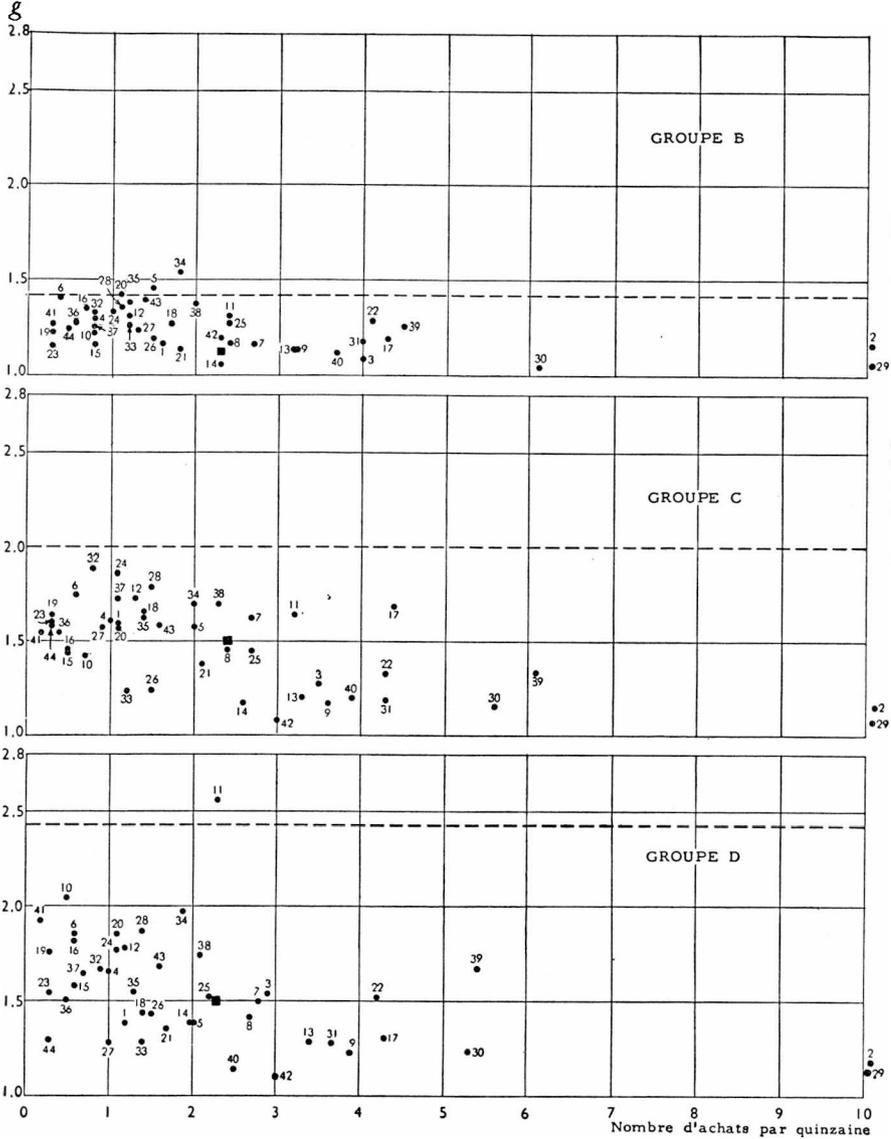
(1) Multiplié par 1/2 pour C et 1/3 pour D.

(2) D'après la note de la page 61 on aurait pu aussi bien prendre pour ordonnée F ou F/k . La signification de ce coefficient nous a semblé moins concrète pour une mesure. Le tableau VIII présente ces valeurs pour les différents produits pour les échantillons C et D.

(3) En ce qui concerne g , le calcul sur les seuls consommateurs donne des valeurs un peu supérieures et ceci surtout pour les produits peu répandus.

GRAPHIQUE IV

**Gain de précision et fréquence d'achats
(calculs portant sur l'ensemble des ménages)**



- | | | | |
|--|---|--|--|
| <ul style="list-style-type: none"> 1 - Restaurant-cantine 2 - Pain 3 - Pâtisserie 4 - Farine 5 - Pâtes 6 - Riz 7 - Carottes 8 - Poireaux 9 - Tomates 10 - Autres légumes 11 - Pommes de terre | <ul style="list-style-type: none"> 12 - Conserves de légumes 13 - Agrumes 14 - Pommes, poires 15 - Autres fruits frais 16 - Conserves de fruits 17 - Bœuf 18 - Veau 19 - Mouton 20 - Porc frais 21 - Jambon 22 - Charcuterie | <ul style="list-style-type: none"> 23 - Conserves viandes 24 - Volailles 25 - Œufs 26 - Poissons 27 - Crustacés 28 - Conserves poissons 29 - Lait 30 - Fromages 31 - Beurre 32 - Huile 33 - Margarine | <ul style="list-style-type: none"> 34 - Sucre 35 - Chocolat 36 - Confitures 37 - Confiserie 38 - Condiments 39 - Vin 40 - Bière, cidre 41 - Liqueurs 42 - Eaux minérales 43 - Café 44 - Thé |
|--|---|--|--|

Le nombre d'achats moyen est d'ailleurs plus faible pour les groupes B et C, dans lesquels les consommateurs rares ne sont pas tous présents. En fait ce n'est pas la discrimination entre consommateurs réels et non consommateurs qui peut jouer sur la précision mais, à l'intérieur du groupe des consommateurs, c'est la distinction entre consommateurs rares ou fréquents, l'abscisse du graphique IV ne donne pas une mesure de la fréquence d'achats dans ce sens. Enfin une analyse correcte du comportement doit distinguer entre les consommateurs réels et les autres ménages dans la mesure où les données le permettent.

La relation entre fréquence d'achats et amélioration de la précision est assez nette et son intensité varie selon les groupes B, C ou D. La fréquence des achats n'est pas le seul facteur en cause, mais également le **mode d'achat**, surtout en ce qui concerne les denrées stockables. Les relations observées sont d'ailleurs meilleures en ce qui concerne les denrées périssables, pour elles la dépense totale effectuée sur l'ensemble de la période est proportionnelle à la fréquence des achats. Pour certains produits stockables où les dépenses sont d'autant plus élevées qu'elles sont rares, les corrélations négatives sont fortes en valeur absolue et le gain de précision plus important que pour les autres dépenses de même fréquence. C'est le cas pour les achats de pommes de terre, de vin, de sucre, de riz et de condiments ; pour les trois groupes de ménages les points correspondants sont au-dessus de l'ensemble à fréquence égale (graphique V). En revanche d'autres produits également stockables tels que le thé, les liqueurs, l'huile et les conserves, les consommateurs rares sont également de faibles consommateurs, il ne semble pas y avoir de stock. Le gain de précision est inférieur pour ces produits à celui auquel on s'attendait du fait de leur fréquence rare.

Il n'y a en définitive que 2 produits pour lesquels les corrélations négatives entre les semaines l'emportent encore sur les autres après 6 semaines d'enquête (groupe D) : ce sont les achats de pommes de terre et ceux de liqueurs, les calculs étant rapportés aux seuls consommateurs ; ces corrélations sont d'ailleurs à peine significatives.

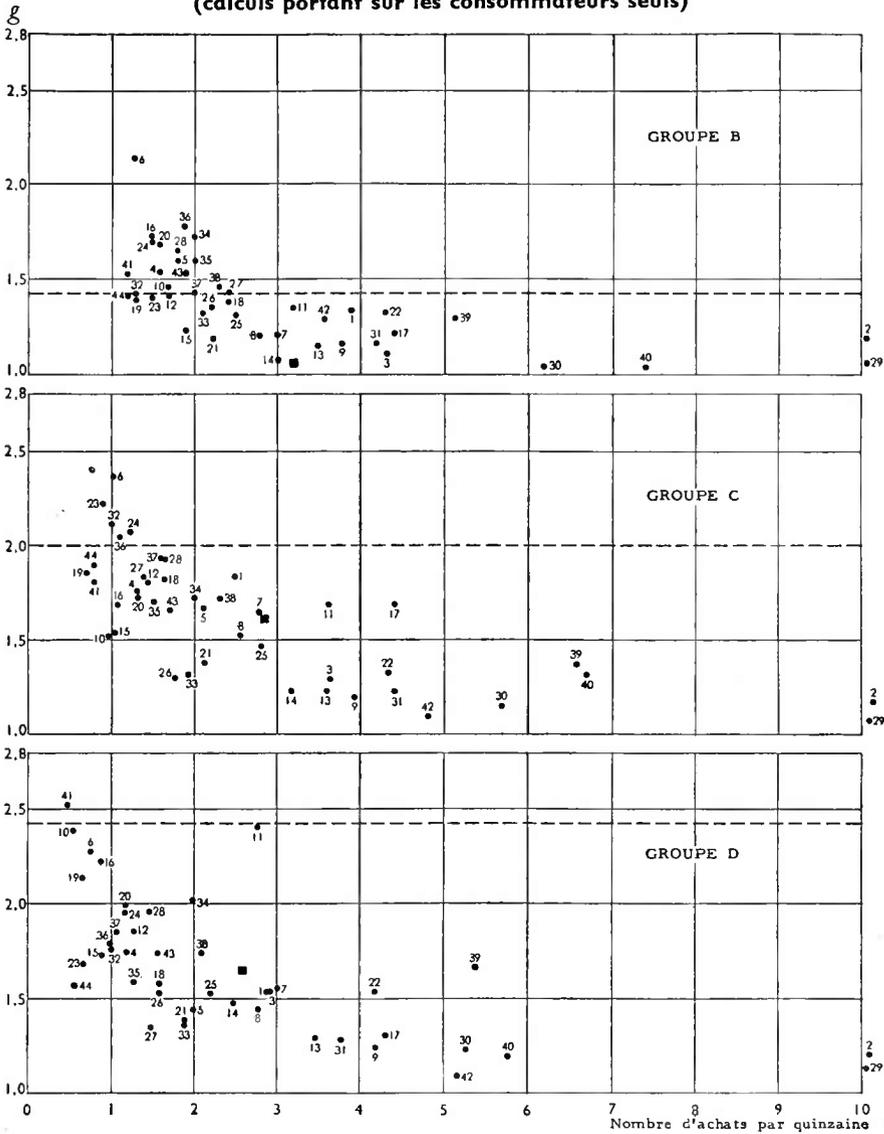
Ces graphiques nous donnent un ordre de grandeur de la réduction des écarts-types des dépenses lorsqu'on passe d'une semaine à deux, quatre ou six. L'écart-type de la dépense hebdomadaire estimée à partir de k semaines sera réduit par rapport à celui d'une estimation sur une seule semaine dans le rapport de $1/g$. Le coefficient de corrélation entre cette dépense et par exemple le revenu sera multiplié par g . En moyenne on trouve pour g les valeurs suivantes :

	Calculs sur l'ensemble	Calculs sur les consommateurs seuls
Groupe B.....	1,24	1,38
Groupe C.....	1,48	1,61
Groupe D.....	1,54	1,65

Les coefficients de corrélation avec le revenu restent cependant très faibles. En ce qui concerne l'échantillon D, ρ^2 est inférieur à 0,25 pour les 44 produits considérés et pour 19 d'entre eux il n'est pas significativement différent de zéro au seuil 1 %.

GRAPHIQUE V (1)

Gain de précision et fréquence d'achats
(calculs portant sur les consommateurs seuls)

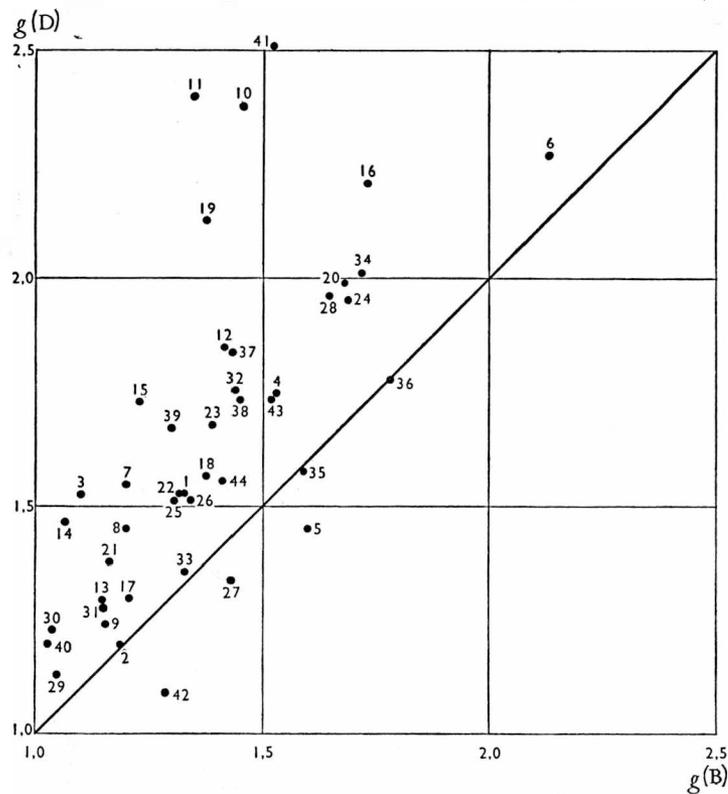


Enfin les graphiques VI montrent que si l'amélioration de la précision est accrue lorsqu'on passe de 2 à 6 semaines (échantillons B et D) la différence n'est pas sensible entre 4 et 6 semaines. S'il s'agissait du même échantillon notre mesure donnerait une amélioration toujours accrue, mais on a vu dans le cas de la dépense totale alimentaire que l'accroissement est de plus en plus faible. **Il ne semble donc pas intéressant d'allonger la période d'enregistrement des comptes alimentaires au delà de 4 semaines.**

(1) Voir nomenclature, graphique IV, page 67.

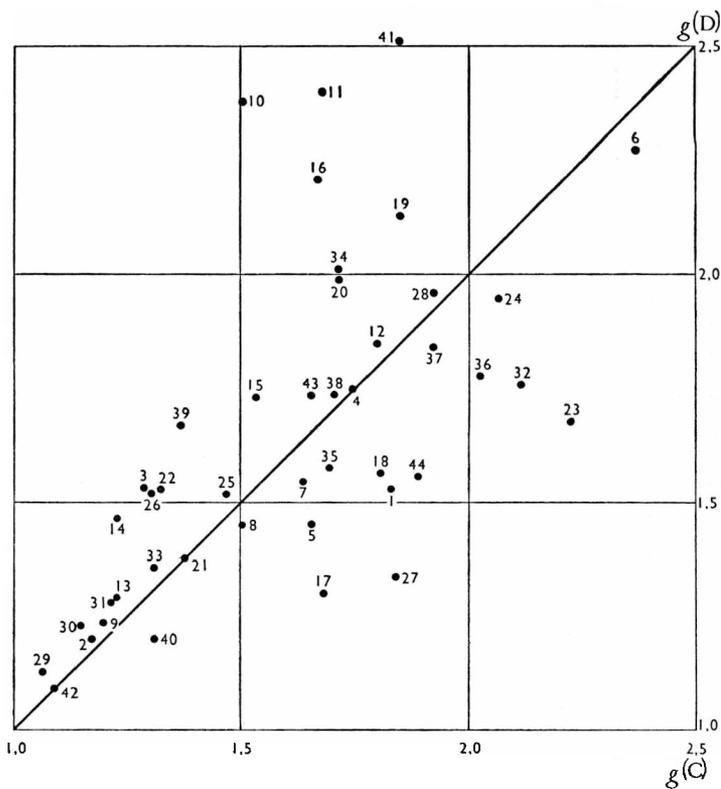
GRAPHIQUE VI-A

Comparaison du gain de précision entre les groupes B et D (1)



GRAPHIQUE VI-B

Comparaison du gain de précision entre les groupes C et D



(1) Voir nomenclature graphique IV page 67.

TABLEAU VIII

Variance entre semaines et variance entre ménages

PRODUITS	ÉCHANTILLON C			ÉCHANTILLON D		
	V'	V	F	V'	V	F
1 restaurant, cantine. . .	24 435	10 563	1,73	19 386	8 277	2,5
2 pain.	201	464	9,24	300	639	12,8
3 pâtisserie	1 008	1 180	4,68	617	370	3,6
4 farine.	76	36	1,85	103	46	2,7
5 pâtes	87	44	2,01	165	138	5,0
6 riz	27	10	1,50	26	8	1,9
7 carottes.	87	50	2,32	107	69	3,9
8 poireaux.	127	83	2,62	132	112	5,1
9 tomates	138	260	7,54	154	254	9,9
10 autres légumes	104	66	2,56	76	20	1,5
11 pommes de terre	3 411	798	0,94	792	118	0,9
12 conserves légumes	237	90	1,52	196	74	2,3
13 agrumes	450	66	0,59	283	349	7,4
14 pommes, poires	269	479	7,10	390	334	5,1
15 autres fruits frais	58	38	2,57	99	42	2,5
16 conserves de fruits	80	52	2,57	96	32	2,0
17 bœuf	9 071	3 218	1,41	3 794	4 414	7,0
18 veau.	3 827	1 489	1,56	1 992	1 572	4,7
19 mouton	1 338	437	1,30	944	295	1,9
20 porc frais.	1 118	581	2,08	1 064	358	2,0
21 jambon	520	429	3,30	378	392	6,2
22 charcuterie.	1 952	1 885	3,86	1 694	1 046	3,7
23 conserves de viande.	70	33	1,86	157	85	3,2
24 volailles.	3 429	1 042	1,22	2 647	1 002	2,3
25 œufs.	320	218	2,73	265	169	3,8
26 poissons	621	851	5,47	493	384	4,7
27 crustacés.	380	193	2,03	138	183	8,0
28 conserv. poissons	230	78	1,36	415	122	1,8
29 lait	180	964	21,42	318	882	16,6
30 fromage	384	899	9,35	507	808	9,6
31 beurre.	1 063	1 815	6,83	1 293	1 738	8,1
32 huile	326	95	1,17	362	139	2,3
33 margarine	72	102	5,67	70	91	7,8
34 sucre	218	85	1,56	262	72	1,7
35 chocolat	181	81	1,79	172	100	3,5
36 confitures	102	47	1,83	94	49	3,1
37 confiserie	136	66	1,93	110	47	2,6
38 condiments.	163	51	1,24	152	67	2,7
39 vins	3 414	2 880	3,37	12 028	3 142	1,6
40 bière, cidre	608	836	5,50	318	774	14,6
41 liqueurs.	2 516	971	1,54	1 281	288	1,3
42 eaux minérales.	133	586	17,62	119	723	36,5
43 café	265	133	2,01	304	133	2,6
44 thé	27	11	1,65	39	23	3,6

L'indice i se rapportant au ménage et variant de 1 à n .L'indice j se rapportant à la semaine et variant de 1 à k :

$$V' = \frac{1}{n(k-1)} \sum_i \sum_j (x_{ij} - \bar{x}_i)^2 \quad F = \frac{kV}{V'}$$

$$V = \frac{1}{n-1} \sum_i (\bar{x}_i - \bar{x})^2$$

Variance entre ménages et variance intra-ménage selon les produits

Comme pour la dépense alimentaire totale (voir p. 59), on a calculé les variances V et V' pour chacun des produits séparément. Le tableau VIII présente ces résultats pour les échantillons C et D. Les calculs sont faits sur l'ensemble des ménages interrogés et non sur les consommateurs seuls. On notera que la variance entre ménages V n'est supérieure à la variance intra ménage V' que pour les produits les plus fréquemment achetés : pain, lait, fromage... on peut comparer les résultats des tableaux VII et VIII). Le coefficient F de Fisher-Snédecor n'est pas toujours significativement supérieur à 1 même pour l'échantillon D (il ne l'est pas pour 5 produits en D et 10 en C). Ces résultats inclinent à penser qu'on ne peut guère envisager de faire des analyses de comportement valables, à partir de nomenclature trop détaillée, sauf en ce qui concerne des produits très courants ⁽¹⁾.

En conclusion, la durée d'enregistrement des dépenses modifie sensiblement la qualité de la mesure avec cependant les réserves suivantes : d'une part, la précision n'est guère améliorée au delà d'une certaine période, 4 semaines semble-t-il pour l'ensemble des produits étudiés précédemment ; d'autre part, cette amélioration est variable selon les produits car la fréquence d'achat joue un rôle non négligeable comme on pouvait s'y attendre. Ce dernier point serait à prendre en compte pour étudier un marché particulier. Les analyses que nous envisageons ne s'intéressent qu'à des groupements de dépenses répondant à une fonction précise pour le consommateur. A ce niveau les résultats qui précèdent laissent penser que les estimations des dépenses alimentaires obtenues pour 3 ou 4 semaines sont satisfaisantes pour l'analyse.

Cette enquête-pilote préparait une enquête plus vaste destinée à l'analyse et pour laquelle nous avons décidé de procéder ainsi : les dépenses alimentaires sont enregistrées au jour le jour pendant une durée de 3 semaines consécutives et les dépenses non alimentaires pendant toute l'année. Pour saisir les modifications du comportement alimentaire en cours d'année on a renouvelé l'enregistrement des comptes alimentaires à 4 mois d'intervalle ⁽²⁾.

La coupure entre dépenses alimentaires et non alimentaires est un peu artificielle puisqu'en fait la variable importante est la fréquence d'achat, mais elle nous a semblé être la plus commode et la plus objective pour imposer à des enquêtés. Nous pensons que cette méthode fournira des estimations satisfaisantes tout en ne demandant pas un effort trop considérable aux familles car les inscriptions de dépenses alimentaires sont de loin les plus astreignantes ⁽³⁾.

(1) Le test a été fait à partir du coefficient z de Fisher :

$$z = \frac{1}{2} (\text{Log } kV - \text{Log } V')$$

au seuil 1 %.

(2) Ce qui fait donc 9 semaines en tout dans l'année pour l'alimentation.

(3) Sur 72 inscriptions hebdomadaires notées par les ménages du groupe D, en moyenne 15 seulement concernaient des dépenses non-alimentaires.

PRÉSENTATION DES ANNEXES

Les annexes A à D donnent les moyennes et les écarts-types de chacune des dépenses alimentaires observées pour chacune des semaines d'enquête pour les 4 groupes de ménages. Pour l'ensemble des semaines on a présenté en outre le nombre moyen d'achats par ménage et le nombre de ménages ayant fait au moins un achat du produit concerné au cours de la période.

Pour ne pas alourdir la présentation, on ne donne ici que les calculs se rapportant à l'ensemble des ménages interrogés.

L'écart-type est calculé à partir de l'expression

$$\sigma_j = \sqrt{\frac{1}{n-1} \sum_i (x_{ij} - \bar{x}_j)^2} \text{ pour la dépense de la semaine } j.$$

Le nombre d'achats par ménage est en fait le nombre de jours au cours desquels un achat a été fait. C'est le quotient $\frac{N}{n}$ où n désigne le nombre de ménages interrogés et N le nombre de jours au cours desquels chacune des personnes ⁽¹⁾ ayant tenu ses comptes a effectué au moins un achat du produit concerné.

L'annexe E est une liste d'observations individuelles présentées pour le groupe D seulement et concernant la dépense alimentaire totale y compris les dépenses de boisson.

(1) On pourrait remplacer personne par ménage car peu de personnes en dehors de la mère de famille ont effectué des achats alimentaires. Il s'agissait plutôt de dépense de tabac, de cantine ou de dépenses personnelles ou d'« argent de poche ».

ANNEXE A

ÉCHANTILLON A

67 ménages ayant tenu leurs comptes pendant 1 semaine

PRODUITS	1 ^{re} SEMAINE			
	1	2	3	4
Restaurant, cantine	62,4	171,9	0,7	19
Pain.	69,3	28,5	5,7	67
Pâtisserie	34,4	34,0	1,8	55
Farine	5,7	7,9	0,6	31
Pâtes	10,6	11,5	0,8	44
Riz	2,4	4,9	0,3	18
Carottes	15,6	11,4	1,7	57
Poireaux	17,7	26,7	1,4	49
Tomates, salades	19,2	22,1	1,8	50
Autres légumes.	4,1	9,7	0,3	13
Pommes de terre	21,5	39,4	1,5	50
Conserves de légumes	10,0	16,1	0,6	29
Agrumes.	24,8	25,4	1,4	48
Pommes, poires	29,6	35,6	1,6	51
Autres fruits frais.	10,5	18,2	0,7	27
Conserves de fruits	9,5	17,9	0,5	25
Bœuf	110,0	67,5	2,0	63
Veau	43,2	53,6	0,7	37
Mouton	15,9	40,9	0,3	16
Porc frais.	19,7	37,2	0,4	22
Jambon.	29,5	73,4	0,9	38
Charcuterie	59,3	46,8	1,9	61
Conserves viande.	3,1	11,8	0,1	6
Volailles, lapin, gibier	56,3	58,4	0,6	38
Œufs	26,9	20,0	1,3	53
Poissons	25,3	23,4	0,8	43
Crustacés, coquillages.	12,2	18,9	0,5	25
Conserves de poissons	12,6	19,9	0,6	32
Lait frais.	66,1	30,1	5,7	66
Fromage.	61,9	49,6	3,3	65
Beurre.	80,4	42,6	2,2	67
Huile	15,7	17,3	0,6	34
Margarine	6,9	9,9	0,6	30
Sucre.	13,1	18,1	0,7	41
Chocolat.	9,6	13,6	0,7	34
Confiture	4,5	8,1	0,3	18
Confiserie.	3,8	7,9	0,4	20
Condiments	11,2	13,6	1,1	45
Vin.	71,3	70,6	2,9	60
Bière, cidre	14,3	23,0	1,4	27
Liqueurs	8,9	29,5	0,1	7
Jus de fruits.	14,3	20,6	1,7	34
Café	23,4	19,0	0,9	48
Thé, chicorée	1,7	4,6	0,2	12

1) Dépense moyenne x par ménage (dizaine de centimes). 2) Écart-type de la dépense σ_x .
3) Nombre d'achats par ménage. 4) Nombre de ménages ayant effectué au moins 1 dépense au cours de la semaine.

ANNEXE B

ÉCHANTILLON B

58 ménages ayant tenu leurs comptes pendant 2 semaines

1 ^{re} SEMAINE		2 ^e SEMAINE		TOTAL DES 2 SEMAINES			
1	2	1	2	1	2	3	4
44,7	74,5	55,2	115,8	100,0	162,9	1,6	24
67,9	30,2	69,0	31,6	135,7	53,0	10,7	58
46,5	46,7	41,4	41,0	87,9	80,4	4,0	53
5,1	7,9	4,3	8,5	9,4	12,7	0,8	31
11,2	13,3	7,6	10,4	18,8	16,3	1,5	48
2,2	5,1	1,7	4,0	3,8	6,5	0,4	19
13,3	12,1	11,5	10,4	24,8	19,4	2,7	52
12,1	10,9	12,8	14,3	24,9	21,5	2,4	51
14,0	19,9	18,4	18,7	32,4	34,3	3,2	49
6,8	11,5	5,7	11,0	12,5	18,4	0,8	26
21,6	52,9	15,3	24,5	36,9	59,2	2,4	43
12,1	18,2	8,3	14,2	20,4	24,8	1,2	40
29,5	27,5	29,0	26,7	58,5	48,0	3,2	53
24,2	31,8	19,0	22,6	43,2	51,7	2,3	44
7,3	12,3	4,6	11,2	11,9	20,5	0,8	26
5,0	9,2	5,5	10,8	10,5	15,0	0,7	26
121,0	89,0	118,3	91,2	239,0	150,9	4,3	57
39,3	51,6	46,7	50,0	86,1	82,7	1,7	41
20,1	61,7	6,3	17,9	26,4	65,0	0,3	16
29,3	35,9	21,1	37,2	50,4	51,9	1,1	39
19,9	27,4	23,6	26,5	43,5	47,3	1,8	48
70,7	59,9	62,1	55,6	132,8	89,4	4,1	56
2,9	8,1	3,6	9,3	6,4	15,1	0,3	13
58,9	83,9	38,7	64,3	97,5	103,7	1,1	39
29,8	22,4	22,9	19,4	52,7	32,6	2,4	56
26,1	31,6	31,9	35,9	58,0	56,4	1,5	41
13,6	27,1	14,2	24,0	27,9	40,6	1,2	30
11,7	17,8	9,4	13,7	21,1	23,3	1,1	36
60,3	34,2	62,3	44,4	122,6	74,9	10,7	58
60,6	41,8	56,8	45,4	117,4	83,6	6,1	57
85,4	58,4	83,2	58,3	168,6	100,1	4,0	55
20,4	37,9	7,7	13,6	28,1	39,1	0,8	36
10,1	19,8	8,2	12,2	18,3	25,9	1,3	36
17,7	19,4	11,7	13,2	29,4	21,0	1,8	51
12,0	18,7	11,7	22,7	23,7	29,9	1,2	36
8,6	17,1	6,8	15,7	15,3	25,5	0,6	20
4,8	11,3	4,9	11,3	9,7	18,2	0,8	22
16,2	19,8	8,9	12,4	25,1	23,5	2,0	50
82,3	119,1	47,3	48,6	129,5	132,9	4,5	51
22,7	40,6	22,6	43,2	45,4	81,8	3,7	29
13,8	44,8	8,6	24,5	22,4	55,5	0,3	13
15,6	24,1	8,5	14,2	24,0	32,0	2,3	36
25,4	39,8	17,9	19,6	43,4	42,6	1,4	44
1,2	3,3	2,8	6,6	4,0	8,0	0,5	21

1) Dépense moyenne \bar{x} par ménage (dizaine de centimes). 2) Écart-type de la dépense σ_x .
 3) Nombre d'achats par ménage. 4) Nombre de ménages ayant effectué au moins 1 dépense au cours des 2 semaines.

ANNEXE C

ÉCHANTILLON C

75 ménages ayant tenu leurs

PRODUITS	1 ^{re} SEMAINE		2 ^e SEMAINE	
	1	2	1	2
Restaurant, cantine	33,0	87,3	58,5	156,9
Pain	66,4	25,5	69,4	25,1
Pâtisserie	39,8	44,7	40,3	45,5
Farine	5,8	9,7	4,4	7,6
Pâtes	10,6	8,6	9,5	11,2
Riz	2,7	6,2	2,3	5,1
Carottes	12,0	11,6	12,9	12,9
Poireaux	15,1	14,7	12,7	13,5
Tomates, salades	18,2	18,2	17,7	22,1
Autres légumes	7,9	16,2	6,8	11,5
Pommes de terre	28,5	104,8	13,6	13,3
Conserves de légumes	10,0	15,6	10,5	18,3
Agrumes	28,1	29,8	27,6	25,6
Pommes, poires	24,9	27,5	24,6	27,6
Autres fruits frais	6,4	11,0	3,2	8,8
Conserves de fruits	4,8	8,4	3,8	9,1
Bœuf	139,7	146,3	112,5	82,4
Veau	49,4	91,3	46,0	48,2
Mouton	14,0	61,7	10,5	32,6
Porc frais	25,9	37,9	27,7	41,1
Jambon	27,4	31,9	23,1	26,1
Charcuterie	74,6	66,3	66,5	54,5
Conserves viande	4,1	10,5	2,5	8,7
Volaille, lapin, gibier	54,9	65,8	46,8	59,2
Œufs	25,9	20,8	27,1	19,8
Poissons	30,9	38,7	27,8	37,7
Crustacés, coquillages	12,2	18,4	12,3	21,7
Conserves de poissons	12,7	15,4	13,1	15,9
Lait	60,1	33,2	60,5	32,6
Fromage	53,4	36,0	52,1	31,2
Beurre	85,9	54,1	75,9	45,4
Huile	11,5	15,7	14,2	19,1
Margarine	8,6	13,7	8,1	13,4
Sucre	14,5	13,9	15,8	19,1
Chocolat	12,4	14,9	9,5	12,8
Confiture	2,5	8,8	2,4	6,8
Confiserie	5,1	9,4	4,9	9,7
Condiments	13,1	12,3	12,0	15,3
Vin	65,2	58,1	78,1	102,5
Bière, cidre	18,8	26,9	19,7	29,3
Liqueurs	20,3	69,1	6,3	30,4
Jus de fruits	16,9	26,5	16,4	25,9
Café	20,8	16,6	22,3	19,5
Thé, chicorée	2,5	8,3	1,5	5,0

1) Dépense moyenne x par ménage (dizaine de centimes). 2) Écart-type de la dépense σ_x . 3) Nombre d'achats

comptes pendant 4 semaines

3 ^e SEMAINE		4 ^e SEMAINE		TOTAL DES 4 SEMAINES			
1	2	1	2	1	2	3	4
72,2	225,6	61,1	180,2	224,9	411,1	2,3	34
66,8	22,1	68,0	26,5	270,7	86,2	22,7	74
38,5	38,1	33,7	47,3	152,3	137,4	7,1	73
6,5	11,1	5,7	9,7	22,3	23,8	2,1	59
9,8	10,5	10,2	11,4	40,1	26,5	3,9	70
2,8	5,7	2,6	5,3	10,4	12,8	1,1	41
11,8	10,7	10,4	10,6	47,0	28,4	5,5	74
11,2	12,3	10,9	12,4	49,8	36,4	4,8	70
16,2	18,8	17,4	17,0	69,6	64,6	7,2	69
5,0	9,9	4,3	8,8	24,0	32,6	1,4	55
20,0	35,9	19,2	31,6	81,3	113,0	6,3	65
12,9	17,5	9,9	13,6	43,3	37,9	2,6	68
32,6	35,2	38,4	36,0	126,7	105,6	6,7	69
20,8	27,4	18,8	20,9	89,2	87,5	5,2	62
3,9	8,7	2,0	6,6	15,5	24,5	1,1	39
3,8	10,7	4,5	13,6	16,9	28,8	1,0	35
112,5	73,5	123,1	79,6	488,0	226,9	8,9	75
37,1	66,5	30,9	47,9	163,5	154,3	2,7	61
6,1	19,6	5,5	22,7	36,2	83,6	0,6	28
26,5	35,8	26,0	36,3	106,2	96,4	2,2	62
25,4	31,3	24,3	24,7	100,2	82,9	4,2	74
70,2	48,4	63,7	60,8	275,0	173,7	8,7	75
2,5	7,1	3,8	10,3	13,0	22,9	0,6	26
49,7	62,5	47,2	52,9	198,5	129,1	2,2	68
25,6	22,0	29,2	22,9	107,7	59,1	5,5	74
30,5	35,9	25,9	32,6	115,1	116,7	2,9	62
11,7	22,3	9,8	24,8	45,9	55,5	1,8	48
11,1	15,6	14,3	16,6	51,2	35,4	3,1	69
63,1	33,6	62,8	33,3	246,5	124,2	22,2	75
46,2	36,8	50,5	33,4	202,3	119,9	11,2	74
70,0	50,4	80,9	53,2	312,8	170,4	8,5	72
10,5	16,5	10,6	21,9	46,8	39,0	1,6	63
7,7	12,4	5,8	10,0	30,2	40,4	2,4	49
15,4	14,4	13,3	15,3	58,9	36,9	4,0	74
13,1	17,5	9,7	13,2	44,7	36,0	2,8	69
4,2	10,3	5,2	16,1	14,3	27,3	0,8	25
6,4	15,9	6,5	13,4	22,9	28,5	2,1	51
11,8	14,2	11,9	13,0	48,9	32,4	4,6	74
63,7	62,9	69,3	62,4	276,0	214,7	12,3	70
26,2	50,4	20,2	32,1	84,9	115,6	7,8	43
4,9	22,2	19,2	71,7	50,6	124,6	0,5	23
14,9	25,5	16,5	26,9	64,6	96,9	5,9	46
18,8	19,9	16,2	16,2	78,2	46,1	3,2	71
1,1	3,6	2,1	4,4	7,1	13,4	0,7	31

par ménage. 4) Nombre de ménages ayant effectué au moins 1 dépense au cours des 4 semaines.

ANNEXE D

ÉCHANTILLON D

80 ménages ayant tenu leurs

PRODUITS	1 ^{re} SEMAINE		2 ^e SEMAINE		3 ^e SEMAINE	
	1	2	1	2	1	2
Restaurant, cantine	41,3	86,1	31,9	77,7	35,7	87,9
Pain	72,4	31,3	74,2	32,8	71,6	29,0
Pâtisserie	29,0	33,3	30,4	31,4	24,8	26,3
Farine	3,9	6,4	7,1	11,3	7,1	12,8
Pâtes	12,7	19,8	12,2	12,6	12,5	16,4
Riz	2,5	4,9	2,0	4,5	2,7	5,5
Carottes	12,1	13,5	14,3	12,3	13,7	11,0
Poireaux	13,7	14,3	13,8	15,1	15,1	16,0
Tomates, salades	18,3	18,2	19,0	21,5	16,2	17,1
Autres légumes	4,7	10,6	4,0	8,0	4,1	9,6
Pommes de terre	17,2	26,5	12,7	20,1	15,7	27,1
Conserves de légumes	11,9	14,7	11,4	14,7	7,8	13,9
Agrumes	21,5	24,0	26,5	24,9	27,9	27,2
Pommes, poires	18,9	25,7	15,5	22,1	21,8	31,0
Autres fruits frais	10,5	17,1	6,0	11,4	3,6	8,7
Conserves de fruits	4,8	14,7	5,2	12,2	4,8	9,2
Bœuf	114,0	83,3	117,5	77,9	121,2	107,2
Veau	40,7	56,6	44,4	65,7	44,0	60,6
Mouton	15,7	45,2	7,9	21,7	14,7	51,1
Porc frais	28,8	42,1	23,8	34,2	18,2	28,4
Jambon	18,7	27,1	20,6	25,1	21,1	26,9
Charcuterie	52,7	49,4	55,6	47,5	59,3	46,1
Conserves viande	3,1	15,9	3,6	11,5	4,5	13,5
Volaille, lapin, gibier	52,4	64,3	46,2	57,8	48,6	57,2
Œufs	22,3	18,6	20,5	21,4	22,7	20,3
Poissons	25,4	30,8	26,1	26,5	21,9	24,3
Crustacés, coquillages	9,4	17,2	9,9	18,6	10,2	19,4
Conserves de poissons	15,9	35,1	10,0	15,7	12,2	14,3
Lait	59,7	32,1	61,0	44,2	59,4	30,8
Fromage	47,6	37,7	46,7	32,4	52,7	37,5
Beurre	80,0	48,6	74,2	56,6	79,9	49,4
Huile	20,2	32,7	11,5	16,5	13,8	19,8
Margarine	7,2	11,7	9,8	13,0	9,5	11,7
Sucre	16,0	21,9	14,7	16,4	15,7	17,1
Chocolat	12,0	18,0	11,3	18,7	8,8	12,2
Confiture	5,8	11,7	5,8	10,9	3,2	7,3
Confiserie	5,2	10,3	3,9	8,2	2,7	6,0
Condiments	11,4	18,0	10,0	16,3	11,4	13,7
Vin	97,1	240,4	61,3	59,2	53,4	50,9
Bière, cidre	12,7	25,3	15,8	40,3	12,8	26,2
Liqueurs	10,0	51,4	11,5	52,9	11,9	40,6
Jus de fruits	17,6	28,2	17,9	30,2	18,3	30,7
Café	17,5	17,3	20,9	18,7	21,8	21,8
Thé, chicorée	2,4	9,7	1,9	8,9	1,3	4,6

1) Dépense moyenne x par ménage (dizaines de centimes). 2) Écart-type de la dépense σ_x . 3) Nombre d'achats

ÉCHANTILLON D

comptes pendant 6 semaines

4 ^e SEMAINE		5 ^e SEMAINE		6 ^e SEMAINE		TOTAL DES 6 SEMAINES			
1	2	1	2	1	2	1	2	3	4
40,7	90,4	35,5	78,6	84,2	333,3	269,2	545,9	3,6	33
72,5	30,6	68,8	26,6	71,6	28,1	431,1	151,7	32,9	79
21,6	26,0	22,7	24,6	32,7	34,1	161,1	115,3	8,6	79
6,4	11,4	6,4	10,3	7,7	14,8	38,7	40,7	3,1	67
13,6	20,3	11,0	13,8	14,0	15,3	75,9	70,6	6,1	79
3,1	5,2	2,0	4,1	4,2	7,5	16,4	17,1	1,7	56
13,3	12,2	13,2	14,7	10,8	11,1	77,4	49,8	8,5	77
14,4	15,9	14,1	14,3	13,2	13,9	84,3	63,6	8,0	76
17,7	17,5	16,1	18,8	18,2	23,3	105,4	95,7	11,6	74
3,6	8,6	3,6	9,0	2,5	8,5	22,5	26,5	1,4	56
14,3	26,0	15,4	32,4	16,0	34,1	91,2	65,3	6,9	65
9,1	17,1	11,6	17,7	10,9	13,6	62,5	51,4	3,7	74
25,8	21,2	31,2	23,7	30,3	22,8	163,3	112,2	10,3	78
18,8	28,7	14,6	22,4	13,7	22,2	103,3	109,5	6,1	65
3,0	8,1	2,5	10,1	1,7	5,7	27,2	38,7	1,8	52
3,6	8,2	4,1	8,5	4,1	9,2	26,4	34,1	1,7	48
117,4	79,0	104,7	72,4	112,3	97,9	687,4	398,6	12,9	80
41,7	54,1	35,8	51,1	43,4	52,3	250,0	237,9	4,1	70
6,1	23,5	4,6	15,8	7,1	22,6	56,0	103,0	0,9	33
29,0	38,7	24,0	34,0	22,3	32,3	146,1	113,6	3,2	73
18,5	30,9	19,3	26,5	19,8	22,8	117,9	118,7	5,0	72
50,9	40,9	63,9	64,4	62,1	45,3	344,4	193,9	12,5	79
6,0	21,9	2,9	11,5	3,3	11,1	23,3	55,3	1,0	34
31,4	44,9	52,7	53,3	52,2	58,9	283,5	190,0	3,2	72
22,8	19,1	18,6	16,6	23,5	22,6	130,5	78,8	6,6	79
26,2	34,2	23,8	23,7	22,9	28,5	146,4	117,5	4,3	71
8,5	14,0	11,3	17,4	8,9	16,6	58,1	81,2	3,1	55
14,7	19,2	12,9	20,5	12,6	18,6	78,3	66,3	4,1	73
61,8	34,6	60,3	31,7	59,6	30,3	361,8	181,7	32,2	79
46,7	34,2	45,2	33,6	49,1	34,8	287,9	170,5	15,8	80
78,1	62,1	71,4	47,6	74,5	52,8	458,2	250,1	11,2	79
10,9	19,5	6,7	11,4	13,5	17,8	76,5	70,8	2,5	71
8,2	10,6	9,3	13,4	10,6	12,6	54,6	57,1	4,3	62
15,3	18,3	12,3	11,1	15,8	15,7	89,9	51,0	5,8	79
8,6	13,4	8,1	14,1	11,7	15,9	60,6	59,9	3,7	74
5,9	17,2	1,9	5,2	5,9	11,2	28,5	42,0	1,6	42
6,1	14,0	5,7	14,3	5,7	14,9	29,2	41,2	2,1	51
11,6	13,8	9,7	10,1	11,6	13,9	65,6	49,2	6,2	80
70,9	84,1	58,0	62,9	66,0	59,9	406,5	336,3	16,1	79
12,8	26,4	15,1	32,4	17,0	39,3	86,2	166,7	7,4	34
1,0	7,3	4,7	23,5	6,0	20,6	45,0	102,0	0,5	23
15,3	28,0	13,8	26,9	15,9	27,6	98,9	161,2	8,9	46
16,1	16,9	20,8	22,2	20,4	20,1	117,5	69,1	4,7	78
1,1	5,2	1,9	6,1	2,5	2,6	11,2	28,8	0,9	43

par ménage. 4) Nombre de ménages ayant effectué au moins 1 dépense au cours des 6 semaines.

