C.R.E.D.O.C.
BIBLIOTHÈQUE

C. R. E. D. O. C.

G. ROTTIER

ENQUÊTES PAR SONDAGE ET ANALYSE DE LA DEMANDE

Sou1963-2050

1963

Enquêtes par sondage et analyse de la demande / G. Rottier. (1963).

CREDOC•Bibliothèque

DE DOCUMENTATION SUR LA CONSOMMATION
8e)

R6 003(1)

C.R.E.D.O.C.
BIBLIOTHÈQUE

ENQUETES par SONDAGE

et

ANALYSE de la DEMANDE

CREDOC
N° 305

R 6 403
(1er ex.)

C.R.E.D.O.C.
BIBLIOTHÈQUE

Université de Paris

Faculté de Droit et des Sciences Economiques

ENQUETES par SONDAGE

et

ANALYSE de la DEMANDE

Essai sur les courbes d'ENGEL

Thèse pour le Doctorat ès-Sciences Economiques

présentée et soutenue le 2 Juin 1961 à 14 h.

par

Georges ROTTIER

Président : M. H. GUITTON, Professeur

Suffragants : MM. J. MARCHAL, Professeur

G. Th. GUILBAUD, Professeur associé

La Faculté n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

C.R.E.D.O.C.
BIBLIOTHÈQUE

ENQUETES PAR SONDAGE ET ANALYSE DE LA DEMANDE

- Essai sur les courbes d'ENGEL -

Table des matières

		Pages
CHAPITRE I :	L'étude économique du comportement du consommateur	9
Section 1.	La place de l'économétrie et des autres sciences humaines dans l'étude de la consommation	11
Section 2.	Les orientations de l'analyse de la demande : problèmes de prévision et problèmes de décision	15
Section 3.	Le renouvellement des méthodes lié au développement des enquêtes par sondage	18
Section 4.	Les bases théoriques de l'analyse de la demande d'un bien de consommation courante	20
Section 5.	Les bases théoriques de l'analyse de la demande d'un bien durable de consommation	28
Notes		37
CHAPITRE II :	L'observation de la consommation	41
Section 1.	Le rôle des enquêtes par sondage dans l'étude de la consommation	43
Section 2.	Notions de théorie des sondages	45

Pages

1 - la notion de sondage probabiliste 45
2 - l'étude d'une moyenne dans un sondage simple 46
3 - la stratification de l'échantillon 49
4 - les sondages à plusieurs degrés 54
5 - sondages et plans expérimentaux 55

Section 3. L'enquête française de 1956 57

1 - le plan de sondage 57
2 - la méthode d'enquête 63
3 - les conditions d' exécution de l'enquête 66
4 - le fichier mécanographique 71
5 - la qualité des résultats 75
6 - le questionnaire d'enquête 77

Section 4. Notes bibliographiques 79

1 - le développement des enquêtes de consommation en France depuis 1954 79
2 - les principales enquêtes étrangères récentes 97

Notes 101

CHAPITRE III : L'explication de la consommation : les courbes d'ENGEL 103

INTRODUCTION 105

PREMIERE PARTIE : Problèmes de méthode 111

Section 1. Les techniques d'estimation statistique 111

1 - l'estimation dans un modèle linéaire 111
2 - l'estimation à partir de données groupées 115

Pages

Section 2. La forme des courbes d'ENGEL 119

1 - généralités 119
2 - modèles linéaires des courbes d'ENGEL 126

A - Lois à élasticité constante 129

B - Lois semi-logarithmiques 139

C - Vérification de la linéarité d'une loi 144

3 - un exemple de modèle non linéaire des courbes d'ENGEL 155

Section 3. La définition des variables introduites 175

1 - l'explication par le revenu ou par la dépense totale de consommation 175
2 - l'évolution des dépenses et des quantités consommées 183

Section 4. L'influence d'autres facteurs que le niveau de vie 188

1 - la taille des ménages et les unités de consommation 188
2 - une tentative de détermination statistique des unités de consommation 199
3 - variables qualitatives affectant les courbes d'ENGEL 202

DEUXIEME PARTIE : Résultats déduits de l'enquête française de 1956 205

Section 1. Les courbes d'Engel pour l'ensemble de la population non-agricole française 205

1 - introduction 205
2 - les méthodes d'analyse choisies 206
3 - les résultats 208

Pages

Section 2. L'effet de l'urbanisation sur la consommation 229

1 - les méthodes d'analyse 229
2 - l'étude des consommations alimentaires 239
3 - l'influence combinée de l'urbanisation et du revenu 257
4 - l'effet d'autres facteurs qualitatifs 263

Notes 267

CONCLUSION 277

BIBLIOGRAPHIE 285

o

o o

C.R.E.D.O.C.
BIBLIOTHÈQUE

CHAPITRE I

L'ETUDE ECONOMIQUE DU COMPORTEMENT DU CONSOMMATEUR (1)

SECTION 1

LA PLACE DE L'ECONOMETRIE ET DES AUTRES SCIENCES HUMAINES DANS L'ETUDE DE LA CONSOMMATION

"L'aboutissement de toute production et de toute distribution est ce personnage central de l'économie du XXème siècle : le Consommateur" (2).

Central, ce personnage l'est sûrement. Connaître l'évolution de ses besoins et de ses choix, ou les conditionner par les techniques de la publicité, est la tâche essentielle des services de recherche commerciale d'innombrables entreprises. Prévoir l'orientation de ses dépenses pour y ajuster le développement à moyen terme de l'économie est la première démarche de l'établissement d'un plan, sitôt qu'une situation de pénurie est dépassée (3).

Consommation et Civilisation

En même temps, le consommateur est un personnage mystérieux, complexe et changeant qui échappe à des règles d'étude simple. Son comportement et ses choix sont le reflet d'une civilisation et d'une époque. Conditionnés par des facteurs économiques nombreux, ils reflètent en outre les valeurs que, consciemment ou non, reconnaît une société. "L'économie de la consommation est en fait le domaine de plusieurs disciplines. Elle introduit la psychologie et la sociologie des besoins et des choix, les décisions économiques des consommateurs qui en résultent et l'effet de ces décisions sur l'allocation des ressources dans une économie" (4).

Dans les chapitres qui suivent, nous nous intéresserons uniquement au consommateur comme objet d'étude, avec les techniques qui lui sont propres, pour l'économiste ou le statisticien. Dans ce premier chapitre, nous tenterons rapidement, en particulier, de décrire certains des schémas ou des modèles qu'introduit l'économètre (car il n'existe plus actuellement d'économie de la consommation qui ne soit une tentative de mesure) pour expliquer les décisions d'achat des consommateurs dans divers domaines.

Consommation et dépense

Remarquons que la phrase précédente introduit une restriction à l'intérieur du domaine d'études propre à l'économiste. Nous sommes, en effet, passés de la "consommation", qui consiste à "faire disparaître en les utilisant des biens déjà existants", (5) à la "dépense du consommateur", c'est à dire à l'acte par lequel la propriété d'un bien économique est transférée d'une entreprise à un ménage. Selon une coutume maintenant trop bien établie pour qu'on l'abandonne, l'économie de la consommation est essentiellement l'étude des décisions qui déterminent les flux monétaires allant des ménages vers les autres secteurs de l'économie. En outre, comme nous le verrons en considérant quelques cas particuliers, ces décisions sont saisies à travers leurs effets objectifs sur les flux monétaires et non à travers une analyse plus profonde des motifs qui les expliquent.

L'intérêt exclusif attaché aux dépenses de consommation mutile indéniablement la réalité. Trois exemples qui sont chacun devenus d'une particulière importance le montreront aisément.

Un des problèmes les plus urgents de notre époque est de comparer et d'expliquer les niveaux de vie et de consommation de populations qui ont atteint des niveaux différents de développement économique. Ces populations peuvent, soit appartenir à des pays différents (ex. : l'Inde, la France et les Etats-Unis), soit représenter des catégories de la population d'un même pays (ex. : les populations agricole et non agricole). Ces comparaisons ne peuvent

être faites en termes de dépenses de consommation. A des structures économiques (nationales ou d'entreprise) distinctes correspondent des degrés très différents dans le recours aux services du marché pour satisfaire aux besoins des consommateurs. L'autoconsommation, la valeur des services fournis par le travail non rémunéré des membres du ménage, les services fournis par les collectivités publiques, les revenus réels découlant de l'utilisation du patrimoine de biens matériels du ménage, doivent être pris en considération. Les techniques habituelles de la comptabilité nationale sont inadéquates pour mesurer ces éléments ; les techniques d'analyse des dépenses des consommateurs renoncent à les expliquer. Les unes et les autres ont été construites en référence à une économie de marché où toute consommation est l'aboutissement d'une transaction.

Mais même à l'intérieur d'une catégorie sociale donnée, dans un pays industriel moderne, considérer seulement les dépenses des consommateurs ne suffit pas. D'une part, la demande de biens durables, comme nous le verrons plus bas, est en partie le reflet de la demande pour le flux de services qui découle de la possession de ces biens. Cela explique la caractéristique fondamentale des modèles expliquant la demande de biens durables ; en même temps, cela justifie l'importance qu'on attache maintenant à l'âge du chef de ménage parmi les facteurs expliquant les dépenses de consommation : la raison principale n'est pas que l'âge et le nombre des enfants expliquent les besoins d'une famille, mais, qu'avec l'âge, s'accroît l'accumulation de biens durables, meubles ou immeubles, dont les services comptent notablement dans le niveau de vie des familles étudiées.

D'autre part, une définition correcte de la consommation devrait tenir compte de la valeur pour les consommateurs des services fournis par des mécanismes collectifs : éducation, santé, urbanisme, etc... De ne pas compter des services, ni de les intégrer dans nos schémas d'explication, il n'y a qu'un pas jusqu'à les négliger ou les sous-estimer, aussi bien dans l'étude que dans l'action. Et le risque est alors grand que "l'économie se trouve orientée vers la satisfaction des besoins humains les moins urgents" (6).

"N'est-il point paradoxal que cette ultime étape du processus économique, objet même de l'activité des hommes, n'ait suscité, semble-t-il, jusqu'à ces dernières années, qu'une curiosité médiocre.

La production des biens et services, son financement, la rémunération des producteurs, ont eu le pas sur la consommation" (7).

Certes, mais identifier consommation et dépenses des consommateurs, c'est à dire consommation et demande sur le marché, est aussi donner le pas à la production des biens et des services par un autre biais. La consommation n'est plus en effet étudiée du point de vue des besoins des consommateurs, de leur hiérarchie et de la façon dont ils sont satisfaits. Elle est seulement étudiée dans la mesure où son évolution conditionne l'activité productrice des divers secteurs.

Il s'agit là d'un domaine qui est loin d'être négligeable. Il convenait cependant de remarquer en introduction qu'identifier la consommation à la demande ne permet pas d'en épuiser l'étude.

SECTION 2

LES ORIENTATIONS DE L'ANALYSE DE LA DEMANDE : PROBLEMES DE PREVISION ET PROBLEMES DE DECISION

L'étude économique des dépenses des consommateurs présente deux volets. Le premier est l'analyse de l'incidence, sur l'économie dans son ensemble, de l'évolution de la consommation ; le second est l'analyse des facteurs économiques qui conditionnent le comportement des consommateurs. Nous considérerons uniquement le second volet dans la suite de ce texte ; la séparation entre les deux ne peut cependant pas être totale, ne serait-ce que pour des raisons accidentelles. L'étude du comportement du consommateur est essentiellement du domaine de l'économie appliquée ; elle exige d'importants moyens matériels et son développement, dans les cinquantes dernières années, a été largement orienté vers la solution de problèmes d'action posée à des administrateurs. Ces problèmes ont été soit des problèmes de prévision, au niveau d'une nation, d'un secteur économique ou d'une entreprise, soit des problèmes de décision. Les deux types de problèmes peuvent être illustrés en considérant l'orientation des travaux d'économie de la demande faite depuis l'entre deux guerres par les services du Ministère de l'Agriculture de Washington. D'une part, estimer l'évolution probable à moyen terme des excédents de production agricole des Etats-Unis, ce qui exige une prévision de la demande des consommateurs américains pour les principaux produits agricoles ; d'autre part, évaluer l'incidence, sur un marché donné, de diverses mesures possibles d'intervention en matière de prix agricoles, c'est à dire aider à prendre une décision en en dégageant les conséquences probables.

Les deux types de problèmes ne sont pas symétriques. Le second est uniquement un problème pour l'économètre ; il s'agit d'étudier, dans un environnement sociologique qu'il est légitime de considérer comme stable, l'effet d'une modification d'une variable économique donnée. Le premier problème, par contre, déborde le domaine d'explication de l'économiste : la prévision de la consommation, si elle doit avoir toute son utilité, ne peut pas être réduite à une perspective (8).

On connaît la distinction entre les deux notions. Le terme de perspective a d'abord été introduit par les démographes : une perspective démographique indique l'évolution future de la population sous l'hypothèse explicite que les taux de nuptialité, de fécondité, de mortalité, ou bien ne se modifient pas, ou bien se modifient d'une façon bien définie, suggérée par l'examen du passé. De la même façon, une perspective de consommation est établie en se donnant l'évolution d'un petit nombre de facteurs économiques : revenus, répartition de la population par âge, etc... ; et en en déduisant l'évolution probable des dépenses des consommateurs sous l'hypothèse explicite que des lois de comportement préalablement établies ne se modifieront pas.

Place des facteurs économiques

L'analogie entre une perspective démographique et une perspective de consommation risque cependant d'être trompeuse, car la seconde mutile plus le réel que la première. Alors qu'à moyen terme (cinq à quinze ans), l'évolution de la population est à peu près déterminée dès que sont connus les principaux taux démographiques, le passage d'une perspective à une prévision est beaucoup plus hasardeux en matière de consommation. Cela vient en partie de ce que le comportement des consommateurs est fondamentalement stochastique et donc imprévisible en toute rigueur : on ne peut que dégager des tendances vraisemblables, à partir desquelles établir une prévision ponctuelle implique un pari. Mais cela vient surtout de ce que les facteurs introduits n'expliquent qu'une partie du comportement des consommateurs.

Ce comportement dépend entre autres de facteurs économiques, tels que les revenus, les prix et les patrimoines. Ces facteurs jouent d'ailleurs d'une façon qui varie notablement d'un marché à un autre, et nous verrons sur quelques exemples qu'il n'existe pas un modèle unique du comportement économique du consommateur, mais un éventail de modèles dinstincts qui sont tous incomplets et imparfaits et dont nous ne considérerons que le plus simple dans les chapitres suivants.

Or l'action de ces facteurs économiques s'insère dans un contexte plus large. Les décisions des consommateurs ont des dimensions psychologiques, physiologiques, sociologiques et psycho-sociologiques qu'on ne peut éviter d'introduire dans une tentative d'explication complète (9). Les éléments sociologiques et psycho-sociologiques sont particulièrement importants. Si l'on se borne à étudier quelques très grands groupes de dépenses, comme l'alimentation, l'habitation, etc... auxquelles s'appliquent les lois classiques d'Engel, on observe certes de grandes analogies d'un pays à l'autre (10). Mais dès qu'on veut entrer plus dans le détail, les facteurs extra-économiques jouent un rôle croissant dans l'explication. La structure des consommations est étroitement liée à une civilisation ; elle reflète les valeurs qui la marquent, les aspirations et les mentalités des hommes qui la vivent. Plus inconsciemment que consciemment, ces éléments motivent profondément les comportements des consommateurs qui, bien que conditionnés par des paramètres économiques, ne traduisent pas, le plus souvent, un calcul économique conscient.

Ces considérations ont une particulière importance pour expliquer la demande de biens durables. Dans une civilisation où ce qu'on a, et surtout ce qu'on a et que l'autre n'a pas, présente plus de signification que ce qu'on est, la demande de biens durables ne s'explique pas seulement par les services objectifs qu'on retire de la possession du bien, mais plus encore par la satisfaction qu'on retire de sa possession indépendamment de l'"usus".

En face de cela, la théorie du comportement du consommateur nous représente un homme qui, avec une structure de choix donnée, modifie ses comportements objectifs sous l'effet de stimuli qui ont des variations des revenus, des prix ou d'autres facteurs économiques. La démarche qu'en déduit l'économètre selon des lignes que nous décrirons est légitime, mais n'épuise pas la réalité. Il convient aussi d'expliquer comment se forment et se modifient ces structures de choix elles-mêmes. Cette tâche n'est pas du domaine de l'économiste et n'a été encore que très imparfaitement assumée par les spécialistes d'autres disciplines. Il en résulte un malaise. Les notions fondamentales de besoins, de niveau de vie, de style de vie, de motif d'achat sont aberrantes par rapport aux catégories et aux concepts de l'économiste. Il ne parvient ni à leur donner de définition opératoire, ni à les incorporer à ses schémas d'explication. Mais il sait en même temps que ces schémas seront partiels et pauvres s'il ne les prend pas en considération.

SECTION 3

LE RENOUVELLEMENT DES METHODES LIE AU DEVELOPPEMENT DES ENQUETES PAR SONDAGE

L'étude de l'influence quantitative, sur les dépenses des consommateurs, des divers facteurs économiques retenus a marqué un progrès considérable dans les deux dernières décennies, grâce au développement des techniques d'enquêtes par sondage auprès des ménages. Les données provenant d'enquêtes présentent en effet une série d'avantages qu'aucune autre source statistique ne partage. KOOPMANS a pu écrire, sans exagération, que "la méthode des sondages peut bien acquérir une importance semblable, pour l'étude du comportement des agents économiques, à celle qu'a acquise le microscope pour l'étude des phénomènes biologiques" (11).

Alors que toutes les autres informations statistiques sont des données globales qui se rapportent à l'ensemble d'un marché ou d'une économie et qui ne permettent de saisir que des relations agrégatives entre des moyennes, les enquêtes fournissent des informations individuelles qui permettent une analyse beaucoup plus profonde des comportements des ménages (12). Quand on se borne à l'analyse de lois de comportement individuelles qui est le centre de l'économie de la consommation, on peut ainsi éviter les difficultés considérables que pose l'agrégation des lois individuelles pour parvenir à une relation sur le marché. Certes, cette difficulté se présente à un stade ultérieur, dès qu'on veut utiliser les enseignements obtenus à partir de l'étude des données individuelles pour expliquer l'évolution de l'ensemble ou d'un secteur de l'économie. Il est cependant bon que les difficultés puissent être ainsi sériées.

Les enquêtes par sondage fournissent ensuite des informations beaucoup plus nombreuses que les autres sources. Alors que les séries de consommation apparente, de prix et de revenu sur un marché fournissent rarement plus d'une vingtaine d'observations de chaque variable, il est fréquent maintenant d'analyser des enquêtes fournissant plusieurs milliers ou même dizaines de milliers d'observations.

La richesse des modèles étudiés est donc beaucoup plus considérable, puisque la taille des échantillons permet d'introduire simultanément un plus grand nombre de variables (13). En outre, en jouant sur la structure de l'échantillon d'une façon inspirée par la technique des plans d'expérience, on peut neutraliser ou isoler l'influence de certaines variables ou choisir l'échantillon de façon à ce que l'estimation des paramètres les plus importants soit aussi précise que possible. C'est ainsi que pour étudier l'épargne personnelle par sondage, il convient de stratifier l'échantillon à la fois suivant le niveau du revenu et l'activité du chef de ménage, afin de tenir compte du fait que la majeure partie de l'épargne provient des ménages à haut revenu et que la répartition du revenu entre l'épargne et la consommation ne suit pas les mêmes lois chez les travailleurs indépendants et chez les salariés.

Enfin, outre les éléments objectifs que nous considérerons seulement dans deux exemples que nous traiterons, les enquêtes permettent aussi de saisir des éléments subjectifs, commes les intentions d'achat, les opinions ou les anticipations des consommateurs. Bien que l'interprétation de ces éléments subjectifs soit délicate, leur considération est nécessaire à la solution de nombreux problèmes. C'est le cas, par exemple, de l'évolution à court terme des achats de biens durables. Les enquêtes sur les intentions d'achat, effectuées depuis une dizaine d'années à l'Université de Michigan, sous la direction de KATONA, et depuis quelques années en France par l'I.N.S.E.E. sous une forme voisine, ont fourni des indications très utiles aux conjoncturistes. De même, des enquêtes d'opinion sur le logement, faites en 1955 par l'I.N.S.E.E. et le C.R.E.D.O.C. (14) (et répétées en 1961), ont fourni à peu près la seule méthode pour évaluer les conditions de l'équilibre sur le marché du logement.

Historiquement, les premières enquêtes faites auprès des consommateurs portaient sur la totalité des budgets familiaux. Quelque importantes qu'elles soient encore, ces enquêtes de "budgets de familles" ne représentent plus qu'une espèce d'un genre plus large. Des enquêtes de consommation plus spécialisées, comme les enquêtes alimentaires, s'y sont d'abord ajoutées, puis des enquêtes plus délicates, visant à la fois à saisir les revenus, les patrimoines et les diverses formes de l'épargne des ménages (15).

Nous évoquerons rapidement les principes selon lesquels on analyse les résultats de ces enquêtes dans les deux paragraphes suivants, qui traiteront respectivement des courbes d'Engel et de la demande de biens durables. Dans les chapitres suivants, nous examinerons en détail les difficultés que présente l'application de ces principes dans le cas des courbes d'Engel.

SECTION 4

LES BASES THEORIQUES DE L'ANALYSE DE LA DEMANDE D'UN BIEN DE CONSOMMATION COURANTE

Bien que l'analyse des courbes d'Engel repose sur des intuitions déjà anciennes, puisque la première publication d'Engel est plus que centenaire, ce n'est que quelques années avant la seconde guerre mondiale qu'ALLEN et BOWLEY (16) lui ont donné une forme économétrique rigoureuse.

Considérons les achats C_i d'un bien de consommation courante, au cours d'un intervalle de temps donné, par un ménage i dont on connaît également le revenu R_i au cours du même intervalle de temps. Si l'on porte les valeurs observées de C_i et R_i pour un échantillon de ménages sur un graphique cartésien, on observe une corrélation assez nette entre ces deux variables. Les observations sont rassemblées autour d'une loi de tendance moyenne Γ qui est la courbe d'Engel pour le produit considéré (fig. 1).

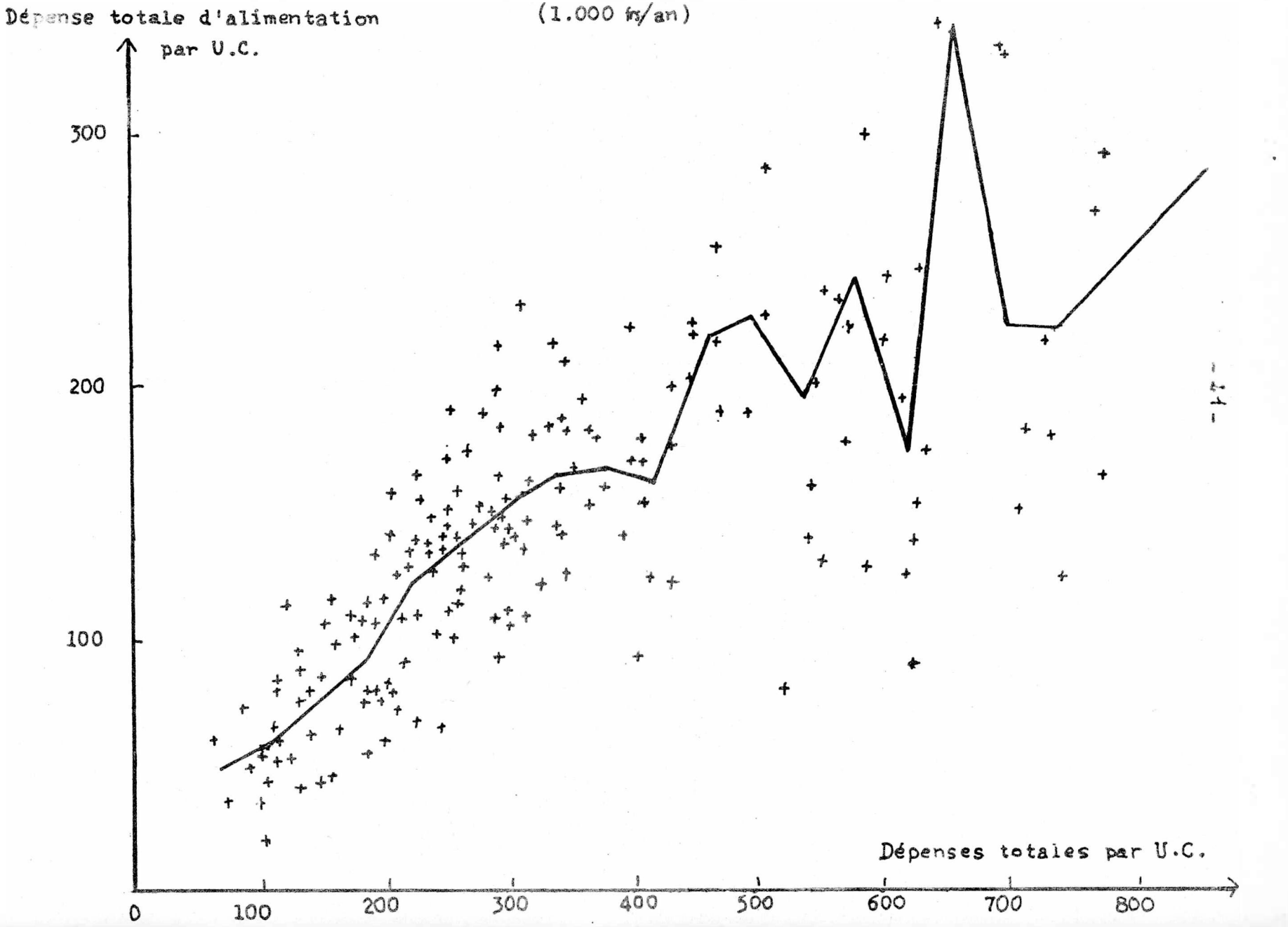
Le graphique ainsi obtenu n'a cependant aucun pouvoir explicatif ; il est uniquement une description de l'échantillon considéré. Pour passer de la description à l'explication, deux étapes doivent être franchies. Il faut d'abord qu'on admette comme raisonnable l'hypothèse que le revenu influence les dépenses des consommateurs, ensuite que la courbe Γ de la figure 1 corresponde bien à la relation théorique ainsi postulée, qu'elle puisse lui être "identifiée".

Le raisonnement proposé par ALLEN et BOWLEY est bien connu. Si l'on considère un individu ayant une surface de choix parétienne, la théorie de la demande du consommateur permet de déduire de cette surface de choix une relation :

$$(1) \qquad q_k = f\ (R, p_1 \dots p_k \dots p_n)$$

Figure 1

Enquête pilote budgets de famille 1956



qui est la courbe de demande individuelle du bien k par l'individu considéré : la consommation q_k est une fonction du revenu de cet individu et de l'ensemble $p_1 \ldots p_n$ des prix des biens de consommation disponibles sur le marché.

Si l'on considère, au même moment, une population de consommateurs, ayant les mêmes surfaces de choix, les prix $p_1 \ldots p_n$ sont, à de faibles variations locales près, les mêmes pour tous ces consommateurs. Leurs consommations et leurs revenus décrivent donc une courbe de niveau de la surface (1) :

$$(2) \quad q_k = f \left[R, (p_1 \ldots p_n = \text{Ctes}) \right]$$

Moyennant quelques hypothèses complémentaires, la courbe Γ de la figure 1 peut être identifiée avec la courbe représentant la fonction partielle de demande (2).

Pour tenir compte du caractère stochastique de la relation entre le revenu et la consommation, que met en évidence la figure 1, on transforme le schéma (2) en écrivant que la consommation C_i d'un ménage de revenu R_i est en fonction de R_i à un aléatoire U_i près. On obtient le modèle :

$$(3) \quad C_i = f(R_i) + U_i$$

Une fois spécifiée la forme de la fonction f (R) et moyennant quelques hypothèses raisonnables sur la distribution des résidus aléatoires U_i, les paramètres numériques de f (R) peuvent être estimés convenablement. Dans leur étude ALLEN et BOWLEY avaient admis que la fonction f (R) était linéaire. Ils retenaient donc le modèle :

$$(4) \quad C_i = \alpha R_i + \beta + U_i$$

et admettaient que l'aléatoire U_i satisfaisait aux conditions classiques qui justifient l'ajustement par les moindres carrés (conditions que nous examinerons dans le chapitre III). Dans le cas, si ρ^2 est le carré du coefficient de corrélation

linéaire entre C et R, on sait que ρ^2 mesure la part de la variance des consommations qui est expliquée par l'influence du revenu. Comme les ajustements, qui sont effectués sur des données groupées, donnent des ρ^2 assez élevés (voisins de 2/3), on est fondé à admettre que l'hypothèse exprimée par le modèle (4) est non seulement quantifiée, mais justifiée empiriquement.

Ce raisonnement classique appelle des critiques qui mettent en cause la nature logique des hypothèses théoriques introduites. Pour les dégager, considérons la signification économique des résidus aléatoires U_i. Il recouvrent l'influence de plusieurs groupes de facteurs de nature très différente. L'introduction de chacun d'entre eux peut amener à nuancer l'hypothèse théorique sur laquelle repose le modèle (4).

Ils recouvrent d'abord de simples erreurs de mesure sur la consommation ou le revenu. Ces erreurs traduisent seulement l'imperfection des méthodes d'enquête et n'affectent pas la nature du raisonnement suivi. Quelque important qu'il soit d'en examiner l'incidence quand on fait du travail concret, nous sommes fondés à les négliger ici.

Les résidus traduisent ensuite l'influence d'ensemble d'un grand nombre de variables secondaires qu'on pourrait, en théorie, introduire dans l'analyse, mais qu'il est plus expédient de considérer en bloc. Nous avons, par exemple, admis que les prix étaient les mêmes pour tous les consommateurs. Cela n'est pas rigoureusement vérifié et les différences locales de prix expliquent donc une partie des écarts des observations à la courbe Γ. En outre, le modèle (1) est lui même incomplet. Il fournit un cadre pour l'étude de l'influence des revenus et des prix "toutes choses égales d'ailleurs". Or, l'intuition suggère un certain nombre de facteurs économiques ou sociologiques, tous mesurables ou repérables, susceptibles d'influencer la consommation des ménages. C'est le cas des structures familiales, de l'âge du chef de ménage, de sa profession, de la région ou de la taille de la commune d'habitat, etc... La loi de demande déduite de la théorie des choix est incomplète ; cependant, elle indique les facteurs les plus importants ; les autres peuvent, en général, être incorporés dans le résidu U_i du modèle (3).

Une hypothèse critique a cependant été introduite pour parvenir à ce modèle : une fois éliminée l'influence des variables secondaires dont nous venons de parler, on suppose que tous les consommateurs ont la même surface de choix (17). Cette hypothèse est manifestement irréaliste ; elle est en même temps inutile. Il suffit en effet d'admettre que les variations individuelles des surfaces de choix sont peu considérables pour la majorité des consommateurs et distribuées convenablement autour d'une tendance moyenne, pour que le modèle (3) reste valable. Le résidu U_i traduit donc aussi des différences de goûts entre consommateurs. La courbe Γ n'est plus rigoureusement une loi de demande individuelle. Elle est une loi de demande "moyenne" pour la population considérée (18). Pour que cette loi puisse être identifiée à partir des observations, il suffit d'introduire quelques hypothèses assez faibles sur la façon dont sont distribuées les variations individuelles des goûts. L'hypothèse la plus importante est que ces variations sont indépendantes des revenus.

Tout ce qui précède traduit une conception déterministe du comportement des consommateurs. Le résidu U_i est une sténographie commode pour résumer l'influence de l'ensemble des variables secondaires et les hétérogénéités de comportement entre les ménages. Mais le modèle (3) reste valable sous des hypothèses beaucoup moins restrictives. De même que nous avons admis des différences de choix entre consommateurs, nous pouvons admettre des différences accidentelles d'une circonstance à l'autre dans le comportement de chaque consommateur étudié, pourvu que ces différences soient distribuées autour d'une moyenne liée de façon stable au revenu. Le point de vue selon lequel le comportement des consommateurs individuels est fondamentalement stochastique a été souvent exprimé par les spécialistes de l'étude de la consommation (19) : il est beaucoup plus réaliste que le point de vue déterministe précédemment défini.

Mais, à ce stade, que reste-t-il de la liaison avec la théorie pure des choix des consommateurs ? Très peu de choses, en somme, et certainement ce peu de choses est plus nuisible qu'utile. Le modèle (3) peut en effet reposer sur une hypothèse beaucoup plus directe et beaucoup moins restrictive : il suffit d'admettre que le revenu explique en partie la consommation, selon une même loi moyenne pour les différents consommateurs. Cette loi ne suppose pas qu'existent des surfaces de choix stables, ni que le comportement du consommateur traduise un calcul conscient entre diverses possibilités qu'il sait ordonner.

Les deux hypothèses possibles ont un trait fondamental en commun : elles font jouer un rôle privilégié au revenu des consommateurs. Cela correspond à un postulat courant de la théorie économique. La théorie pure des choix exprime cependant ce postulat sous une forme qui restreint le pouvoir explicatif du modèle introduit, en le liant à un schéma de calcul individuel conscient, sans permettre, en compensation, de dégager des hypothèses complémentaires qu'il suppose aucune restriction expérimentalement vérifiable sur la forme des courbes d'Engel. C'est qu'en fait, la théorie pure des choix et l'analyse des courbes d'Engel correspondent à deux domaines entièrement distincts de la science économique. Le second sujet est du domaine de la l'analyse empirique du comportement des consommateurs, et le premier du domaine de l'étude de l'allocation des ressources par le mécanisme des prix et de la théorie de l'optimum économique. Le lien habituellement présenté entre les deux domaines est artificiel et conduit à une vue inutilement dépourvue de réalisme du comportement des consommateurs individuels.

Si la théorie des choix restreint inutilement la validité de l'analyse, elle n'aide pas en outre à fournir les hypothèses complémentaires nécessaires. Certaines de ces hypothèses ont été énoncées chemin faisant. Elles sont habituellement admises sans examen assez critique, alors qu'elles devraient poser les problèmes non résolus les plus importants aux spécialistes : nous visons les hypothèses qui ont trait à l'indépendance en probabilité entre le terme stochastique et le revenu dans le modèle (3).

Un second groupe d'hypothèses est relatif à la forme de la fonction $f(R)$ dans ce modèle. Nous avons déjà indiqué que le choix de la forme de la fonction f était essentiel à la formulation complète du modèle économétrique considéré. Ce choix est fait le plus souvent empiriquement, en tenant compte surtout des commodités de calcul. Cependant, les principaux choix possibles peuvent être exprimés en termes d'hypothèses alternatives sur la propension marginale à consommer ou l'élasticité de la consommation du bien étudié. Exprimées en ces termes, ces hypothèses peuvent parfois être testées efficacement.

Enfin, l'interprétation, en termes de loi moyenne de comportement individuel, de la relation obtenue à partir de l'étude de plusieurs consommateurs au même instant, exige qu'on introduise deux hypothèses supplémentaires qui sont fondamentales.

La première est que le comportement des ménages étudiés ne dépend pas de facteurs dynamiques, afin que la consommation d'une période puisse légitimement être expliquée par le revenu de la même période. Cela exclut les phénomènes d'inertie qui aboutissent au maintien d'une consommation malgré des fluctuations du revenu. Ces phénomènes d'inertie peuvent avoir des causes psychologiques, mais surtout des causes techniques ou juridiques. Certaines consommations, comme les loyers ou les primes d'assurance, ont un caractère contractuel. Même si l'on admet, comme on le fait souvent dans la littérature du sujet, que les courbes d'Engel traduisent des lois de longue période (à supposer qu'une telle affirmation ait une signification précise), on doit reconnaître que l'analyse des budgets familiaux ne permet pas d'expliquer ce type de consommation. L'hypothèse exclut aussi le cas où la demande d'un bien repose sur les revenus et les prix prévus pour les périodes ultérieures. C'est justement le cas de la demande de biens durables, qui échappe donc aussi à l'analyse que nous allons décrire. Une difficulté plus fondamentale se présente quand ce n'est pas la demande d'un bien particulier, mais le niveau de vie qui est défini à partir de la prévision du futur. C'est le cas, comme l'a montré M. FRIEDMAN (20), après d'autres auteurs, des populations à revenu instable (comme les travailleurs indépendants).

La seconde hypothèse est que les choix des différents consommateurs sont indépendants. C'est à dire que la consommation d'un consommateur ne dépend pas du niveau de consommation, actuel ou passé, d'autres consommateurs, ni de sa place dans la distribution des revenus. Cela exclut à la fois les consommations ostentatoires, motivées par l'imitation d'individus ou de groupes particuliers, et les comportements altruistes, motivés par la connaissance de la pauvreté ou de la misère d'autrui.

De nombreux travaux ont conduit à mettre en doute l'indépendance des choix individuels (21). Cela ne met cependant pas concrètement en cause l'utilité de l'analyse des courbes d'Engel, mais en limite le domaine d'application. C'est en effet principalement dans le domaine des achats de biens durables et de certaines consommations de loisirs (comme les vacances d'hiver) que ces interdépendances et ces phénomènes d'imitation sont les plus marqués.

SECTION 5

LES BASES THEORIQUES DE L'ANALYSE DE LA DEMANDE D'UN BIEN DURABLE DE CONSOMMATION

L'étude de la demande d'un bien de consommation courante repose sur une hypothèse économique bien définie, même si elle est plus faible et moins axiomatique que celle qu'on déduit souvent de la théorie pure des choix.

L'exemple de l'étude de la demande de biens durables montre facilement qu'une telle hypothèse est nécessaire. Considérons le marché français de l'automobile : une enquête par sondage, faite il y a quelques années, a permis de connaître le nombre et les revenus, d'une part, des ménages qui ont acheté une automobile neuve en 1956, d'autre part, des ménages qui possédaient une automobile en 1956. Les deux variables : parc automobile et achats d'automobiles dans l'année sont en bonne corrélation avec le revenu comme le montrent les figures 2 et 3. Dans chaque cas, les ménages ont été triés dans l'ordre des revenus croissants. La figure 2 indique, en fonction du revenu, le pourcentage des ménages qui ont acheté une automobile neuve en 1956 : la figure 3, le pourcentage des possesseurs d'automobiles en 1956, quelle que soit l'année d'achat.

L'examen empirique des données permet donc de proposer deux explications différentes. Le revenu des ménages peut soit expliquer directement les ventes de voitures automobiles, soit le parc possédé (ou désiré) par les ménages. Selon la seconde explication, les ventes d'automobiles dans l'année, au lieu de dépendre directement du revenu, en dépendent indirectement : les ventes sont la somme de deux termes. Le premier est la variation du parc désiré, qu'on peut calculer à partir des modifications de la distribution des revenus d'une année à l'autre ; le second exprime la demande de remplacement des véhicules hors d'usage (22).

Les deux explications conduisent à des résultats contradictoires pour prévoir l'évolution des ventes d'automobiles neuves sur la même période et l'étude empirique des données ne permet en aucune façon de choisir entre elles. Cela illustre le fait qu'il n'existe pas d'économétrie sans modèle du phénomène considéré.

FIGURE 2

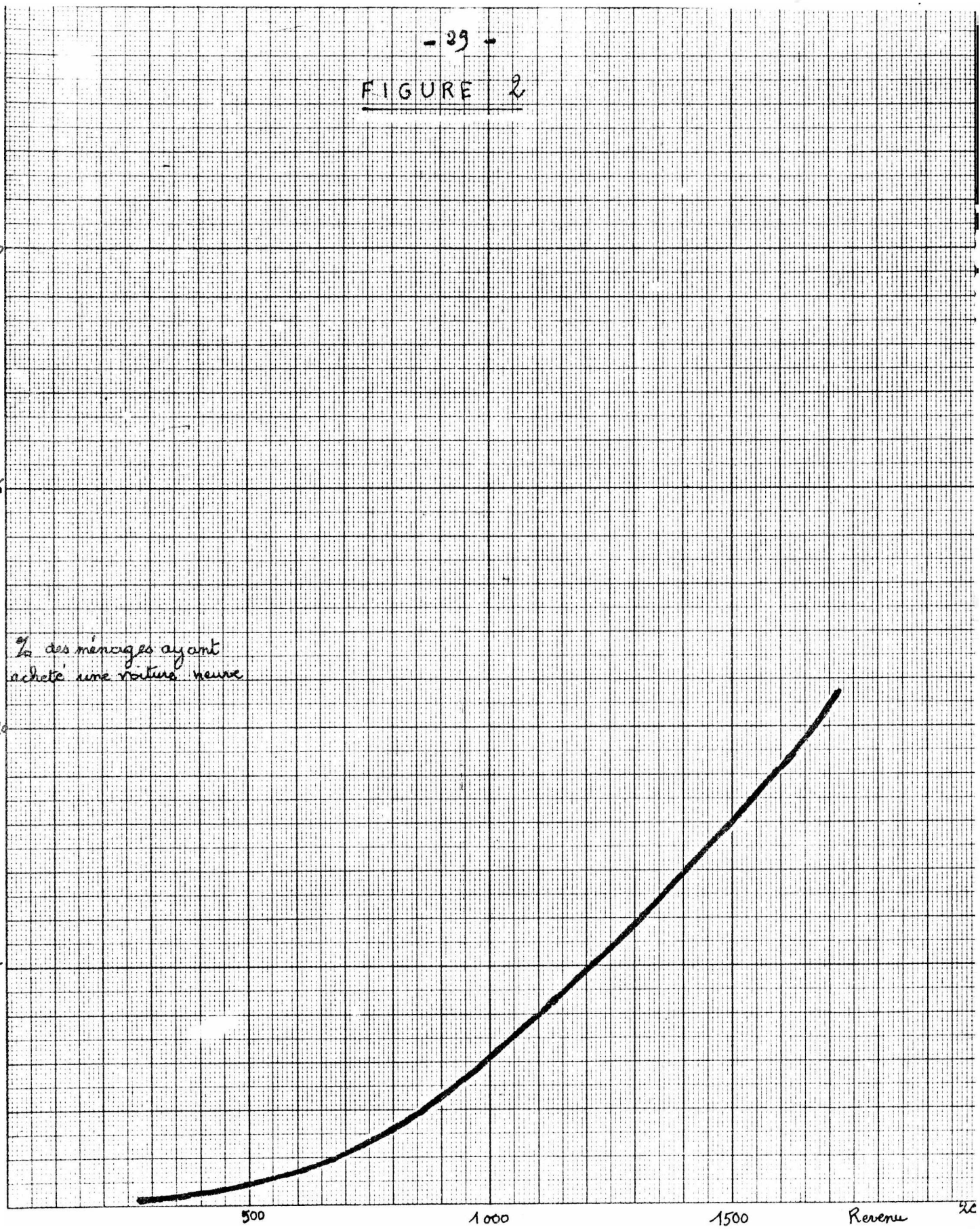
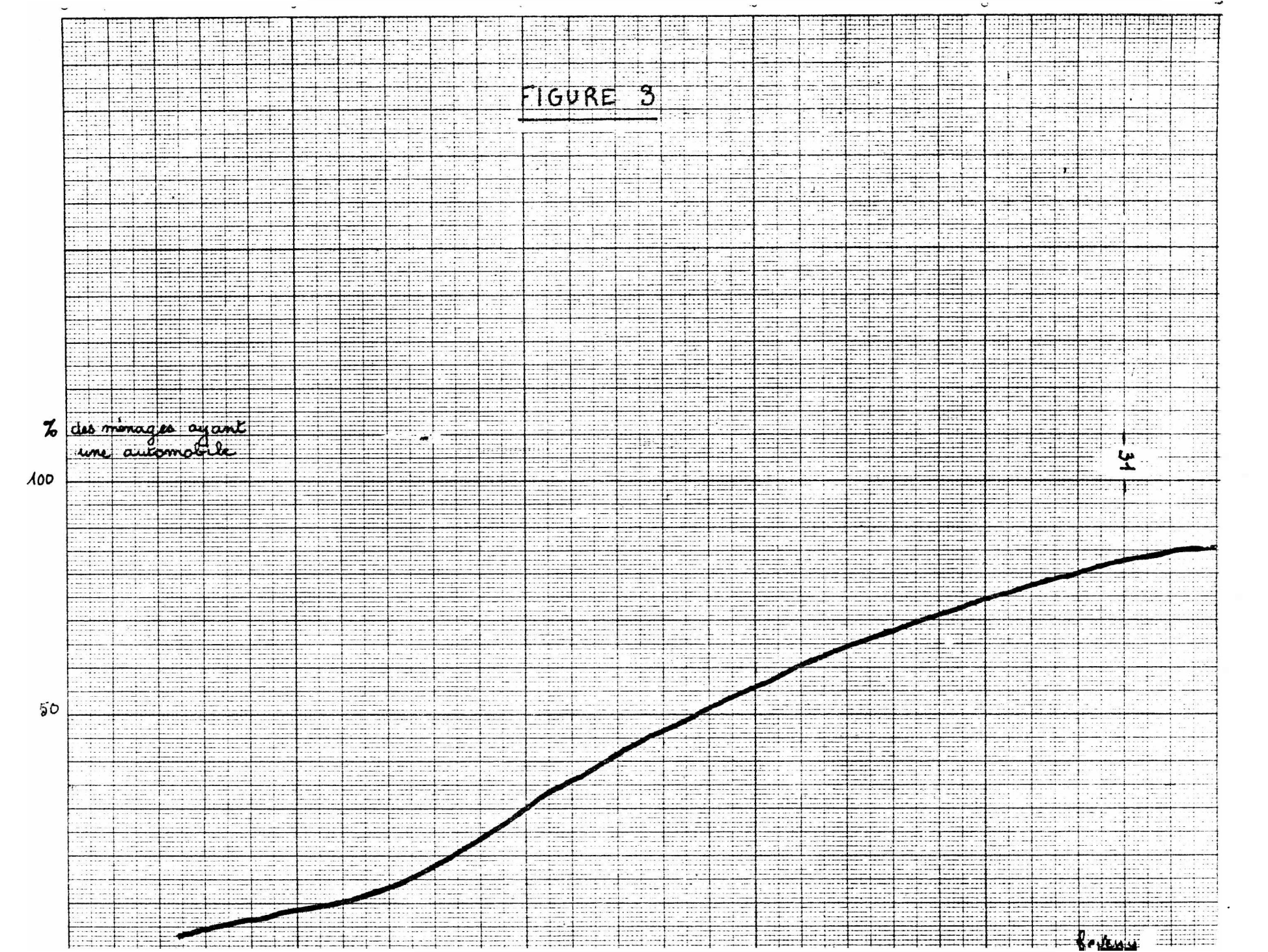

% des ménages ayant
achetés une voiture neuve
500
1 000
1500
Revenu

FIGURE 3



De façon très générale, la seconde explication est seule admise. On peut la justifier selon deux lignes de raisonnement possibles. La première consiste à considérer que la "consommation" qu'il convient d'expliquer par le revenu courant est celle d'un service de transports, fourni par un équipement, la voiture individuelle, qui procurera des services à son propriétaire pendant une succession de périodes.

Le choix entre satisfaire ce besoin de transport par la possession d'un véhicule individuel ou par d'autres méthodes dépend, entre autres, du volume de transport qu'on désire consommer, qui dépend lui-même du revenu courant et des revenus anticipés des périodes ultérieures. L'horizon économique des consommateurs étant assez proche, on peut admettre, sous certaines hypothèses peu restrictives, que le revenu courant soit un bon indicateur de la valeur actualisée du flux des revenus anticipés.

Cette première ligne de raisonnement suppose cependant, implicitement, un calcul économique beaucoup trop conscient de la part des consommateurs. Elle s'appliquerait plutôt à une entreprise qui chercherait à définir une politique optimale en matière de transports, le choix lui étant offert entre posséder son parc de véhicules ou utiliser les services d'entreprises de transport. Dans le cas des consommateurs, une hypothèse plus simple et plus réaliste est d'admettre que la possession d'un bien durable est recherchée pour elle-même, sans considération directe du volume de services qu'on en retire. "La caractéristique d'un bien durable est que l'utilité découle de sa possession plutôt que de son utilisation" (23). Cette seconde ligne conduit à exprimer le raisonnement en termes d'une structure préférée du compte de patrimoine du ménage, plutôt que d'une structure préférée de son compte d'affectation pour l'année en cours, adoptant ainsi l'optique suggérée par BOULDING (24). Le facteur explicatif principal du parc automobile devient alors le capital des ménages. Il s'agit cependant d'une variable qui n'est pas susceptible de mesure expérimentale directe, puisque ce capital est la somme de la fortune au moment de l'observation et de la valeur actualisée de l'échéancier des revenus attendus à partir d'autres sources que cette fortune (et, en particulier, des revenus attendus du travail futur des membres du ménage). Moyennant l'hypothèse complémentaire que nous avons suggérée plus haut, le revenu courant peut être pris pour un indicateur de la valeur actuelle de l'échéancier des revenus attendus.

Nous sommes donc conduits à expliquer l'évolution du parc automobile par les modifications de la structure des revenus et des patrimoines des ménages (25). Afin de préciser cette hypothèse et de l'exprimer en termes de comportements individuels, sans alourdir l'exposé, nous nous bornerons à l'influence du revenu courant ; l'influence du patrimoine s'étudie d'une façon identique.

Dans le cas d'un bien de consommation courante, supposé parfaitement divisible, la variable qu'on cherchait à expliquer était le niveau de consommation qui pouvait varier de façon continue. Dans le cas présent, la situation est plus simple, puisqu'elle s'exprime en termes de tout ou rien. Il s'agit de déterminer le revenu à partir duquel un ménage possédera une automobile. L'hypothèse la plus simple serait d'admettre que ce revenu est le même pour tous les ménages (26). Elle est cependant inacceptable, puisque la courbe de la figure 3 montre qu'on observe des propriétaires de voitures à tous les niveaux de revenus, mais dans une proportion très variable avec le revenu.

Aussi convient-il d'admettre qu'il existe pour chaque ménage un seuil S_i d'accès à la motorisation, qui dépend des préférences du ménage. Pour l'ensemble de la population, le seuil S_i est donc une variable aléatoire dont la distribution doit être étudiée conjointement avec la distribution des revenus R_i des ménages.

Un ménage donné est caractérisé par son revenu R_i et son seuil S_i. Il possédera une voiture si :

$$S_i \leq R_i$$

Il n'en possédera pas si :

$$R_i < S_i$$

L'étude du problème exige l'introduction d'une hypothèse complémentaire : les deux variables aléatoires R et S doivent être indépendantes c'est à dire que le seuil d'accès à la motorisation S_i pour un ménage donné traduit ses préférences mais ne dépend pas de son revenu. Dans ces conditions, la courbe de la figure 3 représente approximativement la courbe cumulée de la distribution de S_i, à condition qu'on dispose d'un assez grand nombre d'observations dans chaque classe de revenus.

En effet, considérons n ménages de revenus R, Puisque S et R sont des variables aléatoires indépendantes, ces n ménages sont un échantillon aléatoire dans la population des S_i. Soit :

$$p = f \quad (S)$$

la courbe cumulée de la distribution des S, p étant la probilité pour que $S_i \leqslant S$. Aux erreurs d'échantillonnage près, pn ménages auront un seuil d'accès à la motorisation inférieur ou égal à R, et n (1 - p) supérieur à R. C'est ce que mesure la courbe 3. Pour achever l'analyse, on admet en général que la distribution des seuils S_i est logarithmico-normale, ce qui permet de transposer les résultats d'une méthode d'analyse classique en expérimentation biologique (27).

L'influence des modifications des revenus sur le parc automobile des ménages s'obtient donc en combinant la distribution des revenus des ménages et la distribution, définie précédemment, des seuils S_i. Cela n'épuise pas l'analyse. Dans le cas français, de 1949 à 1958, le parc automobile des ménages est passé de 1,4 à 3,6 millions de véhicules, alors que l'effet des modifications des revenus et de l'augmentation de la population expliquent le passage de 1,4 à 2,5 millions seulement (28). Une tendance explique donc une part importante de l'augmentation du parc automobile. Cela n'est pas satisfaisant, puisqu'une tendance n'est rien d'autre qu'un aveu d'ignorance et un appel à chercher d'autres explications. Dans le cas présent, plusieurs possibilités se présentent.

La courbe de la figure 3 relie le parc effectif aux revenus. Or, notre modèle théorique se référait au parc <u>désiré</u> par les ménages. En 1956, les deux notions étaient voisines. Par contre, le parc effectif de 1949 était très inférieur au parc désiré. La plupart des voitures neuves étaient destinées aux entreprises et les délais de livraison aux ménages étaient très longs pour tous les modèles. Une partie notable de l'écart correspond au retour à une situation normale sur le marché.

Ensuite, nous n'avons tenu compte que d'une partie de notre hypothèse, puisque nous n'avons pas considéré l'effet des modifications de la structure des patrimoines. Enfin, il peut rester une tendance inéliminable, qui correspond à une modification lente de la forme de la courbe 3, c'est à dire à un changement de l'attitude des consommateurs en face de l'automobile.

Cela nous ramène à notre propos de départ. Le but de l'analyse économique de la consommation n'est pas de fournir une explication totale. En particulier, elle n'a pas à expliquer la formation des préférences des consommateurs. Elle doit seulement étudier la façon dont la variation d'un petit nombre de facteurs socio-économiques module l'expression de cette préférence sur le marché. La tâche est déjà considérable, puisqu'il n'existe pas une théorie d'ensemble pour représenter le comportement économique des ménages, mais une gamme de modèles différents, dont nous avons pu seulement évoquer rapidement deux exemples.

Il est possible que ces modèles différents apparaissent en fait à l'avenir comme des facettes d'une théorie unique plus générale. Cela est indifférent à qui à la tâche d'étudier empiriquement la consommation, puisqu'aucun matériau concret ne lui permet de pressentir les lignes qu'aurait cette théorie plus générale et qu'il se trouve déjà désarmé lorsqu'il veut exprimer ses hypothèses partielles en des termes qui soient à la fois précis, opératoires, et qui soient autre chose que des intuitions peu élaborées.

N O T E S

(1) - Ce texte s'inspire d'un article antérieur : G. ROTTIER: L'Economie de la Consommation - Revue Economique N° 2 1961.

(2) - "Economie et Consommation" - Revue de l'Action Populaire, Juillet-Août 1960 p. 771.

(3) - E.A. LISLE : La prévision de la Consommation. "Revue Economique" - N° 2 1961.

(4) - Ruth P. MACK : Economics of Consumption. ap. "A Survey of Contemporary Economics", vol. II, B.F. HALEY editor. Irwin & C°, 1952, p. 39.

(5) - E. MALINVAUD : Initiation à la Comptabilité Nationale, Paris 1957, p. 17.

(6) - J.K. GALBRAITH : The Affluent Society, Houghton Mifflin, 1958. p. 308.

(7) - F.L. CLOSON : "Avant-Propos" : Consommation, N° 1 - 1958, p. 3.

(8) - Cf. G. ROTTIER : La prévision de la Consommation, Revue de l'Action Populaire, Juillet-Août 1960, particulièrement pp. 807-808.

(9) - En nous limitant à la France, nous devons au moins citer à ce propos, d'une part les travaux de la Section de Nutrition de l'Institut National d'Hygiène, dirigés par le Dr. TREMOLIERES, d'autre part les apports du Groupe d'Ethnologie Sociale du C.N.R.S., dirigé par P.H. CHOMBART de LAUWE (cf. en particulier "Conditions de vie des familles ouvrières", Paris - 1958).

(10) - Cf. HOUTHAKKER : An International Comparison of Household Expenditure Patterna, Commenorating the centenary of ENGEL'S Law - Econometrica, 1957.

(11) - T.J. KOOPMANS : Three Essays on the State of Economic Science, Mc Graw HILL, 1957, p. 206.

(12) - Dans l'étude de la consommation, le terme "individuel" signifie, en fait, qu'on se réfère au ménage. Le ménage défini comme un groupe de personnes vivant ensemble et mettant leurs ressources en commun, est considéré comme l'unité de décision. Il s'agit là d'une abstraction nécessaire pour l'analyse. La mise en commun des ressources n'est jamais entière et leur affectation à telle ou telle utilisation est rarement le résultat d'une décision commune . Savoir qui décide en fait telle dépense, parmi les membres du ménage est un champ d'étude intéressant pour le sociologue.

(13) - On ne peut cependant pas aller très loin. En effet, les ajustements effectués sur des données individuelles sont toujours beaucoup moins précis que ceux effectués sur des moyennes. Il faut donc notablement plus d'observations pour obtenir une estimation du même paramètre avec la même vraisemblance. Les faibles corrélations observées sur les données individuelles sont d'ailleurs instructives en soit, puisqu'elles montrent le caractère trop schématique des relations établies sur des moyennes.

(14) - J. ALBERT - G. ROTTIER - N. TABARD : La demande de logements en France. Annales de Recherches et de Documentation sur la Consommation - N° II-1957.

(15) - La première enquête française notable dans ce domaine a été faite en 1959. Ses résultats ont été publiés dans "Consommation" N° 4 (Octobre-Décembre) 1960.

(16) - ALLEN, R.G.D. et BOWLEY, A.L. : Family Expenditure, Londres 1935.

(17) - Cette hypothèse est explicitement introduite par WOLD, H. et JUREEN, L : Demand Analysis - New-York, 1953, p. 8.

(18) - Cette loi de demande moyenne ajustée sur des données individuelles est bien entendu entièrement distincte d'une loi de demande sur le marché qui relie la consommation totale d'un produit dans une population à son revenu global. Le passage d'une notion à l'autre introduit des difficultés d'agrégation.

(19) - Cf. Ruth MACK, op. cit. et J.S. CRAMER : A statistical model of Major Consumer Durables ; Thèse Amsterdam, non publiée, 1961, Chapitre I, paragraphe 2.

(20) - M. FRIEDMAN : A Theory of the Consumption Function, 1957.

(21) - J.S. DUESENBERRY : Income, saving and the Theory of Consumer Behavior - Harvard, 1949.

- D. BRADY and M.FRIEDMAN : Savings and the Income Distribution ap. Studies in Income and Wealth, vol. 10, 1947.

(22) - Cf. H. FAURE : Un modèle prospectif du marché de l'automobile - Consommation, N° 4, 1959.

(23) - M.J. FARRELL : The Demand for motor cars in the United States. Royal Statistical Society, 117 A II, p. 171.

(24) - K. BOULDING : A Reconstruction of Economics, 1950.

(25) - Cette hypothèse est testée et quantifiée par J.S. CRAMER, op. cit.

(26) - Cette méthode simplifiée a été utilisée par :

Savini : un modello per prevision di circolazione automotociclista in Italia - Industria, 1957,

et Melle J. MORICE : La demande d'automobiles en France, Paris, 1957.

(27) - Cf. J. AITCHISON and J.A.C. BROWN : The lognormal Distribution, Cambridge, 1957, Chapitre 7.

(28) - Cf. H. FAURE, op. cit.

C H A P I T R E II

L'OBSERVATION DE LA CONSOMMATION

SECTION 1

LE ROLE DES ENQUETES PAR SONDAGE DANS L'ETUDE DE LA CONSOMMATION

Les enquêtes par sondage, dont les techniques se sont considérablement développées et perfectionnées dans les dernières décennies, représentent une méthode d'observation statistique dont les applications s'étendent à des domaines très variés : agronomie, technologie de l'entreprise, sociologie et psycho-sociologie, etc... Pour un ensemble de raisons complémentaires, l'étude économique du consommateur représente cependant leur terrain d'élection. Les ménages forment en effet une population statistique nombreuse dont les éléments sont assez homogènes pour que ces techniques s'appliquent avec succès à leur étude. En même temps, aucune autre source statistique ne permet de les observer commodément. L'étude exhaustive, par recensement, est une opération extrèmement coûteuse qui ne peut être engagée qu'à de longs intervalles et qui ne permet d'obtenir que des résultats, certes indispensables (ne serait-ce que fournir une base à des sondages ultérieurs), mais peu nombreux, simples et de nature plus sociologique et démographique qu'économique. Les statistiques d'origine administrative, pour leur part, touchent rarement les ménages. Lorsqu'elles le font, ce qui est le cas de certaines statistiques fiscales, elles fournissent des informations limitées dont la qualité est difficilement contrôlable.

L'enquête par sondage, dont le coût est faible par rapport à celui de la collecte des autres renseignements, offre l'avantage d'une grande souplesse. Le statisticien est relativement maître de la définition des concepts qu'il introduit. Le choix des variables étudiées peut être déterminé de façon à permettre la vérification quasi-expérimentale de telle ou telle hypothèse théorique posée a priori. Moyennant le recours à des enquêteurs très expérimentés, des informations correctes peuvent être obtenues sur des sujets apparemment très confidentiels ou indiscrets, comme la structure des patrimoines, les habitudes d'épargne ou les dépenses médicales. Enfin, comme la collecte des renseignements

convenablement formés et contrôlés, permet de réduire considérablement les erreurs de mesure et d'observation, au prix d'une erreur d'échantillonnage qui est le plus souvent négligeable quand on ne vise pas à étudier des sous-populations trop restreintes.

Malgré ces avantages, l'exécution correcte des enquêtes par sondage est une opération délicate qui correspond à une technique particulière très spécialisée dont la pratique représente un véritable métier. La théorie et la pratique des sondages forment le sujet d'une littérature très étendue dont il ne nous appartient pas de rendre compte ici (1) ; nous nous bornerons à rappeler quelques éléments indispensables pour la description de l'enquête sur les budgets familiaux, faite en 1956 par le C.R.E.D.O.C. et l'I.N.S.E.E., dont l'analyse a inspiré cet essai (2).

Cette enquête visait à étudier toutes les dépenses de consommation des ménages ordinaires résidant en France métropolitaine et à étudier l'autoconsommation (en particulier de produits alimentaires) quand elle était susceptible de se présenter. L'objectif de l'enquête était double. D'une part, mesurer la consommation moyenne des différents produits, pour l'ensemble de la population aussi bien que pour des sous-populations particulières ; l'enquête devait ainsi fournir des matériaux pour une meilleure connaissance de postes importants de la Comptabilité nationale française. D'autre part, l'enquête visait à relier la consommation des différents produits à des variables propres à chaque ménage et susceptibles de contribuer à l'expliquer : revenu, nombre de personnes composant le ménage, région ou taille de la commune d'habitation, catégorie socio-professionnelle du chef de ménage, etc...

Le second point de vue nous intéressera seul ici. Mais il importe de souligner que l'enquête s'intéressait également au premier. En effet, les deux types d'études, comme nous le verrons, auraient justifié des méthodes d'enquête et des plans d'échantillonnage tout à fait différents. En fait, la méthode adoptée représente un compromis, la balance ayant cependant penché plus en faveur du premier point de vue que du second (3).

Avant de décrire l'enquête elle-même (section 3), nous devons rappeler quelques notions simples sur la pratique des sondages, dont l'énoncé facilitera notre exposé.

SECTION 2

NOTIONS DE THEORIE DES SONDAGES

§ 1 - La notion de sondage probabiliste

Contrairement aux études de marché, où des méthodes empiriques moins coûteuses sont fréquemment utilisées, les grandes enquêtes nationales se font toujours à partir d'un sondage probabiliste de ménages. On définit ainsi une méthode de sondage telle que chaque individu de la population étudiée ait une probabilité connue et non nulle d'appartenir à l'échantillon. Dans la plupart des applications, les tirages sont faits avec une fraction de sondage constante. Dans ce cas, on peut considérer que chaque ménage de la population est représenté par une boule dans une urne ; le sondage consiste à extraire un certain nombre de ces boules.

Il existe diverses possibilités pratiques pour réaliser le tirage sans trop de frais, en s'assurant que la probabilité de sortie de chaque élément de la population soit bien la même. Ces procédés ne nous intéressent pas ici. Par contre, il est clair que les sondages probabilistes exigent qu'on dispose d'une liste complète, sans ommission ni répétition, de la population étudiée. Or il n'existe pas de fichier des ménages. On atteint donc, en pratique, les ménages de façon indirecte, en effectuant soit un tirage d'individus, soit un tirage de logements.

Un tirage d'individus peut se faire à partir des listes électorales, qui sont publiques. Leur utilisation est cependant un pis aller. En effet, outre qu'elles sont imparfaitement mises à jour, elles ne permettent pas de tirer correctement des ménages à partir d'un tirage d'individus, puisque le nombre d'électeurs par ménage est variable.

En pratique, on remplace donc le sondage de ménages par un sondage de logements. Trois fichiers peuvent être utilisés : les fichiers d'abonnés à l'E.D.F., les fichiers d'assujettis à la cote mobilière et les feuilles de logements du recensement de la population. Seule cette dernière ressource est utilisée dans les enquêtes faites par les organismes publics. Son utilisation entraine deux conséquences. La première est qu'on est lié par la définition du ménage adoptée au recensement, qui est plus extensive que celle qui conviendrait pour une analyse proprement économique (cf. infra, p. 57) (4). La seconde est que la base de sondage vieillit, puisque certains logements recensés ont pu être détruits ou affectés à d'autres utilisations que l'habitation depuis le recensement, cependant que des logements neufs ont été construits. Les changements d'affectation ou les destructions ne posent de problème sérieux que si leur importance relative varie beaucoup d'une localité à l'autre : l'omission des logements neufs pourrait entraîner par contre des erreurs importantes. Aussi complète-t-on maintenant la base de sondage à partir des statistiques de permis de construire.

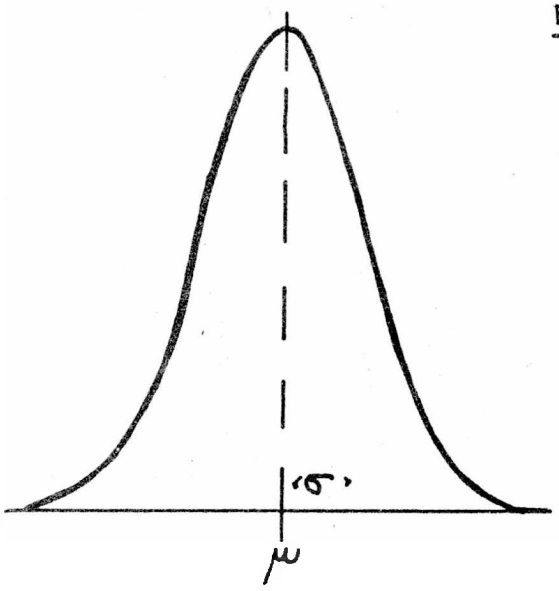
Les imperfections liées à l'ancienneté de la base de sondage, ainsi que le refus de répondre à l'enquête de la part de certains ménages font que l'échantillon de ménages ayant répondu n'est pas identique à l'échantillon tiré qui était lui, vraiment aléatoire. Des corrections sont nécessaires pour éviter qu'il en résulte une erreur systématique dans certaines estimations : nous indiquerons comment ce travail a été effectivement réalisé en 1956.

§ 2 - L'étude d'une moyenne dans un sondage simple

Afin de justifier certains aspects, apparemment complexes, du tirage de l'échantillon, nous commencerons par rappeler comment se présente l'estimation de la moyenne d'un caractère quantitatif à partir d'un sondage probabiliste simple. Nous supposons que nous avons une population mère dans laquelle le caractère X considéré est distribué autour d'une moyenne m avec un écart type σ (Distribution I, figure II-1). Nous tirons un échantillon de n observations, en admettant que ce tirage soit fait avec remise dans l'urne, c'est à dire d'une façon telle que les n observations soient indépendantes en probabilité. Soit II la distribution de x dans cet échantillon ; cette distribution est centrée autour d'une moyenne $\bar{x}$ avec un écart type.

Si nous considérons tous les échantillons possibles de n observations, la moyenne $\bar{x}$ de ces échantillons est une variable aléatoire dont la distribution (distribution III, figure II-1) possède les propriétés suivantes :

FIGURE - II - 1



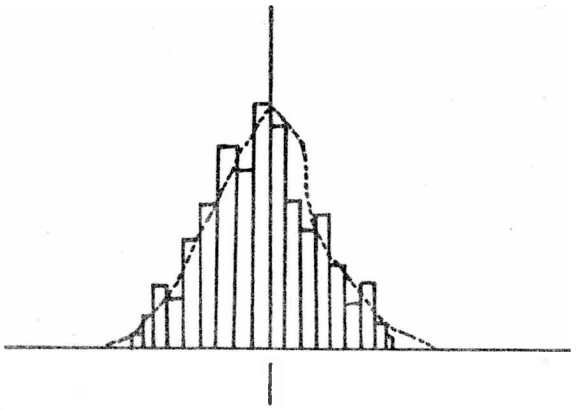
Distribution I

Population mère, infinie

Connu : 1) La distribution est normale

2) Son écart type σ

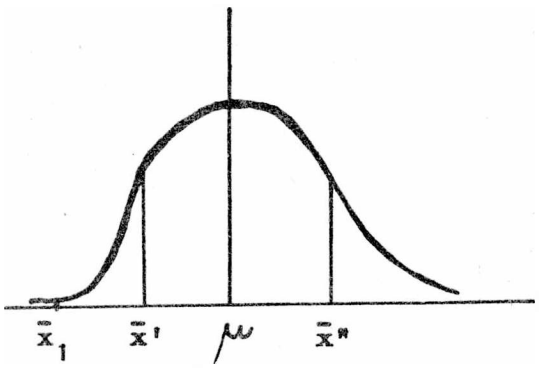
Inconnu : La moyenne μ qu'on cherche à estimer



Distribution II

Distribution des valeurs de x observées dans l'échantillon.

- Complètement connue



Distribution III

Distribution des valeurs de $\bar{x}$ pour l'ensemble, infini, de tous les échantillons différents de n observations tirés de I .

Inconnue - Mais le calcul permet d'en définir certaines propriétés :

moyenne μ

écart type $\frac{\sigma}{\sqrt{n}}$

etc...

a. La distribution III, sous des hypothèses très peu restrictives, tend vers une distribution normale quand n augmente. Elle peut pratiquement être considérée comme normale dès que $n > 30$.

b. La distribution III a une moyenne égale à μ.
$E(\bar{x}) = \mu$. La moyenne de l'échantillon est un estimateur sans biais de μ.

c. La distribution III a un écart type égal à $\frac{\sigma}{\sqrt{n}}$, qu'on peut estimer par $s\sqrt{\frac{n}{n-1}}$

d. Aucun estimateur de la moyenne μ n'a de variance plus faible que $\bar{x}$ dans le cas d'un tirage simple (5).

La condition soulignée dans la proposition d. est fondamentale. C'est elle qui justifie les développements modernes de la théorie des sondages. Si la moyenne de l'échantillon est en effet un estimateur efficient de la moyenne de la population dans un sondage simple, la précision de l'estimation peut être encore augmentée en adoptant un plan d'échantillonnage plus complexe. Un problème de sondage est fondamentalement un problème de calcul économique : il s'agit de déterminer le plan de sondage qui donne la précision maximale pour un coût d'enquête donné (ou, ce qui revient au même, qui permet une précision donnée pour le coût le plus faible) (6).

Deux procédés principaux accroissent l'efficacité du sondage : la stratification et le tirage à plusieurs degrés.

§ 3 - La stratification de l'échantillon

Avant d'effectuer une enquête, on possède habituellement un certain nombre d'informations, au moins qualitatives sur le phénomène étudié. On sait, par exemple, que les habitudes de consommation peuvent varier avec la région et l'urbanisation. Il est utile de tenir compte de ce genre d'informations dans le choix de l'échantillon, car cela peut accroître considérablement l'efficience de l'enquête.

Si la population étudiée peut être analysée en sous-populations plus homogènes du point de vue du phénomène étudié, on gagnera de la précision en tirant séparément l'échantillon qui correspond à ces sous-populations, dites "strates". La probabilité de tirage de chaque individu de la population reste bien connue et non nulle, mais la probabilité de tirage de certains ensembles d'individus qui formeraient des échantillons aberrants est rendue nulle.

Dans l'exemple que nous avons cité, la stratification empêche l'apparition d'échantillons comprenant un trop fort pourcentage de ruraux ou d'habitants d'une région particulière, qu'un sort défavorable pourrait faire apparaître dans un sondage simple.

La stratification peut se faire selon des critères qualitatifs, comme la division de la France en régions. Quand on connaît, à partir de la base de sondage, la valeur, pour tous les individus, d'une variable qu'on pense être en corrélation avec la variable étudiée, on peut aussi stratifier l'échantillon selon les valeurs de la variable connue, dite "variable de contrôle". C'est, par exemple, le cas d'une stratification suivant la taille des communes de résidence des ménages.

Il est fréquent de stratifier l'échantillon suivant plusieurs critères à la fois. Nous verrons plus bas que l'échantillon de logements tiré pour l'enquête de 1956 a été stratifié à la fois suivant la taille des communes et la région.

Le calcul de l'erreur d'échantillonnage peut être facilement étendu au cas d'un sondage stratifié. Nous utiliserons les notations suivantes :

Population

N			individus en tout
N_i			individus dans la strate i
$\bar{X}$	et	σ_X	moyenne et variance de X dans la population
$\bar{X}_i$	et	σ_{X_i}	moyenne et écart type de X dans la strate i

Echantillon

n	individus tirés
n_i	individus tirés dans la strate i
$\bar{x}_i$ et s_{x_i}	moyenne et écart type de x dans le sous-échantillon i.

On cherche l'estimateur optimal de $\bar{X}$. Les formules du paragraphe 2 sont valables à l'intérieur de chaque strate. Dont $\bar{x}_i$ est un estimateur non biaisé de $\bar{x}_i$; sa variance est :

$$\frac{s^2_{x_i}}{n_i}$$

$\bar{x} = \sum_i \frac{N_i}{N} \bar{x}_i$ est un estimateur non biaisé de $\bar{X} = \sum \frac{N_i}{N} X_i$

$$E(\bar{x}) = E\left\{\sum \frac{N_i}{N} \bar{x}_i\right\} = \sum \frac{N_i}{N} E\, x_i = \sum \frac{N_i}{N} X_i$$

L'estimateur étant une somme de variables aléatoires indépendantes, sa variance s'estime sans difficulté

$$V(\bar{x}) = \sum_i \frac{N_i^2}{N^2} V(\bar{x}_i)$$

$$\text{de } V(x_i) = \frac{\sigma^2 X_i}{n_i} \text{ (d'après a)}$$

$$V(\bar{x}) = \sum \frac{N_i^2}{N^2} \frac{\sigma^2 X_i}{n_i}$$

qui est toujours inférieur ou égal à $\frac{\sigma_X^2}{n}$

La répartition de l'échantillon entre les strates

Si on prend la même fraction de sondage dans chaque strate ($\frac{n_i}{N_i}$ = cte) on a un échantillon dit "représentatif". C'est une technique classique, car elle facilite considérablement l'exploitation de l'enquête, l'échantillon pouvant se dépouiller comme un recensement. Si on a des informations sur l'ordre de grandeur des variances $\sigma^2_{x_i}$ dans chaque strate, on peut cependant gagner à diviser l'échantillon différemment. NEYMAN a étudié la répartition optimale d'un échantillon d'effectif n donné entre p strates, c'est-à-dire celle qui minimise $v(\bar{x})$:

Le problème est de trouver le minimum de :

$$\sum_i \frac{N_i^2}{N^2} \frac{\sigma^2 x_i}{n_i}$$

Les variables n_i étant soumises à la condition $\sum n_i = n$

En introduisant un multiplicateur de Lagrange, cela revient à annuler :

$$\frac{\partial}{\partial n_i} \left[\sum_i \frac{N_i^2}{N^2} \frac{\sigma^2 x_i}{n_i} - \lambda \left(\sum n_i - n \right) \right]$$

On obtient immédiatement :

$$\frac{N_i^2 \sigma^2 X_i}{n_i^2} = \text{cte}$$

Les fractions de sondage doivent être proportionnelles à l'écart type de x dans chaque strate. Cet écart type est inconnu, mais on peut le remplacer par une estimation grossièrement approchée : l'écart type, dans chaque strate, de la variable de contrôle utilisée pour définir les strates.

Un cas particulier important se présente souvent : Si on a des raisons de penser que l'écart type et la moyenne de X sont dans un rapport fixe dans chaque strate, c'est à dire que les distributions sont homothétiques (dans le cas où elles sont normales), la condition NEYMAN :

$$\frac{n_i}{N_i \sigma_{X_i}} = \text{Cte} \quad \text{s'écrit} \quad \frac{n_i}{N_i \bar{X}_i} = \text{Cte}$$

puisque

$$\frac{\sigma_{X_i}}{\bar{X}_i} = \mu$$

Les n_i doivent être proportionnels à la valeur cumulée des X_i dans chaque strate. Si l'on fait, par exemple, une enquête auprès de détaillants pour connaître le stock d'un produit, et si l'on admet que le stock d'un détaillant est à peu près proportionnel à ses achats annuels du produit, on stratifiera la population des détaillants selon le montant de leurs achats et on tirera dans chaque strate des effectifs proportionnels aux achats de l'ensemble des détaillants de la strate dans l'année écoulée.

Un exemple d'enquête auprès d'entreprises, cité par HANSEN et HURWITZ (7), montre l'avantage d'un échantillon optimal :

	Effectifs dans la population	Echantillon non stratifié	Echantillon représentatif	Echantillon Neyman	Echantillon avec $\frac{n_i}{N_i \bar{X}_i} = \text{cte}$
Strates 1	538		15	244	286
et ef- 2	4.756		131	288	409
fectifs 3	30.964		854	468	305
Total	36.258	1.000	1.000	1.000	1.000
$\frac{\sigma_{\bar{X}}}{\bar{X}}$		9,9 %	7,1 %	3,0 %	3,3 %

On voit dans cet exemple que le choix d'un échantillon proche de l'optimum apporte un gain plus considérable que la simple stratification.

Dans la pratique, signalons qu'afin de simplifier l'exploitation, on prend des fractions de sondage en rapports simples les unes avec les autres, tout au moins chaque fois que l'enquête est importante. Dans l'exemple ci-dessous, on prendrait par exemple des fractions de sondages égales respectivement à $\frac{1}{2}$ $\frac{1}{10}$ et $\frac{1}{100}$.

§ 4 - Les sondages à plusieurs degrés

Au lieu de tirer directement les individus, on peut d'abord tirer des groupes d'individus, puis en deuxième stade, tirer des individus à l'intérieur de chaque groupe. Ex : dans un tirage de logements en France, on peut d'abord tirer des communes, puis tirer des logements à l'intérieur de chacune de ces communes. Bien entendu, la probabilité de tirage de tous les logements français doit rester connue et non nulle. Dans le cas où cette probabilité est la même pour toutes les strates, on peut par exemple tirer les communes avec une probabilité proportionnelle au nombre de logements qu'elles contiennent, le même nombre de logements étant ensuite tiré dans chaque commune.

Les communes sont dites "unités primaires" et les logements "unités secondaires".

L'avantage de ce procédé est de concentrer l'échantillon sur un moins grand nombre de communes qu'un tirage direct de logements et donc de réduire le coût du travail sur le terrain.

Bien entendu on peut faire des tirages à plus de deux degrés (ex : tirage de cantons ; tirage de communes dans ces cantons ; tirage de logements dans ces communes).

Contrairement à la stratification qui accroît la précision, le tirage à plusieurs degrés a une précision inférieure, pour une taille d'échantillon donné, au tirage direct. Mais le coût de l'enquête diminue plus vite que la précision. On obtient donc un résultat final plus précis pour le même coût.

§ 5 - Sondages et plans expérimentaux

La combinaison des procédés que nous avons rapidement décrits permet, au prix d'une notable complication du plan de sondage, de réduire considérablement l'erreur d'échantillonnage sur l'estimation d'une moyenne (8). Cependant, trois remarques conduisent à nuancer légèrement les conclusions qu'on pourrait déduire des paragraphes précédents.

En matière de sondages auprès de consommateurs, on ne peut jamais, d'abord, pousser la stratification aussi loin qu'un raisonnement théorique le conseillerait. Selon le raisonnement de NEYMAN, il serait souhaitable de stratifier l'échantillon selon la variance probable des consommations dans les différentes strates. Cela est impossible pour deux raisons. La première est que le résultat auquel on parviendrait ne serait vraisemblablement pas le même pour toutes les consommations. Si la notion d'un plan de sondage optimal est claire quand on mesure un seul caractère (comme le chiffre d'affaires d'entreprises, l'emploi par branche ou la production de blé dans des exploitations agricoles), il semble difficile de lui attribuer une signification non ambigüe dans le cas d'une enquête de budgets familiaux. La seconde raison est qu'on ne peut établir de stratification qu'à partir de renseignements que fournit le fichier qui sert de base de sondage. En fait, une stratification géographique, suivant la région et la catégorie de communes, est particulièrement simple à réaliser concrètement, puisque les documents du recensement de la population sont justement classés par communes. Mais rien n'excluerait qu'on stratifie l'échantillon selon d'autres critères observés lors du recensement, comme la catégorie sociale du chef de ménage ou la taille des familles. Cela n'a jamais été fait en France, mais fournirait probablement un avantage illusoire. En effet, il faut retenir qu'on ne tire pas des ménages, mais des logements, et qu'on interroge l'occupant de ces logements, même s'il n'est plus le même qu'à l'époque du recensement. Les changements de domicile sont suffisamment fréquents pour que des caractères liés au ménage et non au logement soient rapidement caducs : c'est pourquoi, outre les raisons de coût, on ne stratifie l'échantillon qu'en fonction de caractères assez permanents, liés au logement.

Ensuite, il convient de ne pas limiter son attention aux erreurs d'échantillonnage. En matière de consommation, les erreurs, systématiques ou accidentelles, de mesure, liées aux défauts de mémoire des enquêtés, au manque de soin des enquêteurs ou à la difficulté propre du sujet, sont beaucoup plus considérables que

les erreurs d'échantillonnage. L'intérêt d'un bon plan de sondage est qu'il assure une erreur d'échantillonnage négligeable pour une taille d'échantillon compatible avec une formation et un contrôle sérieux des enquêteurs. C'est vers cet objectif que devraient tendre les principaux efforts. Notre impression est que les enquêtes extensives faites dans la dernière décennie aux Etats-Unis, en Angleterre et en France (9) ont porté sur des échantillons trop nombreux pour assurer ce contrôle avec tout le soin voulu. Le souci d'obtenir des estimations correctes des consommations, non seulement pour l'ensemble de la population, mais aussi pour des sous-populations restreintes, a conduit à faire porter l'enquête sur des échantillons très lourds par rapport à l'importance des moyens d'enquête et de contrôle disponibles.

Enfin, on remarquera que nous n'avons considéré les problèmes d'échantillonnage que du point de vue de l'estimation d'une moyenne. C'est effectivement ainsi qu'ils sont étudiés en pratique et les résultats en sont regrettables pour les études économétriques. Un échantillon efficient pour l'estimation d'une moyenne ne l'est pas en effet pour l'estimation d'une régression et, a fortiori, pour la comparaison des régressions dans différentes sous-populations. Pour les études économétriques, l'échantillon devrait être choisi selon des méthodes inspirées par la technique des plans d'expérimentation. Cela conduirait à augmenter le nombre d'observations dans les tranches de revenus les plus fortes (où la variance des consommations est élevée) et les plus faibles, et à réduire le nombre d'observations voisines de la moyenne. Cela conduirait également à tirer des échantillons de taille voisine pour des sous-populations d'effectifs très différents, mais dont on voudrait efficacement comparer les lois de consommations.

Bien que les remarques précédentes ne soient plus originales, elles n'ont pas encore, à notre connaissance, été effectivement appliquées à des enquêtes de consommations. Il y a à cela deux raisons principales. La première est le coût élevé d'une enquête établie selon ces méthodes. La seule façon de procéder consisterait en effet à tirer d'abord un échantillon surabondant de ménages, à observer la valeur approximative des variables de contrôle choisies, à définir des strates selon ces variables et à tirer un sous-échantillon avec une fraction de sondage variable selon les strates. C'est auprès de ce sous-échantillon que devrait porter l'enquête principale. La seconde raison est qu'on n'est pas assez sûr du choix des hypothèses théoriques à vérifier par une telle enquête pour oser encore la préparer.

SECTION 3

L'ENQUETE FRANCAISE DE 1956

Cette enquête, financée par une dotation du Fonds National de la Productivité, a été exécutée en collaboration par l'I.N.S.E.E. et le C.R.E.D.O.C. En fait, la conception de l'enquête et l'exploitation mécanographique ont été faites en commun, le choix de l'échantillon et le travail sur le terrain ont été la responsabilité de l'I.N.S.E.E. ; la détermination du plan d'exploitation, l'analyse et l'interprétation des résultats la tâche du C.R.E.D.O.C. Dans cette section, nous nous inspirerons étroitement du texte que nous avons fait établir pour présenter le compte rendu de cette enquête (10).

§ 1 - Le plan de sondage

La taille de l'échantillon a été fixée à environ 22.800 ménages ce qui correspond à un taux de sondage uniforme de 1 : 625. Ce chiffre s'applique à l'ensemble de l'enquête qui a eu lieu en quatre vagues successives, une pour chaque trimestre. Cette répartition avait pour but d'alléger la charge des enquêteurs et surtout d'éliminer l'effet des variations saisonnières des résultats de l'enquête. L'échantillon des ménages a été renouvelé à chacune des quatre vagues, mais les quatre échantillons ont rigoureusement les mêmes caractéristiques techniques.

L'échantillon a été tiré dans le fichier des logements établi à l'occasion du recensement de la population de Mai 1954. Comme ce fichier n'a pas été tenu à jour les ménages occupant les logements construits depuis le recensement ont échappé à l'enquête, ce qui est un défaut d'ailleurs relativement secondaire. Il s'agit d'environ 500.000 ménages. Une conséquence plus importante du choix de cette base de sondage est que la définition du ménage selon le recensement est imposée à l'échantillon et par suite à l'enquête entière. En principe, un ménage est l'ensemble de toutes les personnes vivant dans un seul logement. C'est donc un concept plus large que la famille. Un ménage peut comprendre des pensionnaires

ou sous-locataires, des domestiques et même parfois plusieurs couples, par exemple dans le cas où un jeune ménage est logé chez ses parents. Dans les rares cas où un pavillon ou un appartement est divisé en deux locaux indépendants, ayant chacun sa propre entrée, on les compte pour deux logements ; ceci exige d'ailleurs que les occupants des deux parties mènent des vies nettement distinctes et ne prennent généralement pas leurs repas ensemble. Signalons enfin que le fichier des logements comprend les logements vacants, les résidences secondaires et les chambres d'hôtels ou de pensions affectées à une occupation permanente par un ménage.

Le tirage de chaque échantillon suit la procédure commune appliquée à toutes les enquêtes auprès des ménages faites par l'I.N.S.E.E. L'échantillon est déterminé en trois étapes successives :

- la première définit un échantillon de cantons ;
- dans la seconde étape, on tire un échantillon de communes parmi les cantons échantillons ;
- Enfin, un échantillon de logements est tiré dans les communes échantillons.

Cette procédure assure d'une part, une répartition équilibrée selon les régions et la taille des agglomérations, et évite d'autre part une dispersion trop forte au niveau local qui rendrait l'exécution de l'enquête coûteuse et moins efficace. L'échantillon de cantons qui représente la première étape a été déterminé une fois pour toute en 1954 : il a fourni un cadre permanent à toutes les enquêtes auprès des ménages entreprises par l'I.N.S.E.E. depuis cette date jusqu'en 1959, les deux autres degrés de tirage étant renouvelés pour chaque enquête. En janvier 1960, un nouvel échantillon maître a été établi.

Le plan de sondage de l'échantillon maître de cantons est assez complexe ; nous ne pouvons qu'en indiquer les principales caractéristiques. On a d'abord découpé le territoire en strates assez homogènes. Ces strates peuvent être réparties en deux groupes, suivant le degré d'urbanisation :

a. Le premier groupe consiste en 55 cantons ou groupes de cantons comprenant chacun une agglomération de plus de 50.000 habitants. Les 55 villes principales de ces agglomérations sont incluses d'office dans l'échantillon et sondées à un seul degré. Les communes secondaires (banlieue) sont sondées à deux degrés : dans chaque banlieue, on tire au sort quelques communes échantillons, et dans ces dernières un échantillon de logements. Il en est de même des communes situées à la périphérie de chaque canton ou groupes de cantons, en dehors de l'agglomération de 50.000 habitants et plus (sondage à deux degrés : communes, logements).

b. Les autres cantons sont répartis en 54 strates par une double classification selon la région et le degré d'urbanisation. La France est divisée en huit régions indiquées sur la carte II-2. Dans chaque région, les cantons sont classés suivant la taille de leur plus grande agglomération et, éventuellement, suivant la proportion de la population agricole. Ces strates ont été sondées à trois degrés : on a d'abord tiré au sort trois ou quatre cantons par strate lors de l'établissement de l'échantillon-maître permanent, ensuite on tire pour chaque enquête deux communes dans chaque canton et des logements dans ces communes.

Le tirage des cantons et communes dans les strates du dernier groupe et le tirage des communes dans les strates de zones agglomérées du groupe a) ont été effectués de telle sorte que la chance d'être tiré soit proportionnelle au nombre de logements. L'échantillon de logements est ensuite déterminé par tirage systématique dans le fichier de logements des communes, ce fichier étant classé par îlot. Dans les grandes villes intégrées d'office dans l'échantillon-maître, le nombre de logements est proportionnel à la population, mais dans les communes tirées au sort, soit dans les zones agglomérées soit dans les strates du groupe b) le même nombre de logements est prélevé dans chaque commune d'une même strate. Comme ces communes ont été choisies à l'étape antérieure avec des chances proportionnelles à leur population, en fin de compte chaque logement recensé en 1954 a la même chance de faire partie de l'échantillon.

Signalons que l'échantillon ainsi tiré du fichier des logements consiste en une liste de logements identifiés par leur adresse et le nom du ménage qui l'occupait lors du recensement. L'échantillon comprend également des logements vacants ou secondaires au recensement ; ils sont maintenus dans l'échantillon afin que les enquêteurs puissent vérifier qu'ils ne sont pas occupés comme logement principal à la date de l'enquête.

FIGURE II-2

LES HUIT REGIONS DE SONDAGE

Notons enfin que l'échantillon dont la détermination a été décrite ci-dessus est l'échantillon initial. Cet échantillon subit une déformation due aux divers déchets et refus dont nous parlerons dans le paragraphe 3. Il est ensuite redressé pour compenser partiellement l'effet de ces déformations comme nous l'indiquons dans le paragraphe 4. C'est l'échantillon redressé qui a servi de base à tous les résultats que nous utiliserons.

§ 2 - La méthode d'enquête

Une fois l'échantillon tiré, les ménages désignés ont été interrogés par des enquêteurs spécialisés. La plupart d'entre eux font partie du réseau permanent d'enquêteurs de l'I.N.S.E.E., mis en place après le tirage de l'échantillon-maître de cantons qui a servi à toutes les enquêtes auprès des ménages de 1954 à 1960. On verra plus loin comment l'enquête s'est effectivement déroulée ; nous ne considérerons pour l'instant que les procédures d'enquête prévues. Le questionnaire d'enquête est reproduit hors texte.

Chaque ménage tiré a été prévenu du passage de l'enquêteur par une brève lettre d'avis. L'enquêteur se rend au domicile et essaie d'abord de vérifier l'identité et les caractéristiques principales du ménage (composition, profession du chef), telles quelles lui sont indiquées d'après le fichier du recensement de 1954. Si l'enquêteur trouve un autre ménage installé dans le logement désigné, celui-ci est substitué au ménage indiqué. Il se peut évidemment que l'enquêteur ne parvienne pas à identifier le logement indiqué sur la liste ; ou qu'il le trouve vacant ou affecté à un autre usage que l'habitation, ou enfin qu'il ne réussisse pas à joindre le ménage qui l'occupe. Dans ce dernier cas, l'enquêteur doit cependant entreprendre plusieurs visites avant de renoncer.

Quand l'enquêteur a identifié le ménage et s'il ne s'est pas heurté à un refus immédiat, il présente l'enquête, explique le carnet de comptes et aborde l'interview prévue dans le questionnaire. Le carnet de comptes est destiné à être rempli par la ménagère qui est priée d'y inscrire les dépenses courantes et, dans le cas d'exploitants agricoles ou de commerçants, les quantités de produits alimentaires autoconsommés ou autofournis : par exemple, le poulet pris dans la basse-cour ou la boîte de conserve prélevée dans le magasin. Le carnet doit être rempli jour par jour pendant sept jours consécutifs, à partir du lendemain de la visite de l'enquêteur. Vers le milieu de cette période l'enquêteur fait une

seconde visite pour vérifier le remplissage du carnet et, dans un délai de deux jours après la fin de la période, il vient une dernière fois pour le ramasser. Seules les dépenses courantes (principalement les dépenses d'alimentation) sont de cette façon relevées par la ménagère elle-même, bien qu'elle soit aidée par l'enquêteur;tous les autres renseignements sont obtenus par l'enquêteur par interview.

Le carnet de comptes peut prendre deux formes différentes. La ménagère peut inscrire les dépenses courantes directement dans les pages 2 à 6 du questionnaire qui constituent un carnet de comptes journaliers avec une classification détaillée de produits ; mais si cette procédure paraît trop difficile, l'enquêteur peut y substituer un carnet-brouillon, ne mentionnant aucun produit, dans lequel la ménagère doit relever toutes les dépenses du ménage. Dans ce cas, l'enquêteur transcrit plus tard les dépenses courantes de ce brouillon sur le questionnaire.

Parcourons rapidement le questionnaire. La première page comporte des renseignements généraux sur le ménage, comme sa composition et la profession de son chef ; l'enquêteur doit le remplir lors de sa première visite, de sorte que ces renseignements soient disponibles même en cas d'échec ultérieur de l'enquête. Les pages 2 à 6 servent à l'inscription des dépenses courantes et des quantités autoconsommées ; les huit pages suivantes correspondent à l'interview qui est fort longue et fort compliquée. Toutes les autres dépenses du ménage sont relevées dans cette partie du questionnaire et la liste de produits, jointe comme aide-mémoire à l'instruction aux enquêteurs, est utilisée à plusieurs reprises. L'enquêteur est libre de répartir l'interview à son gré entre ses trois visites bien que ses instructions penchent vers une répartition à peu près égale. La dernière page du questionnaire comporte des renseignements sur les ressources du ménage ; cette partie a été formellement réservée à la dernière visite car on a craint qu'elle provoque un grand nombre de refus.

L'objectif de l'enquête est de saisir l'ensemble des dépenses de consommation du ménage. On remarquera que le questionnaire ne porte que sur des dépenses, sauf pour l'autoconsommation où le relevé de quantités s'impose. L'enquête est simplifiée aussi du fait qu'on a renoncé à interroger tous les membres du ménage et à isoler leurs dépenses. Par contre, le souci de renseigner sur <u>toutes</u> les catégories de dépenses a conduit à une interview fort compliquée, notamment en ce qui concerne le logement et les achats à crédit.

La période pendant laquelle les dépenses sont relevées varie selon les produits. Son choix a de l'importance. Une période trop longue exige un effort de mémoire excessif, et on peut craindre que les réponses de la personne interrogée sous-estiment la dépense effective. Si la période est courte, par contre, les dépenses des produits dont les achats sont peu fréquents sont estimées avec des erreurs d'échantillonnage importantes. En outre, il semble qu'une période trop courte puisse conduire à une surestimation systématique des dépenses légèrement moins fréquentes ; ceci serait dû au fait que les personnes interrogées auraient tendance à déclarer des dépenses en fait antérieures à la période visée, en préférant dans le doute fournir des réponses positives (mais inexactes) aux questions qui leur sont posées. Le même motif est en jeu quand la vue d'une liste de produits conduit la ménagère à anticiper certains achats qu'elle aurait normalement faits après la période d'enregistrement des dépenses. Rappelons que le carnet de comptes qui fait partie du questionnaire comporte une telle liste suggestive de produits, mais que le carnet brouillon ne mentionne par contre aucun produit.

Ces considérations ont amené à distinguer quatre périodes différentes, les produits étant répartis selon la fréquence habituelle d'achat :

- Une semaine pour les dépenses courantes et l'autoconsommation alimentaire à relever par la ménagère. Ces dépenses courantes comprennent l'alimentation, les achats de produits d'entretien, mercerie, pharmacie, journaux, fleurs, services personnels, tabacs, dépenses relatives aux loisirs, transports, P.T.T. et quelques autres catégories (l'ensemble énuméré dans les pages 2 à 6 du questionnaire).

- Un mois pour les biens durables peu coûteux (livres, vaisselle, etc...), pour une partie des dépenses de santé (page 13 du questionnaire) et pour l'essence et l'huile parmi les frais d'utilisation de voiture (page 11).

- Trois mois pour les vêtements, les chaussures et les combustibles (pages 9 et 12 du questionnaire).

- Douze mois pour certaines dépenses de logement (page 8), pour les biens durables (page 10), les vacances et voyages (pages 14 et 15) certains frais d'utilisation de voiture (page 11) et de santé (page 13).

- Finalement, il y a certaines catégories de dépenses (loyer, eau, gaz, etc...) (pages 7 et 9), pour lesquelles ont a demandé d'indiquer le montant de la dernière quittance ainsi que la période à laquelle elle se rapportait.

Les questions relatives aux ressources du ménage représentent une innovation ; dans les enquêtes antérieures on avait évité des questions précises à ce sujet par crainte de provoquer des refus. L'enquête pilote n'a pas confirmé cette crainte : 6 % seulement des ménages ont refusé de répondre aux questions sur leur revenu. En même temps, ces résultats ont indiqué qu'il valait mieux demander les détails des revenus d'un ménage que poser la question de façon globale : même si les estimations obtenues restent incertaines, une analyse détaillée des ressources du ménage réduit le risque d' omissions importantes sans augmenter sensiblement le taux de refus. Le questionnaire de l'enquête définitive a par conséquent analysé les revenus personnels des membres du ménage selon leurs sources. Ces questions visent les revenus des douze derniers mois ou de la dernière année du calendrier.

§ 3 - Les conditions d'exécution de l'enquête

Comme le réseau comprenait environ 300 enquêteurs, il était prévu que chacune des quatre vagues de l'enquête, touchant 5.700 ménages, s'étalerait sur quatre semaines, représentatives d'un trimestre ; cependant, comme la première se situait en Avril, la seconde a du être avancée d'un mois afin d'éviter que les départs en vacances ne réduisent le taux de réponses. En outre, certaines Directions Régionales (D.R.) de l'I.N.S.E.E. n'ont pas pu faire les enquêtes dont elles étaient chargées dans le délai fixé. En fin de compte, l'enquête a eu lieu dans les périodes suivantes :

- Première enquête : 12 Avril-10 Mai 1956, sauf à la D.R. de Paris où elle n'a été terminée que fin Mai.

- Deuxième enquête : 14 juin-12 juillet 1956 ; la D.R. de Rennes l'a poursuivie jusqu'au 11 Septembre, la D.R. de Dijon l'a retardée jusqu'à la période du 1er Septembre au 8 Octobre.

- Troisième enquête : 1er Octobre-29 Octobre 1956. Pour la D.R. de Paris : 17 Septembre-5 Novembre ; pour la D.R. Montpellier : 25 Novembre-23 Décembre.

- Quatrième enquête : 7 Janvier-4 Février 1957.

Cette distribution des quatre vagues dans l'année élimine convenablement les variations saisonnières de la plupart des dépenses de consommation pour l'ensemble de l'enquête. Comme les enquêtes se situent à des intervalles proches d'un trimestre, un élément saisonnier ne peut évidemment subsister que pour les dépenses relevées sur une période inférieure. Pour plusieurs produits alimentaires, relevés pendant une semaine, une vérification est possible à l'aide des coefficients saisonniers mensuels établis à partir d'autres sources (11). On trouve en effet que la moyenne de ces coefficients pour les quatre périodes visées (avril-Mai, Juin-Juillet, Octobre et Janvier) est en général assez proche de la moyenne annuelle, comme le montre le tableau ci-dessous :

Moyennes des coefficients saisonniers mensuels pour l'ensemble des quatre périodes de l'enquête (moyenne annuelle = 100)

Alimentation générale et épicerie	101,1
Farine, pain, patisserie et confiserie	99,6
Boucherie	98,9
Charcuterie	104,1
Produits de la mer et de l'eau douce	99,1
Produits laitiers, basse-cour, gibier	99,1
Fruits et légumes	103,5

o

o o

Les enquêteurs ont subi des échecs dus à plusieurs causes. Rappelons que l'échantillon initial est une liste de logements qui comprend des logements vacants ou secondaires, chaque logement étant défini par son adresse et, le cas échéant, par le nom du ménage qui l'occupait lors du recensement de 1954. En premier lieu, il se peut que l'enquêteur ne parvienne pas à retrouver le logement indiqué. Ce cas est rare et il concerne le plus souvent des logements vacants à la date du recensement. En second lieu, le logement identifié par l'enquêteur peut ne pas héberger de ménage : logement resté ou devenu vacant ou secondaire, logement converti en local commercial ou détruit depuis le recensement. Parmi les logements habités qui restent, il y a encore deux catégories d'échecs. D'abord il y a des logements apparemment occupés où l'enquêteur n'a pas pu joindre le ménage, personne n'étant jamais présent au domicile. Ces cas peuvent être dus à une absence de longue durée ; ils correspondent souvent à des résidences secondaires ou, ce qui est plus important, à des ménages qui par leur nature sont difficiles à toucher : Célibataires ou couples où les conjoints ont tous les deux un emploi. Ensuite, certains ménages refusent de répondre à l'enquête, soit au début, soit au cours de l'opération. Ce dernier groupe se confond avec la catégorie, beaucoup moins nombreuse, des ménages dont les questionnaires, imparfaitement remplis, ont été rejetés lors de la vérification ultérieure. Le tableau page 69 montre l'incidence des quatre catégories d'échecs énumérés dans chaque vague de l'enquête.

L'échantillon se trouve ainsi réduit ; certains échecs sont cependant d'un caractère plutôt technique, liés à l'usage d'un fichier de logements déjà ancien comme base de sondage. Pour évaluer la réussite de l'enquête auprès du public, on ne doit considérer que les refus inscrits à la ligne 5 du tableau, ce qui conduit à un taux de réponses de 82,6 % parmi les ménages effectivement touchés par les enquêteurs (12). Ce résultat est satisfaisant ; il ne s'applique cependant qu'à l'ensemble de l'enquête. L'importance des refus réside surtout dans le fait qu'ils ne sont pas distribués uniformément dans l'échantillon initial et que leur incidence risque de déformer la composition de l'échantillon effectif. Dans la mesure où ils traduisent une caractéristique du ménage (célibataires isolés, conjoints ayant tous les deux un emploi), les cas d'absence peuvent également influencer la composition de l'échantillon d'une façon systématique. L'effet de ces deux catégories de déchets sur l'échantillon est évidemment difficile à estimer, étant donné qu'on dispose de peu de renseignements sur les ménages auxquels ils correspondent. On peut cependant se faire une idée de la déformation de l'échantillon due à toutes les causes d'échecs réunies en comparant l'échantillon des ménages observés avec le recensement de 1954. Cette comparaison est à la base du redressement de l'échantillon que nous décrivons au paragraphe 4.

(Voir tableau page 69)

	1ère enquête	2e enquête	3e enquête	4e enquête	Ensemble	En % de l'ensemble
1. Echantillon initial	5.641	5.723	5.723	5.704	22.791	100
2. Logements dont on ne peut retrouver aucune trace	90	69	53	64	276	1,2
3. Logements vacants, secondaires ou convertis	323	309	368	356	1.356	6,0
4. Ménages impossibles à joindre, absence de longue durée	253	360	292	293	1.198	5,3
5. Ménages refusant de répondre, questionnaire imparfaitement rempli	811	853	981	817	3.462	15,2
6. Echantillon effectif	4.164	4.132	4.029	4.174	16.499	72,3

Chez les 16.500 ménages qui ont collaboré à l'enquête, la collecte des renseignements ne s'est pas toujours déroulée selon la procédure prévue. Même si on ne leur oppose pas un refus immédiat, les enquêteurs trouvent parfois que l'accueil ne s'améliore pas avec la réflexion et que l'interview devient de plus en plus pénible lors des visites successives. Beaucoup d'enquêteurs ont par suite pris l'habitude de placer toute l'interview autant que possible lors de la première visite. En effet, le plus souvent seul le carnet de comptes brouillon et non pas le questionnaire a été remis à la ménagère. Quant à l'utilisation du carnet de comptes, s'il a matériellement contribué à un relevé exact des dépenses hebdomadaires, il a aussi dans un bon nombre de cas été, en fait, rempli ou complété en interview par l'enquêteur lors de sa dernière visite. On ne peut certes pas prétendre que l'enquête ait fourni 16.500 questionnaires remplis sans erreurs ni lacunes. Ces dernières avaient d'ailleurs été prévues, au moins dans l'interview ; l'instruction aux enquêteurs distingue nettement entre dépenses enregistrées, dépenses nulles et dépenses non déclarées. Pour la question relative aux ressources où cette catégorie risque d'être plus forte, une quatrième option a été prévue :

c'est celle des ressources estimées par l'enquêteur. Ce chiffre suffit pour classer le ménage dans une catégorie de revenu. Cette possibilité a permis de réduire la proportion des questionnaires qui ne comportent aucune indication du revenu à 5 %.

Sur le fond, la méthode d'enquête peut être jugée, à l'expérience, de la façon suivante :

1. Le carnet de comptes a été utile, même s'il a en fait été rempli par l'enquêteur : il a conduit la ménagère à prêter plus d'attention à ses achats dont elle s'est très souvent efforcée de conserver une trace. On peut craindre cependant que le questionnaire, sous la forme utilisée, n'ait parfois influencé le comportement de la ménagère en lui suggérant certains achats surtout pour des produits stockables dont la périodicité d'achats habituelle est certainement supérieure à la semaine. De ce point de vue, l'utilisation uniforme du carnet de comptes brouillon, sans indication de produits pour l'enregistrement de tous les achats, aurait été préférable.

2. Les périodes de référence adoptées dans l'interview ont donné satisfaction, sauf pour la lingerie et la bonneterie où il aurait fallu enregistrer les dépenses d'un mois plutôt que d'un trimestre. Le changement de période d'une question à une autre a cependant sensiblement alourdi l'interview, et rendu plus difficile la tâche des enquêteurs.

3. L'étendue de l'étude a conduit à un questionnaire trop synthétique et trop elliptique ; sur ce point, l'aide-mémoire de produits donné en annexe des instructions aux enquêteurs a été moins efficace que l'aurait été une rédaction plus détaillée du questionnaire. Dans le questionnaire, tel qu'il a été utilisé, la nomenclature est parfois trop succincte, ce qui est le cas de la rubrique vêtements principaux, parfois le dessin des tableaux prête à confusion et favorise l'oubli, comme dans le relevé de la consommation hebdomadaire qui porte à la fois sur les dépenses et sur les quantités autoconsommées d'alimentation, ou comme dans le tableau II qui groupe les véhicules, l'équipement ménager et les meubles.

4. Bien que le questionnaire n'admette en principe aucune équivoque à ce sujet, il semble qu'il ait en fait été parfois difficile de séparer les dépenses de consommation des dépenses d'exploitation, notamment chez les agriculteurs. Dans la mesure où l'exploitant agricole nourrit son personnel, l'autoconsommation peut également confondre des frais d'exploitation avec la consommation du ménage.

5. Les questions sur le revenu ont provoqué moins de refus qu'on ne le craignait. Ceci ne veut cependant pas dire que les réponses obtenues donnent satisfaction. Il semble que les ressources soient en général nettement sous-estimées ; ici encore, les ménages agricoles posent des problèmes particulièrement difficiles et leurs ressources ont été plutôt mal appréhendées.

6. Ne pas avoir interrogé tous les membres du ménage a conduit à des omissions, dans le cas notamment de certaines consommations faites hors du domicile. Il en résulte que les dépenses de tabac et de boissons sont nettement sous-estimées.

Malgré ces réserves, l'enquête a rassemblé une documentation importante qui faisait défaut. Le coût de collecte a été modeste : pour les 16.500 questionnaires remplis, le prix de revient a été de 1.500 frs par pièce.

§ 4 - Le fichier mécanographique

A partir des questionnaires remplis et acceptés on a d'abord constitué un fichier de cartes perforées, dites "cartes détail" qui contiennent presque tous les renseignements portés sur le questionnaire. En nous référant à ce dernier, nous signalerons les omissions et simplifications suivantes :

- N'ont pas été perforés : page 1, le département, la composition du ménage autrement que par effectifs selon le groupe d'âge, la catégorie socio-professionnelle des autres membres que le chef du ménage ; page 7, les réponses aux questions 1 et 3.

- Ont été supprimés les détails des tableaux II (page 10), III (page 11), VII (page 14) et X (page 15), dont on n'a perforé que les totaux ; du tableau IV (page 12), dont seule la dernière colonne a été perforée.

Les cartes perforées sont du type classique, divisées en 80 colonnes permettant chacune un choix de 12 positions. Dans le cas des données numériques, les colonnes sont réparties en zones ; un nombre de cinq chiffres est ainsi porté dans une zone de 5 colonnes, chaque colonne correspondant à une unité d'un rang donné (centaines, milliers, etc...). On peut également perforer d'autres indications telle que la catégorie socio-professionnelle du ménage, en traduisant ces caractéristiques par des codes chiffrés : par exemple, exploitants agricoles 01, salariés agricoles 02, etc...

Les données ont été réparties sur 45 types différents de cartes au total, correspondant chacune à un groupe de produits consommés. A l'exception d'un seul type de carte, établi d'office pour tous les ménages, les cartes n'ont été perforées que dans le cas où elles portaient au moins une dépense non nulle. De ce fait le nombre effectif de cartes par ménage est considérablement réduit et la moyenne est de 19 seulement.

Les 29 premières colonnes sont identiques pour tous les types de cartes-détail. Elles portent des renseignements généraux : numéro d'identification du ménage qui comporte le numéro d'ordre des quatre vagues de l'enquête, la Direction Régionale, la région de sondage de 1 à 8 et la catégorie de commune, composition du ménage, catégorie socio-professionnelle du chef, nombre de salariés, nombre d'unités de consommation, niveau de revenu.

Les autres colonnes portent le détail des dépenses, des quantités auto-consommées et des autres renseignements. La plupart des montants ont été perforés en zones de quatre colonnes, les unités étant en francs pour les dépenses courantes, 100 grammes ou décilitres pour l'auto-consommation, en cent francs pour l'interview. Les données chiffrées n'ont pas été transformées en unités homogènes, de sorte que les dépenses perforées se rapportent selon le cas à 7 jours, un mois, un trimestre ou douze mois.

C'est sur ce fichier mécanographique de cartes-détail qu'on a opéré le redressement de l'échantillon de ménages, dont le but est d'éliminer l'influence des déchets survenus au cours de l'enquête ainsi que certaines déformations qui ont pû se produire lors du tirage de l'échantillon initial. Ce redressement a été effectué séparément pour chacune des quatre vagues de l'enquête. Les ménages interrogés sont d'abord classés suivant trois critères.

- la catégorie de commune,
- la catégorie socio-professionnelle,
- le nombre de personnes,

Cette classification est ensuite comparée à la distribution, selon les mêmes critères, des ménages occupant les logements principaux au recensement de 1954. L'écart entre les deux distributions traduit les déformations que l'échantillon a subies ; on élimine l'écart et on redresse l'échantillon par duplication des cartes de ménages appartenant aux catégories déficitaires. Les ménages dont les cartes sont dédoublées sont tirés systématiquement dans l'échantillon de ménages interrogés. Une application rigoureuse de cette procédure aurait cependant conduit, dans certains cas, à la multiplication par trois ou quatre de certains ménages isolés, ce qui risquerait d'augmenter les erreurs aléatoires. On s'est donc borné à un redressement de l'échantillon qui fait coïncider seulement les distributions marginales selon chacun des trois critères avec les distributions théoriques, c'est à dire les distributions des logements principaux au recensement de 1954. Le redressement a en outre été effectué de telle façon que le nombre de ménages dans l'échantillon redressé correspond à un taux de sondage de 1 : 625 (1 : 2.500 pour chacune des quatre vagues) des logements principaux en 1954. Le tableau ci-après indique le nombre de ménages dédoublés selon les trois classifications.

Le fichier redressé de cartes perforées ainsi obtenu a servi de base à toutes les exploitations de l'enquête. Nous indiquons seulement les traits principaux de la technique d'exploitation mécanographique. Les données relatives aux ménages individuels ont d'abord été additionnées telles qu'elles sont perforées dans les cartes détail sur des cartes récapitulatives qui portent l'ensemble des dépenses de chaque produit pour un groupe de ménages défini par une classification donnée de l'échantillon. Il y a un grand nombre de telles classifications à double ou triple critère de tri, dont chacune donne lieu à un jeu de cartes récapitulatives : les tableaux 1 et 2 du chapitre III correspondent par exemple à un tel jeu.

	Echantillon observé (1)	Correction (2)	Echantillon redressé (1) + (2)	Echantillon théorique (3)
Catégorie de commune				
- Communes rurales	7.785	971	8.756	8.748
- Communes urbaines de moins de 10.000 habitants	1.897	543	2.440	2.432
- Communes urbaines de 10 à 50.000 habitants	2.066	534	2.600	2.594
- Communes urbaines de 50 à 100.000 habitants	717	211	928	932
- Villes de province de 100.000 habitants et plus	1.976	980	2.956	2.951
- Agglomération Parisienne.	2.060	1.760	3.820	3.812
Ensemble	16.501	4.999	21.500	21.469
Effectif du ménage				
- 1 personne	2.362	1.638	4.000	4.111
- 2 personnes	4.210	1.723	5.933	5.856
- 3 personnes	3.280	889	4.169	4.246
- 4 personnes	2.790	404	3.194	3.162
- 5 personnes	1.896	189	2.085	1.978
- 6 personnes et plus	1.963	156	2.119	2.116
Ensemble	16.501	4.999	21.500	21.469
Catégorie socio-professionnelle du chef de ménage				
- Exploitants agricoles	2.741	91	2.832	2.818
- Salariés agricoles	673	51	724	772
- Industriels, artisans, commerçants	1.693	564	2.257	2.337
- Professions libérales, cadres supérieurs	501	195	696	724
- Cadres moyens	909	272	1.181	1.101
- Employés	1.181	409	1.590	1.528
- Ouvriers	4.662	1.203	5.865	5.884
- Salariés divers	773	230	1.003	971
- Inactifs : rentiers, retraités	3.368	1.984	5.352	5.334
Ensemble	16.501	4.999	21.500	21.469

(1) Correspondant au nombre de questionnaires remplis pour l'ensemble des 4 enquêtes.
(2) Nombre de jeux de cartes-détail doublés.
(3) Effectif théorique selon le recensement de 1954, d'après le dépouillement au 1/20 du fichier des logements principaux.

Les dépenses ont ensuite été regroupées selon une nomenclature sommaire de produits à partir de ces cartes récapitulatives.

L'exploitation mécanographique a été compliquée par le fait que les données de base perforées dans les cartes détail ne sont pas homogènes, mais sont exprimées en unités diverses. Il aurait été difficile de chiffrer tous les montants dans une seule unité avant de perforer les cartes, au moins pour les quantités d'autoconsommation : l'application des divers prix que cette procédure exigerait n'est guère praticable. Cependant la transformation en unités homogènes des données relatives à chaque ménage aurait été préférable. Une transformation mécanographique des données avec reproduction des cartes-détail l'aurait permise, mais n'a pas été possible pour des raisons de coût. C'est en effet seulement à partir de données homogènes qu'on peut déterminer facilement la dépense totale par ménage et trier les ménages selon ce critère. Dans la procédure adoptée, ceci a été impossible. Nous avons dû calculer la dépense totale comme moyenne pour des groupes de ménages, à partir des cartes récapitulatives. Dans la définition des groupes, nous avons dû nous contenter de classer les ménages selon les ressources au lieu de la dépense totale.

Signalons enfin une omission dans les cartes récapitulatives : dans l'addition des dépenses de chaque poste par groupe de ménages, le nombre de ménages n'ayant pas déclaré des dépenses pour ce poste n'a pas toujours été repéré. Sauf en ce qui concerne les ressources, on n'a donc pas pu tenir compte du nombre de non déclarations, ce qui peut introduire une sous-estimation systématique dans les dépenses moyennes. Il semble cependant que cette erreur, qui ne s'applique qu'aux dépenses relevées en interview, soit peu importante.

En outre, il faut admettre que malgré tout effort de vérification des erreurs ont été parfois commises dans les opérations que nous venons de décrire. Ces erreurs peuvent être gênantes, sans cependant avoir trop d'importance réelle.

§ 5 - La qualité des résultats

Le principal résultat de l'enquête est une estimation de la consommation des ménages français analysée en une centaine de produits. La précision de ces chiffres est difficile à apprécier. La procédure qu'on a suivie a certes réduit les erreurs d'échantillonnage au minimum. Mais les erreurs lors de la collecte des renseignements de base sont, nous l'avons dit plus haut, plus difficiles à éviter.

Nous avons déjà indiqué plusieurs possibilités d'erreurs de ce genre. Certaines d'entre elles auraient pu être éliminées par une vérification plus stricte des questionnaires avant leur acceptation. Quand au chiffrement des questionnaires et le traitement mécanographique des données, les quelques erreurs qu'on peut y signaler ont peu d'importance.

Outre cette appréciation qualitative de l'enquête il serait évidemment souhaitable d'apprécier la précision des chiffres obtenus par une confrontation directe avec des statistiques indépendantes. Pour quelques produits particuliers (habillement, viande, produits laitiers) on dispose des enquêtes antérieures, mais la seule source qui permettrait une comparaison systématique de l'ensemble de la consommation est l'estimation de la consommation territoriale globale établie par le C.R.E.D.O.C., en collaboration avec le Service des Etudes Economiques et Financières du Ministère des Finances (S.E.E.F.) (13). Ces chiffres sont cependant eux-mêmes des estimations de synthèse, parfois d'une précision douteuse. En effet la confrontation des deux sources a pour plusieurs postes amené une modification des estimations globales dont on trouvera le reflet dans les séries définitives qui seront prochainement publiées. En outre, l'enquête d'une part, et les estimations globales d'autre part, ne mesurent pas tout à fait les mêmes grandeurs. L'enquête ne porte que sur la consommation des ménages résidant en France ; les chiffres globaux comprennent en outre la consommation de la population des institutions (militaires, détenus, hospitalisés), des travailleurs frontaliers et des touristes étrangers, ainsi que les avantages en nature touchés par les ménages. Finalement les renseignements de l'enquête ne permettent pas de répartir les consommations en dehors du ménage (café, restaurants, cantine) selon les produits. Bien qu'on puisse tenir compte de ces différences en ajustant les estimations d'une part ou d'une autre, il reste une certaine marge d'incertitude, qui rend toute comparaison précise illusoire. Dans ces conditions des écarts de 5 ou 10 % doivent être considérés comme normaux et acceptables ; seuls les cas dans lesquels un chiffre global bien fondé accuse un désaccord important avec l'enquête mérite l'attention.

Ces cas sont pour la plupart explicables. La consommation de boissons alcoolisées et de tabac est nettement sous-estimée par l'enquête. On sait qu'il s'agit là d'un phénomène universel dans les enquêtes par sondage. Une sous-estimation de 30 à 50 % est apparue dans les enquêtes de budgets familiaux réalisées récemment en Angleterre et aux Etats-Unis ; le pourcentage est du même ordre de grandeur dans l'enquête française. La sous-estimation de la consommation de boissons

semble notamment plus forte pour la consommation prise à l'extérieur que pour les achats destinés à être consommés à domicile. Comme dans le cas de la consommation de tabac, l'erreur semble être due au fait que dans la plupart des cas les renseignements de l'enquête ont été recueillis auprès de la ménagère seule.

Pour la plupart des autres postes, la concordance est satisfaisante. On observe cependant, pour quelques produits d'alimentation stockables (huile, riz), une surestimation qui serait dûe à des achats dépassant les besoins habituels et suggérés à la ménagère par le carnet de comptes. Les dépenses d'habillement, obtenues par interview, sont par contre sous-estimées par rapport à l'enquête spéciale de 1953, l'erreur portant principalement sur les achats de lingerie et de bonneterie. Cette sous-estimation peut être attribuée à des oublis purs et simples, la période de trois mois à laquelle l'interview se rapportait étant probablement trop longue.

L'estimation des revenus des ménages par l'enquête pose un problème qui n'est pas facile à résoudre. En général, les ressources déclarées sont inférieures aux dépenses totales, dans une proportion qui varie peu d'un groupe de ménages à un autre. On peut cependant constater que l'écart est le plus élevé pour les ménages des catégories de revenu modeste et que, pour les ménages riches, les ressources l'emportent sur les dépenses. Il semble que les ressources déclarées correspondent le plus souvent à la dernière année du calendrier et non pas aux 12 derniers mois. Cela peut expliquer une faible partie de l'écart qu'on constate, mais ne suffit pas pour résoudre le problème. En effet, pour l'ensemble de l'échantillon le résultat est inacceptable : on ne peut pas admettre qu'en général et pour l'ensemble des ménages, les dépenses dépassent les ressources, impôts compris. Il faut donc conclure qu'en général les ressources sont nettement sous-estimées ; malheureusement l'absence de toutes spécification des ressources dans l'exploitation mécanographique des données exclut pour l'instant une analyse détaillée de ce phénomène. Quant à la variation de l'écart selon les catégories de revenu, elle correspond aux résultats obtenus dans toutes les enquêtes de ce genre ; son interprétation reste cependant incertaine.

§ 6 - Le questionnaire d'enquête

INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES ÉCONOMIQUES
29, quai Branly. Paris (VII')

COMITÉ NATIONAL
DE LA PRODUCTIVITÉ

CENTRE DE RECHERCHE
ET DE DOCUMENTATION
SUR LA CONSOMMATION

ÉTUDE

SUR LES DÉPENSES DES MÉNAGES

Re

Commune : ..... Enquête n° .....

N° N° du ménage

Département : .....

DB Rg C. C.

COMPOSITION DU MÉNAGE

Ages

PERSONNES VIVANT AU FOYER	SEXE	ÂGE	PROFESSION INDIVIDUELLE
1. Chef de ménage.....			
2.			
3.			
4.			
5.			
6.			
7.			
8.			

NT NA NE

CSC

N Ac

Ty

UC (1) UC (2)

CB

ATTENTION!

Les renseignements que vous fournirez seront considérés comme STRICTEMENT CONFIDENTIELS.

D'ailleurs, ni votre nom, ni votre adresse ne doivent figurer sur ce questionnaire. Ils ne seront donc connus que de l'enquêteur, et ce dernier est astreint par le secret professionnel à la discrétion la plus absolue.

Enquêteur : M. .....

Visa n° 56 615 du Secr. d'État aux Aff. éc. (I.N.S.E.E.).

J. U. 610211. [39836]

Semaine du ..... au .....

A. — Achats de denrées alimentaires et autoconsommation

NATURE	1er JOUR	2e JOUR	3e JOUR	4e JOUR	5e JOUR	6e JOUR	7e JOUR	*Ne rien écrire dans cette colonne*	
								ACHATS	AUTOCONSOMMATION
PRODUITS À BASE DE CÉRÉALES									
01. Pain									
02. Biscottes, biscuits, pain d'épices									
03. Pâtisserie (achetée chez le boulanger ou le pâtissier)									
04. Farine ordinaire									
05. Farines composées, semoules, flocons d'avoines, maizena, etc.									
06. Riz									
07. Pâtes alimentaires (ou à potage)									
VIANDES ET POISSONS									
08. Bœuf									
09. Veau									
10. Mouton									
11. Cheval									
12. Porc frais (non compris lard gras)									
13. Jambon									
14. Autre charcuterie									
15. Triperie									
16. Volaille									
17. Lapin, gibier									
18. Conserve de viande de charcuterie en boîte									
19. Poisson, crustacés, coquillages									
20. Conserves de poisson									

A. — Achats de denrées alimentaires et autoconsommation (*suite*)

NATURE	1er JOUR	2e JOUR	3e JOUR	4e JOUR	5e JOUR	6e JOUR	7e JOUR	*Ne rien écrire dans cette colonne*	
								ACHATS	AUTOCONSOMMATION
ŒUFS, LAIT, CORPS GRAS									
21. Œufs									
22. Lait frais									
23. Lait en poudre, condensé									
24. Fromage									
25. Petit suisse, yoghourt, crème fraîche, fromage blanc									
26. Beurre									
27. Margarine									
28. Huile									
29. Lard, saindoux									
30. Autres graisses animales									
31. Graisses végétales (Végétaline, cocose, etc.)									
LÉGUMES ET FRUITS									
32. Pommes de terre									
33. Légumes verts									
34. Racines									
35. Légumes secs									
36. Conserves de légumes									
37. Oranges, mandarines, bananes, citrons, pamplemousses									
38. Autres fruits frais									
39. Fruits secs ou en conserve (noix, amandes, marrons, dattes, pruneaux)									

A. — Achats de denrées alimentaires et autoconsommation (*fin*)

NATURE	1er JOUR	2e JOUR	3e JOUR	4e JOUR	5e JOUR	6e JOUR	7e JOUR	*Ne rien écrire dans cette colonne* — ACHATS	AUTOCONSOMMATION
BOISSONS									
40. Vin. . . . Vin ordinaire									
Appellation contrôlée.									
41. Bière									
42. Cidre									
43. Eau minérale, jus de fruits et autres boissons sans alcool . . .									
44. Alcool [1] (eau-de-vie, cognac) . .									
45. Liqueurs [1]									
46. Apéritif [1]									
PRODUITS DIVERS									
47. Café									
48. Thé, chicorée, malt									
49. Chocolat (en tablettes)									
50. Cacao, petits déjeuners									
51. Sucre									
52. Confitures, miel, compote [2] . . .									
53. Confiserie									
54. Potage « express », plats préparés.									
55. Sel, condiments, vinaigre									
56. Extraits de viande									
57. Glaces et crèmes glacées									
58. Glace à rafraîchir									
Repas payant pris à l'extérieur — Cantine — Nombre . .									
Cantine — Dépense . .									
Restaurant — Nombre . .									
Restaurant — Dépense . .									

[1] Non pris au café. — [2] Achats seulement.

B. — Autres dépenses courantes

NATURE	1er JOUR	2e JOUR	3e JOUR	4e JOUR	5e JOUR	6e JOUR	7e JOUR	*Ne rien écrire dans cette colonne*
ENTRETIEN (1)								
01. Savons								
02. Détersifs synthétiques (Omo, Mir, etc.)								
03. Lessive, eau de Javel								
04. Autres produits (détachant, cirage, encaustique, essence à briquet, etc.)								
MERCERIE								
05. Aiguilles, fil, boutons, etc								
06. **PHARMACIE** (2)								
07. **JOURNAUX, REVUES** (3)								
08. **FLEURS**								
SERVICES PERSONNELS								
09. Coiffeur								
10. Bains, douches								
11. Blanchisserie, laverie automatique								
TRANSPORTS (4)								
12. Réguliers pour le travail et l'école								
13. Petits voyages occasionnels (moins de 100 km aller-retour)								
14. Taxi								

(1) Produits et articles d'entretien du linge, des vêtements, de la maison.
(2) Produits d'USAGE COURANT achetés SANS ORDONNANCE et de ce fait non remboursés par la Sécurité sociale. Exemple : cachets d'aspirine, éther, alcool, etc.
(3) Achats au numéro (abonnements exclus).
(4) Compter les transports en commun seulement, pour tous les membres du ménage, transports non remboursés par l'employeur. Si les cartes de transports réguliers sont payées annuellement, faire l'estimation hebdomadaire.

3. — Autres dépenses courantes (*fin*)

NATURE	1er JOUR	2e JOUR	3e JOUR	4e JOUR	5e JOUR	6e JOUR	7e JOUR	*Ne rien écrire dans cette colonne*
LOISIRS								
15. Cinéma, théâtre								
16. Sports (spectacle sportif)								
17. Consommations prises au café [1]								
18. Billets de loterie (Loterie Nationale et autres)								
19. Argent de poche des enfants								
20. Tabac, cigarettes, allumettes								
DIVERS								
21. Timbres, télégrammes, frais postaux [2]								
22. Services domestiques [3] (femme de ménage, bonne, jardinier, chauffeur, etc.)								
23. Leçons (auto, danse, sport) [4]								
24. Dons (quêtes, charité)								
25. Autres dépenses diverses								

[1] Y compris les consommations prises dans les lieux publics (salon de thé, cinéma, kermesse, rue).
[2] Non compris les transferts d'argent.
[3] Y compris la cotisation patronale à la Sécurité sociale.
[4] Non compris les leçons particulières données aux enfants.

C. Dépenses mensuelles, trimestrielles, annuelles

(Interview)

Introduction

1. Avez-vous du mal à boucler votre budget? .....

.....

2. Comment y parvenez-vous?
(Rayer les mentions inutiles.)

Difficilement.....	1	
A peu près.....	2	☐
Facilement	3	
Sans réponse.....	×	

3. Qu'est-ce qui paraît peser le plus fortement sur votre budget? .....

.....

4. Votre niveau de vie s'est-il amélioré depuis un an?
(Rayer les mentions inutiles.)

Oui.....	1	
Est resté stationnaire...	2	☐
A diminué	3	
Sans réponse.....	×	

LOGEMENT

Les questions 1 à 8 ne concernent que le logement dans lequel le ménage habite actuellement.)

1. Payez-vous un loyer ou une redevance d'occupation?
(Rayer les mentions inutiles.)

Oui.....	1	
Non (logé gratuitement).	2	☐
Non (propriétaire).....	3	
Sans réponse.....	×	

a. Pour les personnes ayant répondu « oui » à la question précédente :

2. Quel est le montant, sur votre dernière quittance, du loyer principal ....
(Charges et prestations non comprises.)

Le loyer indiqué ci-dessus correspond à une période de ..... mois ☐☐☐

3. Indiquez, éventuellement, le montant **annuel** de l'allocation logement..... .....

4. Indiquez le montant des payements effectués pendant les **12 derniers mois** (ou sur une période récente couvrant une année entière) au titre :

des charges [1]..... .....

des prestations [2]..... .....

5. Recettes de sous-location dans les **12 derniers mois**..... .....

b. Questions à poser à toutes les personnes, qu'elles soient propriétaires ou locataires de leur logement

6. Indiquez le montant pour les **12 derniers mois** (ou pour une période récente, couvrant une année entière) :

	MONTANT EN FRANCS	OBSERVATIONS [3]
De la contribution mobilière.....	.....	.....
Des réparations et aménagements d'intérieur [*].....	.....	.....
Des réparations et aménagements d'extérieur [*].....	.....	.....
De la contribution foncière (le cas échéant).....	.....	.....
Des primes d'assurances. Immeuble (incendie).....	.....	.....
Des primes d'assurances. Mobilier (vol et incendie)...	.....	.....
Des charges (co-propriétaires).....	.....	.....

7. Avez-vous fait construire ce logement [4] en bénéficiant d'un prêt de l'État ou d'un organisme de crédit?

Oui... 1 Non... 2
(*Rayer la mention inutile.*)

Si oui, avez-vous encore à payer des **annuités** de remboursement?

Oui... 1 Non... 2
(*Rayer la mention inutile.*)

Si oui, quel est le montant de la dernière **annuité** de remboursement que vous avez payée?..... .....

---

[1] Charges : taxes locatives, eau froide, ascenceur, etc.
[2] Prestations : chauffage central, eau chaude.
[3] Mentionner dans cette colonne la période correspondant au montant déclaré (1 mois, 3 mois, etc.) lorsqu'il n'aura pas été possible de l'obtenir pour 12 mois consécutifs.
[4] Il s'agit de constructions postérieures à 1948.

Pendant combien d'années avez-vous encore à payer? .... ans | | |

8. Quel est le montant de votre dernière quittance :

de gaz .... : ..... fr., correspondant à une période de ..... **mois**; | | | |

d'électricité . : ..... fr., correspondant à une période de ..... mois; | | | |

d'eau froide. : ..... fr., correspondant à une période de ..... mois; | | | |

de téléphone : ..... fr., correspondant à une période de ..... mois. | | | |

9. Indiquez, pour les **trois derniers mois**, vos achats de combustible, en quantité et en valeur :

	QUANTITÉ ACHETÉE	UNITÉ	DÉPENSE (EN FRANCS)
1. Boulets, briquettes.....			
2. Cokes.....			
3. Charbons (classés et grains).....			
4. Butane.....			
5. Bois.....			
6. Fuel (mazout).....			
Essence C (pour poêle catalytique).....			

10. Si vous avez changé de logement dans les **12 derniers mois**, quels ont été vos frais de déménagement? .... | | | |

11. Êtes-vous locataire ou propriétaire d'autres logements que celui dans lequel vous habitez actuellement?

Oui... 1 Non... 2 | |

(Rayer la mention inutile.)

12. Si oui, indiquez, pour ce ou ces logements le montant total des dépenses **annuelles** (réparations, loyer, impôts, etc.)..... ..... | | | |

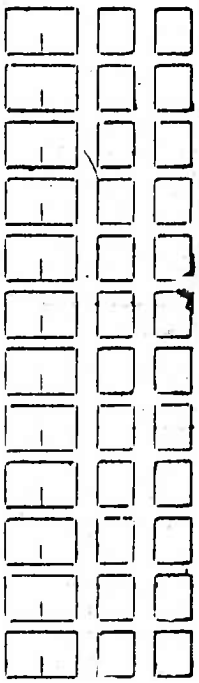
Indiquez éventuellement, le montant **annuel** des recettes correspondantes (sous-location, location, etc.)..... ..... | | | |

II. ACHATS DE VÉHICULES, DE BIENS D'ÉQUIPEMENT MÉNAGER AU COURS DES DOUZE DERNIERS MOIS

1. (L'enquêteur montrera la liste des véhicules, appareils ou meubles donnée dans l'annexe de l'instruction aux enquêteurs, et remplira le tableau ci-dessous.)

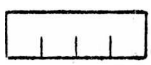
TABLEAU I

VÉHICULE, APPAREIL OU MEUBLE (*A indiquer en clair*)	NEUF OU OCCASION (*Indiquer N ou O*)	ACHAT COMPTANT OU À TEMPÉRAMENT (*Indiquer C ou T*)	PRIX AU COMPTANT	POUR LES ACHATS À TEMPÉRAMENT : Montant total des mensualités versées au cours des 12 derniers mois.



2. Avez-vous eu encore à payer au cours des **12 derniers mois** des mensualités correspondant à des achats effectués depuis plus d'un an ? Oui . . . 1 Non . . . 2
(*Rayer la mention inutile.*)

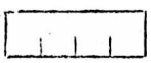
3. Si oui, indiquez le montant total des versements effectués au cours des **12 derniers mois** . . . . . .



4. Parmi les véhicules, meubles ou appareils de la liste, marquez ceux que vous avez vendus d'occasion au cours des **12 derniers mois** et la recette correspondante :

TABLEAU II

VÉHICULES, APPAREILS OU MEUBLES VENDUS	RECETTE CORRESPONDANTE
TOTAL	



5. Indiquez au tableau III, les dépenses de réparation et d'entretien de biens durables (sauf véhicules individuels) effectuées au cours des **12 derniers mois :**

TABLEAU III

APPAREILS OU MEUBLES	DÉPENSE DE RÉPARATION OU D'ENTRETIEN
TOTAL.....	

III. MOYENS DE TRANSPORTS PERSONNELS

L'enquêteur posera la question : «Possédez-vous une auto, une moto, un scooter, un cyclomoteur ?». Dans le cas d'une réponse affirmative, il utilisera pour répondre aux questions du tableau ci-après, autant de colonnes qu'il y a de véhicules dans le ménage. Il indiquera clairement en tête de la colonne : «Auto» ou «Moto», etc.

Ne rien inscrire dans les cases de chiffrement.

01. *Indiquez en titre de la colonne :* auto, moto, scooter ou cyclomoteur.				
02. Puissance fiscale et marque..				
03. Année de fabrication.....				
04. Acheté neuf (N) ou d'occasion (O).....				
05. Si occasion, année d'achat...				
06. Nombre de km parcourus par an.....				
07. Prime d'assurance **annuelle**..				
08. Frais de garage **annuels**				
09. **Frais annuels** — Frais de réparation.....				
10. **Frais annuels** — Achats de pneumatiques, accessoires et fournitures...				
11. **Frais au cours du dernier mois** — Essence.....				
12. **Frais au cours du dernier mois** — Huile, vidange, graissages....				
11 *bis.* *Ne rien écrire sur ces lignes* — Ess....				
12 *bis.* *Ne rien écrire sur ces lignes* — H.....				

13. Si *vous êtes « non salarié »*, indiquez approximativement la proportion (en %) des dépenses 7 à 12 correspondant à vos frais professionnels : .....

14. Si *vous êtes « salarié »*, indiquez le montant **annuel** des remboursements de vos frais de voiture ou moto par votre employeur : .....

IV. ACHATS DE VÊTEMENTS, CHAUSSURES, TEXTILES

(Indiquez la dépense en francs pour chaque groupe, au cours des **trois derniers mois**)

ARTICLES	ADULTES		ENFANTS DE MOINS DE 14 ans	ENSEMBLE
	HOMMES	FEMMES		
1. Vêtements principaux (*) de confection ou de demi-mesures, fourrures				
2. Vêtements principaux sur mesures, tissu non fourni par le client				
3. Dépenses de façon (couturière, tailleur, etc.).				
4. Vêtements de travail, tabliers				
5. Lingerie, bonneterie, layette (*).				
6. Chaussures et pantoufles				
7. Accessoires divers du vêtement (*)				
8. *a.* Laine à tricoter				
b. Tissus de confection				
c. Tissus d'ameublement				
d. Linge de maison (*)				
e. Couvertures				

V. ACHATS D'AUTRES BIENS DURABLES, AU COURS DU DERNIER MOIS

Qu'avez-vous dépensé, au cours du **dernier mois** dans les groupes suivants?

	DÉPENSE (EN FRANCS)		DÉPENSE (EN FRANCS)
01. Petits appareils électriques ou à gaz (*)		08. Livres, disques, jouets	
02. Quincaillerie (*)		09. Articles de sports et de voyage [5]	
03. Articles d'entretien [1]		10. Optique, photo [6]	
04. Produits d'hygiène [2] sauf savon		11. Bijoux, montres	
05. Articles de toilette [3]		12. Teinturerie	
06. Vaisselle et articles divers [4]		13. Ressemelage et autres dépenses chez le cordonnier	
07. Encre, papeterie, fournitures scolaires		14. Graines, plants, produits pour le jardin [7]	

[1] Brosse, balai, éponge, etc.
[2] Produits destinés aux soins corporels : eau de Cologne, dentifrice, produits de beauté, etc.
[3] Brosse à dents, lames de rasoir, etc.
[4] Seau, bassine, lessiveuse, etc.
[5] Valise, articles de pêche, de camping, de sport, etc.
[6] Optique, matériel de photo et de cinéma, etc.
[7] Jardin exclusivement. — Les dépenses professionnelles des exploitants agricoles ne doivent pas apparaître ici.

VI. DÉPENSES DE SANTÉ AU COURS DU DERNIER MOIS

Indiquez dans le tableau ci-dessous, vos dépenses de santé et éventuellement le montant des remboursements de la Sécurité sociale, des caisses de Mutuelle, etc.

	DÉPENSES	REMBOURSEMENTS
Au cours du **dernier mois.** Médecin		
Pharmacien		
Dentiste		
Au cours des **12 derniers mois :** Hôpital, clinique, cure		
Ne rien écrire sur ces 2 lignes		

D | | | |

R | | | |

VII. VACANCES

Indiquez, dans le tableau ci-dessous, pour chacune des personnes du ménage, le montant des dépenses de vacances, au cours des **12 derniers mois** [1]

PERSONNES DU MÉNAGE (LIEN AVEC LE CHEF DE MÉNAGE)	VOYAGE (TRANSPORTS en commun)	LOCATION DE MAISON	HÔTEL	COLONIES DE VACANCES	TOTAL
1. Chef de ménage . .					
2.					
3.					
4.					
5.					
6.					
ENSEMBLE du ménage.					

VIII. VOYAGES AU COURS DES 12 DERNIERS MOIS

Indiquez pour les voyages de plus de 100 km (aller-retour), autres que ceux signalés dans la rubrique « Vacances » les *frais de transports en commun* pour l'ensemble du ménage : .....

IX. DIVERS

Indiquez pour les **12 derniers mois** (ou pour une période récente couvrant une année entière) le montant ;

1. De vos impôts *personnels* [2] (autres que la contribution mobilière et foncière) .....
2. Des dépenses d'éducation et enseignement (frais scolaires, études surveillées, cours du soir, internat, etc.) .....
3. De vos cotisations (mutuelles, syndicats, associations sportives et autres, sauf Sécurité sociale) .....

(1) On utilisera les colonnes « Total » ou « Ensemble » lorsque la ventilation par poste ou par personne ne sera pas possible (dans le cas de voyage à prix forfaitaire par exemple).

(2) Surtaxe progressive, taxe proportionnelle, taxe radio.

4. De vos abonnements (journaux et périodiques, bibliothèque, etc.)..... .....

5. De vos frais de justice..... .....

6. Des pensions diverses (pensions alimentaires, etc.)..... .....

7. Autres dépenses (précisez).... .....

.....

....

X. RESSOURCES DU MÉNAGE AU COURS DES 12 DERNIERS MOIS (ou de la dernière année du calendrier)

	CHEF DE MÉNAGE	CONJOINT	AUTRES PERSONNES du ménage
A. REVENU DU TRAVAIL			
1. Salaires, traitements (y compris activités d'appoint).....			
2. Revenus agricoles (indiquez *grosso modo* un revenu moyen).....			
3. Revenus commerciaux.....			
4. Revenus des chefs d'entreprises industrielle ou artisanale.....			
5. Revenus des professions libérales.....			
B. REVENUS DU CAPITAL			
1. Mobilier.....			
2. Immobilier.....			
C. AUTRES REVENUS			
1. Allocations familiales.....			
2. Pensions, retraites, rentes, etc.....			
3. Autres (précisez) :			
.....			
D. AVANTAGES EN NATURE (*En faire l'estimation*)			
Nourriture, logement, chauffage, etc.....			
TOTAL.....			

TOT. : .....

Indiquez si votre revenu global, par rapport à l'année précédente
(*Rayer les mentions inutiles*)

a augmenté..... 1
est resté stationnaire.... 2
a diminué..... 3

Observations de l'enquêteur

Imprimerie Nationale J. U. 61021 f

SECTION 4

NOTES BIBLIOGRAPHIQUES

§ 1 - Le développement des enquêtes de consommation en France depuis 1945

Cette annexe présente une liste commentée des plus importantes enquêtes de consommation réalisées en France entre 1945 et 1960 par des organismes publics. Cette liste n'est pas complète mais elle permet, par de brèves notes descriptives, d'apprécier le développement de la méthode de ces enquêtes. La première partie de la liste décrit 16 enquêtes générales, dont l'objet est de saisir le budget des ménages interrogés ; une place toute particulière a été faite à l'enquête de 1948 (N° 5) dont les méthodes ont été reprises pratiquement dans toutes les enquêtes de ce genre jusqu'en 1952. La deuxième partie décrit plus rapidement 8 enquêtes sur la consommation de produits particuliers (textiles, produits laitiers et viande, etc...).

On trouvera des éléments bibliographiques supplémentaires dans trois articles :

- Premiers résultats concernant les dépenses des familles de salariés de la région parisienne au printemps 1948, Bulletin de la Statistique Générale de la France, supplément trimestriel, Janvier-Mars 1949, page 37.

- Les enquêtes alimentaires en France, par A. BOUCHE et H. DESROCHE. Bulletin de l'Institut National d'Hygiène, Octobre-Décembre 1953, page 705. Pour la période 1900-1945 cette étude se réfère également à toutes les enquêtes "budgets de familles".

- Liste des enquêtes de consommation, Consommation, Avril-Juin 1959, page 96. Cette liste sommaire des enquêtes réalisées depuis 1948 comprend les listes des organismes officiels ainsi qu'un nombre d'enquêtes plus spécialisées réalisées par des organismes d'étude de marché.

ABREVIATIONS DES NOMS D'INSTITUTS ET DE PUBLICATIONS

I.N.S.E.E.	: Institut National de la Statistique et des Etudes Economiques.
I.N.E.D.	: Institut National d'Etudes Démographiques.
I.N.H.	: Institut National d'Hygiène.
B.S.G.F.	: Bulletin de la Statistique Générale de la France (I.N.S.E.E.)
B.M.S.	: Bulletin Mensuel de Statistique (I.N.S.E.E.)
B.H.S.	: Bulletin Hebdomadaire de Statistique (I.N.S.E.E.)
Suppl. Trim.	: Supplément Trimestriel.

ENQUETES "BUDGETS DE FAMILLE"

1. Dépenses et consommation des familles de salariés de la région parisienne. Printemps et Automne 1946. B.S.G.F., Avril-Juin 1948. Enquête de l'I.N.S.E.E.

Aucune condition n'était imposée quant à la composition de la famille. Le niveau social devait correspondre au maximum à celui d'ouvrier qualifié ou d'employé subalterne.

L'échantillon - 1.502 ménages en Mars-Avril, 1077 ménages en Novembre - était désigné en partie par tirage au hasard dans le fichier de la Caisse Interdépartementale de la Sécurité Sociale, qui est un fichier de personnes et non de ménages, en partie choisi parmi les connaissances des enquêteurs. Les familles touchées appartenaient surtout à la partie la plus favorisée de la population que l'on se proposait d'étudier.

Les dépenses courantes, et notamment les dépenses alimentaires, étaient enregistrées sur carnet de comptes pendant quinze jours. Les autres postes du budget : congés annuels, acquisition de biens durables, faisant l'objet d'une interview sur un an, semblent avoir été saisis de manière beaucoup moins satisfaisante (les résultats de l'enquête de printemps n'ont d'ailleurs pas été publiés sur ce point). En ce qui concerne l'alimentation, on notait non seulement les dépenses mais également les quantités (ce qui permit d'établir un grossier bilan alimentaire) et les prix.

2. Les dépenses de 1080 familles en mars et juin 1946. Population avril-juin 1947. Enquête de l'I.N.E.D.

Enquête dans la région parisienne sans limitation de la composition de la famille ou de milieu social.

L'échantillon était recruté parmi les connaissances des enquêteurs "auxquels des instructions précises étaient données".

Les techniques d'enregistrement sont peu différentes de l'enquête I.N.S.E.E. Sur le carnet des comptes étaient enregistrés dépenses, quantités et prix.

3. Les dépenses de 939 familles de militaires en novembre 1946. Population, juillet-septembre 1947. Enquête de l'I.N.E.D.

L'échantillon consiste en familles d'officiers et sous-officiers ayant accepté de répondre à l'enquête.

Le carnet de comptes utilisé est le même qu'au numéro 2 ; une page spéciale du carnet reproduisait la feuille de solde mensuelle des militaires.

4. Dépenses et ration alimentaire en janvier-février 1948. Population, avril-juin 1948. Enquête de l'I.N.E.D.

Tous les milieux sociaux de la région parisienne ; en revanche les célibataires étaient éliminés ainsi que certaines familles de structure complexe. 774 questionnaires effectivement utilisés.

Les techniques d'échantillonnage et d'enregistrement des données étaient les mêmes qu'en 2. On s'efforçait d'obtenir une déclaration des ressources familiales.

5. Dépenses et consommation de familles de salariés de la région parisienne au printemps 1948, B.S.G.F., janvier-mars 1949 et juillet-septembre 1949, et à l'automne 1948, B.S.G.F., octobre-décembre 1949 et janvier-mars 1950. Enquête réalisée par l'I.N.S.E.E. et l'I.N.H.

Cette enquête est très importante, les techniques d'enregistrement ayant été conservées dans la suite jusqu'en 1952. Par ailleurs, les résultats de l'enquête ont été utilisés conjointement avec les résultats de l'enquête n° 8, pour déterminer les coefficients de pondération de l'indice des prix à la consommation familiale à Paris Base 100 en 1949 (Cf. sup. B.M.S., janvier-mars 1951). A partir de l'enquête d'automne, une étude a été réalisée sur le marché des fruits et légumes frais à Paris (sup. B.M.S., Avril-Juin 1950).

L'enquête était limitée à la Seine, aux familles de salariés de situation inférieure à celle de contremaître ou de comptable, dont la composition était conforme à l'un des quatre types suivants : HF, HFE, HFEE, HFEEE, (H = homme, F = femme, E = enfant de moins de 16 ans).

L'échantillon étant désigné par tirage au sort dans les feuilles de ménages du recensement de 1946, l'enquête de 1948 est la première enquête française sur les budgets de familles réalisée auprès d'un échantillon aléatoire. Il y eut environ 40 % de refus ; les enquêteurs procédaient à des remplacements en utilisant pour cela des listes complémentaires. L'échantillon effectivement utilisé fut de 831 ménages au printemps, et de 758 en automne (dont 348 déjà interrogés au printemps avaient été volontaires pour être interrogés une deuxième fois). L'auteur de l'étude pense que les refus n'ont pas entraîné de distorsion notable mais que les volontaires interrogés à nouveau en automne appartenaient en moyenne aux familles les plus aisées, d'où un biais non négligeable (surestimation des dépenses de l'ordre de 5 %).

Les enquêtes se sont réparties respectivement assez également entre le 5 avril et le 11 juillet, et entre le 11 octobre et le 19 décembre - périodes de relative stabilité des prix.

Le questionnaire s'adressait au ménage entier : aucune question ne tentait de faire préciser les consommations de chacun des membres de la famille.

Un carnet de comptes recevait mention sur une semaine des acquisitions de denrées alimentaires et des variations du stock familial de produits alimentaires. Pour chaque denrée, la désignation précise était demandée ainsi que la quantité acquise et le coût réel de cette quantité (compte tenu du cours du marché parallèle). Au questionnaire était jointe une feuille sur laquelle les menus devaient être consignés jour par jour : on notait pour chaque repas le nombre de rationnaires, les repas pris au domicile par des invités et les repas pris à l'extérieur par des membres de la famille.

Pour les autres dépenses courantes l'enquêteur devait, à partir des dépenses réalisées pendant l'enquête, procéder à une évaluation des dépenses hebdomadaires moyennes de la famille.

Enfin, les dépenses dont la périodicité dépasse largement la semaine : habillement, achat de biens durables, loyer, dépenses médicales et pharmaceutiques... étaient obtenues par interview sur un an. Toutefois, le détail était demandé par trimestre.

L'enquêteur devait le plus possible se reporter aux documents écrits, livres de comptes, quittances ... Il devait noter sur une feuille à part ses impressions sur la qualité des renseignements fournis, ainsi que sur les circonstances particulières de nature à agir sur le comportement de la famille.

6. Les dépenses et l'alimentation des ménages de salariés: Marseille, Lille et Saint-Etienne pendant l'automne de 1948. Suppl. trim. B.M.S., janvier-mars 1950. Enquête de l'I.N.S.E.E. et de l'I.N.H.

Les mêmes méthodes qu'au numéro 5. Nombre de questionnaires remplis : Marseille 161, Lille 220 (presqu'exclusivement des ménages ouvriers), Saint-Etienne 117.

7. Les dépenses et l'alimentation des familles de mineurs à Saint-Etienne au printemps 1949. Suppl. trim. B.M.S., janvier-mars 1950. Enquête de l'I.N.S.E.E. et de l'I.N.H.

Echantillon aléatoire tiré dans le fichier du personnel d'un seul puits de mine - fichier de personnes et non de ménages. On avait exclu les cadres et les familles d'une personne. Taux de refus 20 % ; 184 questionnaires dépouillés.

Techniques d'enregistrement comme au n° 5 ; la direction de la mine avait bien voulu communiquer le salaire des mineurs.

8. Les dépenses de familles de la région parisienne, hiver 1949-1950 et automne 1950. Suppl. trim. B.M.S. ; janvier-mars 1953. Enquête de l'I.N.S.E.E. et de l'I.N.H.

Mêmes méthodes qu'au numéro 5. Questionnaires dépouillés : 601 (hiver 1949-1950) et 411 (automne 1950).

L'auteur de l'article conclut "qu'il est difficile de porter un jugement définitif, à partir de ces résultats, sur l'évolution du niveau de vie des salariés de la région parisienne".

9. Les budgets de familles de salariés dans quelques grandes villes de province, hiver 1949-1950 et automne 1950. Suppl. trim. B.M.S. ; janvier-mars 1953. Enquête de l'I.N.S.E.E. et de l'I.N.H.

Enquête réalisée parallèlement à la précédente et suivant des méthodes identiques. Questionnaires disponibles pour l'hiver 1949-1950 : Lyon 234, Lille 249, Marseille 113, et pour l'automne 1950 : Bordeaux 105, Marseille 158, Rouen 199.

L'auteur conclut que "les enquêtes, effectuées à des époques différentes, ont enregistré de ce fait certaines variations saisonnières en matière de dépenses et de consommations qui viennent se superposer aux différences purement géographiques, ces particularités rendant difficiles les comparaisons d'une ville à une autre".

10. Une enquête sur les budgets de familles de salariés à Strasbourg au printemps 1950. B.H.S., n° 146 du 17 février 1951.

Enquête du même type que la préoédente, réalisée auprès de 187 ménages.

11. Une enquête par sondage sur l'alimentation et les dépenses des ménages d'ouvriers mineurs du groupe de Lens en juin-juillet 1951. Suppl. trim. B.M.S. : octobre-décembre 1952. Enquête de l'I.N.S.E.E. et de l'I.N.H.

L'échantillon comprend 370 ménages. Aucune limitation n'a été introduite quant à la composition de la famille: des ménages, de compositions très diverses, figurent dans l'échantillon.

L'échantillon était désigné par tirage au sort dans les fichiers du personnel des mines ; la probabilité pour les ménages d'être retenus pour l'enquête était proportionnelle au nombre de mineurs qu'ils comprenaient On a supprimé cette erreur systématique lors de l'exploitation, en donnant des poids différents aux divers types de ménages.

L'enquête ayant été réalisée en période de hausse rapide des prix, toutes les dépenses ont été exprimées en francs de juillet 1952.

Contrairement à la méthode employée pour les enquêtes précédentes, on n'a pas appliqué de majorations systématiques sur les divers postes du tableau des dépenses annuelles.

Des questions sur les revenus des divers membres du ménage ont été posées. Malgré les circonstances favorables (accord des syndicats et de la direction, qualité des enquêteurs), cette tentative fut un demi-échec : 70 % de dépenses environ sont couvertes par les revenus déclarés. En utilisant les renseignements donnés par le Bureau de Documentation Minière on a toutefois pu reconstituer une masse des revenus voisine de la masse des dépenses.

12. Les dépenses et les consommations des ménages à Paris, Rennes et dans 17 grandes villes de province, novembre-décembre 1951. Suppl. trim. B.M.S. ; octobre-décembre 1953, et études statistiques, avril-juin 1957. Enquête de l'I.N.S.E.E.

Cette enquête est la première enquête française sur les dépenses des ménages réalisée auprès d'un échantillon aléatoire de ménages de toutes catégories sociales et de toutes compositions. Les fichiers des feuilles de ménages du recensement de 1946 étant trop anciens, le tirage a eu lieu dans les calepins des Contributions Directes avec le taux de sondage suivant :

Agglomération parisienne.....	1/2000
Agglomération de Lyon et de Marseille.....	1/1000
Rennes.....	1/100
Autres agglomérations.....	1/500

Les refus ont été nombreux, notamment dans les plus grandes villes (spécialement à Paris) et dans les catégories de population les plus aisées. Le dépouillement a porté sur 281 questionnaires à Paris, 349 questionnaires à Rennes et 1949 questionnaires pour les autres villes.

Les techniques d'enregistrement étaient les mêmes qu'en 1948. Il n'y a pas eu de difficultés spéciales à se servir d'un même formulaire pour des ménages de conditions sociales différentes : toutefois, il y aurait eu intérêt à développer certaines rubriques : transports, loisirs.

Les résultats ont été exprimés en francs de décembre 1951 (l'année précédant l'enquête ayant connue une inflation permanente). Les résultats sont établis séparément pour Paris, Rennes et pour l'ensemble des 17 villes de province convenablement pondérées ; certains résultats sont établis en fonction du niveau de vie, lui-même obtenu en divisant la dépense totale par le nombre d'unités de consommation ; on avait constitué dans chaque ville cinq classes, de niveaux de vie décroissants, d'effectifs à peu près égaux. Un essai de relevé direct des revenus avait échoué totalement.

13. Une enquête sur les dépenses des ménages d'exploitants agricoles en 1952. Suppl. trim. B.M.S. ; juillet-septembre 1954, B.H.S. n° 243 du 27 décembre 1952 et n° 267 du 13 juin 1953. Enquête de l'I.N.S.E.E.

Deux enquêtes-pilotes ont été faites dans le milieu rural ; l'une dans la Vienne (100 exploitants agricoles, 29 salariés agricoles, 70 autres ruraux) et l'autre dans le bas Rhin (123 exploitants agricoles et 44 ruraux non-agricoles). Ces échantillons ont été désignés par tirage au sort. On trouve les résultats de ces enquêtes-pilotes dans les numéros B.H.S. indiqués ci-dessus.

L'enquête générale, en revanche, a été réalisée auprès d'un échantillon recruté par les maisons familiales d'apprentissage rural, complété par l'I.N.S.E.E. et contrôlé suivant la superficie. L'échantillon dut d'ailleurs être redressé, il contenait trop peu de petites exploitations.

652 questionnaires ont été recueillis ; ils sont géographiquement assez mal répartis. (on n'a pas pu donner de résultats par région pour les régions de sondage 1, 2, 4 et 6).

Les consommations étaient enregistrées sur carnet de comptes sur une semaine. On distinguait : les achats (poids et dépenses) et l'autoconsommation (poids et valeur estimés par l'enquêteur au prix du marché local). Les menus, nombre de rationnaires, variations de stocks au cours de la semaine étaient soigneusement notés. Les autres dépenses courantes étaient enregistrées sur une quinzaine : semaine de l'enquête et semaine précédente. Enfin, les dépenses d'habillement, logement, santé... et les dépenses d'exploitation étaient enregistrées sur un an par interview.

La méthode des quotas, jointe à la qualité des enquêteurs bien introduits dans le monde rural, a permis d'obtenir des questionnaires fort bien remplis.

14. Enquêtes sur les dépenses des ménages à Paris et dans 17 grandes villes de province en 1953. Etudes Statistiques, avril-juin 1957. Enquête de l'I.N.S.E.E.

Cette enquête était limitée aux villes sièges des Directions Régionales de l'I.N.S.E.E., aux catégories socio-professionnelles d'ouvriers, employés et cadres moyens, et aux ménages de plus d'une personne.

L'échantillon - 1.695 ménages à Paris et 3.241 ménages en province - a été choisi par la méthode des quotas. Les contrôles portaient sur la catégorie socio-professionnelle du chef de ménage et sur l'effectif du ménage. Les résultats ont été ensuite redressés, notamment en fonction du nombre de salariés par ménage.

L'enquête portait sur l'année 1953 dans son ensemble, l'échantillon étant renouvelé chaque mois. A Paris, l'enquête s'est déroulée à partir de mai 1953 et s'est prolongée en 1954 et plus loin ; en province, l'enquête a été interrompue au mois d'août.

Les dépenses étaient enregistrées entièrement par la méthode du carnet de comptes sur un mois.

L'enquête n'a pas fait l'objet d'une publication d'ensemble : toutefois, les dépenses alimentaires ont servi à des calculs d'élasticité publiés dans "Etudes Statistiques" Avril-Juin 1957.

15. Les dépenses des ménages d'ouvriers et employés en 1954 dans l'agglomération parisienne. B.H.S. n° 424 du 16 juin 1956. Enquête de l'I.N.S.E.E.

Cette enquête continue la précédente : elle concerne les ménages d'ouvriers et employés résidant dans la Seine, célibataires exclus. L'enquête a été réalisée auprès d'un échantillon de volontaires contrôlé par la méthode des quotas. Elle n'a pas donné d'aussi bons résultats que l'enquête en milieu rural, n° 13., dans laquelle on disposait de très bons enquêteurs.

Le questionnaire utilisé se présentait sous la forme d'un carnet de comptes mensuel sur lequel devaient être notés tous les achats ou débours. (on ne faisait donc pas appel à la mémoire).

L'échantillon était renouvelé chaque mois.

Au total 1.509 carnets mensuels ont été remplis.

Cette enquête a servi au calcul des coefficients de pondération de l'indice des prix de détail des 250 articles dans l'agglomération parisienne.

16. L'enquête sur les budgets familiaux des ouvriers de la sidérurgie, des mines de fer et de charbon de la C.E.C.A. : mai 1956-mai 1957. B.H.S. n° 550 du 29 novembre 1958. Enquête de la Division des Statistiques de la Haute Autorité de la Communauté Européenne du Charbon et de l'Acier, avec la collaboration des instituts nationaux de la Statistique : en France réalisée par l'I.N.S.E.E.

Le domaine étudié était très étroitement limité. Les familles interrogées appartenaient au milieu des ouvriers de la sidérurgie, des mines de fer et de charbon, de plus on imposait aux familles de présenter un certain nombre de critères. Ces critères sont les suivants :

- famille de 4 personnes (les conjoints et deux enfants à charge de moins de 13 ou 14 ans) ;

- seul le chef de famille travaillait ;

- le salaire du chef de famille devait se situer à l'intérieur d'une marge de dispersion de 15 à 20 % autour du salaire moyen du groupe auquel il appartenait ;

- l'ouvrier devait avoir la nationalité du pays dans lequel il travaillait ;

- l'enquête était limitée aux familles dont le chef avait la possibilité de rentrer chaque soir au foyer familial.

Les familles interrogées étaient volontaires, l'échantillon utilisé (440 pour la France) n'étaient donc pas désignées par le sort.

Les conditions d'enregistrement des données étaient particulièrement favorables : les familles volontaires : (largement rémunérées pour la première fois dans une enquête française) ayant tenu un carnet de comptes pendant un an. Il n'a donc été en rien fait appel à la mémoire.

Le caractère continu de l'étude, joint à la connaissance directe des ressources des familles permettaient de très sérieux contrôles de vraisemblance. Pour toutes ces raisons ont peut affirmer que les réponses individuelles ont été beaucoup plus exactes que dans toute autre enquête de même nature réalisée en France.

ENQUETES SPECIALES

17. Une enquête par sondage sur le marché des textiles en février 1948. B.S.G.F. ; juillet-septembre 1948. Enquête de l'I.N.S.E.E.

L'étude était consacrée aux achats de lingerie, chemiserie, tissus, vêtements, linge de maison, laine à tricoter, toile cirée, articles de layette.

Le domaine de l'étude comprenait l'ensemble des ménages français. L'échantillon de 4.086 ménages a été prélevé par tirage au sort dans le fichier des feuilles de ménages du recensement de 1946. Les refus ayant été fort peu nombreux (ils étaient remplacés par les enquêteurs), il s'agit donc d'un échantillon vraiment représentatif.

La méthode d'enregistrement des données (souvenirs sur 7 mois) sans effort pour répartir les achats entre les différents membres du ménage, était moins satisfaisante, bien qu'une liste détaillée des principaux articles ait été proposée comme aide-mémoire. Notons enfin que seules les quantités (nombre d'articles) et non les dépenses étaient enregistrées.

18. Une enquête par sondage sur le ravitaillement, décembre 1948, B.S.G.F., avril-juin 1949. Enquête de l'I.N.S.E.E.

L'étude portait sur une partie de l'alimentation seulement : viande, charcuterie, triperie, volaille, gibier, poisson, oeufs ; vin, bière et cidre, à l'exclusion des dépenses réalisées dans les restaurants, cantines, mais y compris les produits du jardin, de l'exploitation.

Le domaine d'étude était constitué par l'ensemble des ménages français. L'échantillon de 4.237 ménages était prélevé par tirage au sort dans le fichier des feuilles de ménages du recensement de 1946 : il était fort dispersé (335 communes) et les échecs ont été peu importants : refus 51, absences de longue durée 71 (ils étaient remplacés par les enquêteurs).

Les consommations de viande, poisson, oeufs, étaient enregistrées par interview portant sur une semaine, on enregistrait les noms des morceaux, la dépense et le poids. En ce qui concernait les boissons on enregistrait les consommations hebdomadaires considérées comme habituelles.

19. Enquête sur les dépenses d'habillement des Français en 1953. Brochure de l'I.N.S.E.E., 1955.

L'étude porte sur les achats de vêtements, articles de lingerie et bonneterie, tissus, laine à tricoter et layette. Il est donc question uniquement des articles achetés dans le commerce et l'on a exclu du champ de l'enquête : les tissus d'ameublement, linge de maison, chapeaux, accessoires du vêtement. Le champ de l'enquête était constitué par l'ensemble des ménages vivant en France métropolitaine (à l'exception de la Corse).

L'échantillon a été prélevé par tirage au sort dans le fichier déjà ancien des feuilles de ménages du recensement de 1946.

Quatre enquêtes ont été réalisées :

1ère Enquête : du 13 au 23 avril (pour enregistrer les achats du premier trimestre) :

2ème Enquête : du 15 au 27 juin 1953 (pour enregistrer les achats du deuxième trimestre) :

3ème Enquête : du 5 au 17 octobre 1953 (pour enregistrer les achats du troisième trimestre) :

4ème Enquête : du 4 au 16 janvier 1954 (pour enregistrer les achats du quatrième trimestre).

Chacune des quatre enquêtes portait sur un échantillon d'environ 5.500 ménages répartis dans 250 communes. L'échantillon était complètement renouvelé chaque trimestre. 83 % des ménages interrogés ont accepté de répondre. L'échantillon a été redressé par un procédé voisin de celui qui a été employé dans l'enquête sur la dépense des ménages de 1956.

Les données étaient enregistrées par interview faisant appel à la mémoire sur un trimestre. Un questionnaire individuel était rempli pour chaque membre du ménage âgé de plus de 18 mois : on y portait les achats de vêtements, lingerie, bonneterie (dépenses et quantités) mais il n'était pas imposé aux enquêteurs d'interroger individuellement chacun des membres du ménage.

La nomenclature utilisée était précise et détaillée : elle était bien entendu variable suivant le sexe. En ce qui concerne les vêtements, on distinguait :

- la confection ;
- les vêtements sur mesure (dont l'étoffe était procurée par la couturière ou le tailleur) ;
- les vêtements sur mesure (dont l'étoffe était fournie par le client) ;
- les achats de tissus pour habillement, laine à tricoter, layette et vêtements pour petite enfants étaient notés globalement pour le ménage.

Les études publiées portent non seulement sur les dépenses mais sur les quantités et sur les prix.

20. Enquête sur l'équipement ménager, décembre 1954. B.H.S. n° 419 du 12 mai 1956. Enquête de l'I.N.S.E.E.

L'étude portait sur l'équipement des ménages : appareils installés, date d'installation, prix d'achat pour les appareils installés en 1953-1954. Elle portait également sur les dépenses de gaz et d'électricité, sur les dépenses d'ameublement en 1953-1954, sur les dépenses d'aménagement, d'installation et de réparation au cours de l'année 1954 et sur les achats d'un certain nombre d'instruments ménagers : vaisselle, articles de cuisine au cours des trois derniers mois.

Seuls les résultats relatifs à l'équipement ont été publiés.

Le domaine d'étude était constitué par l'ensemble des ménages résidant en France métropolitaine, Corse exclue.

L'échantillon de 20.000 ménages était prélevé par tirage au sort dans le fichier des feuilles de logement du tout récent recensement de 1954. Un taux de sondage plus élevé était appliqué dans les catégories sociales aisées.

21. Enquête sur les conditions de logement des Français non agriculteurs en octobre 1955. B.H.S. n° 421 du 26 mai 1956, et Etudes Statistiques, avril 1957 et Annales de Recherches et de Documentation sur la Consommation, avril-juin 1957. Enquête de l'I.N.S.E.E.

Le domaine d'étude était constitué par l'ensemble des logements français, non compris ceux des agriculteurs et salariés agricoles. L'échantillon (10.000 logements environ) était prélevé par tirage au sort dans le fichier des feuilles de logement du recensement de 1954 ; cet échantillon comprenait également un certain nombre de logements construits depuis le recensement.

Le questionnaire comportait : la description de la famille et du logement (nombre de pièces, superficie, confort...) et le relevé des dépenses de logement : loyer, charges, assurances, contributions mobilières, etc... ; pour les personnes occupant un local neuf construit après 1948 à l'aide d'emprunt de l'Etat ou d'un organisme public : annuités de remboursement. Le questionnaire comportait une question directe très simple sur les ressources du ménage.

22. Enquête sur la consommation et les dépenses de viande et des produits laitiers en 1955. B.H.S. n° 427, du 7 juillet 1956 et Consommation avril-juin 1959.

L'étude portait sur les consommations de viande (avec un grand détail quant à la nature des morceaux), charcuterie, triperie, lapin, volaille, gibier, oeufs, beurre, crème, fromage, lait. Le domaine d'étude était constitué par l'ensemble des ménages ordinaires résidant en France métropolitaine.

L'échantillon était prélevé par tirage au sort dans le fichier des feuilles logement du recensement de 1954. Deux échantillons indépendants d'environ 6.000 ménages chacun ont été interrogés respectivement en avril, mai et en octobre 1955.

Les quantités et dépenses étaient enregistrées par carnet de comptes sur une semaine. On tenait compte de l'autoconsommation mais non des dépenses réalisées en cantine, au restaurant. Le questionnaire comportait une question directe sur les revenus du ménage.

23. Enquête sur l'automobile. B.H.S. n° 654 du 17 décembre 1960, et Consommation octobre-décembre 1960.

Un échantillon de véhicules (voitures particulières et commerciales) a été tiré dans le fichier des duplicatas de cartes grises, tenu par l'I.N.S.E.E.

L'enquête a porté sur :

- la distribution du parc automobile (par modèles et marques) à l'intérieur de la population.

- le kilomètrage parcouru et sa répartition entre utilisations professionnelles et privées.

- les dépenses d'utilisation.

Six échantillons distincts de 2.000 utilisateurs chacun ont été étudiés tous les deux mois en 1959. Un carnet de comptes a été utilisé conjointement avec une interview.

§ 2 - Les principales enquêtes étrangères récentes (14)

ETATS-UNIS 1951

Cette enquête a été faite au début de 1951 par le Bureau of Labor Statistics. Son objet principal était de fournir les données nécessaires à la révision de l'indice des prix de détail du B.L.S. ; en même temps, on envisageait déjà lors de la préparation de l'enquête une analyse économique approfondie des résultats, ce qui a conduit à étendre le champ de l'enquête au delà d'un simple relevé d'achats et de quantités consommées.

L'enquête a porté sur toute la population urbaine des Etats-Unis, y compris les habitants des petites villes de quelques milliers d'habitants. Un plan de sondage, fort compliqué, a défini l'échantillon initial des ménages. Ces ménages ont été tirés au hasard parmi plusieurs sous-populations sélectionnées ; dans son ensemble, l'échantillon n'est donc pas un échantillon aléatoire de la

population urbaine entière, mais sa composition a été ajustée de façon à permettre une analyse efficace des résultats selon plusieurs critères tels que la région, la taille des villes et autres.

On a sollicité la collaboration de quelques 16.000 ménages ; non compris les déchets involontaires (absences, maladies), 15 % des ménages ont refusé de prêter leur concours à l'enquête ou ont fourni des renseignements contradictoires.

Les renseignements ont été recueillis au cours de plusieurs interviews avec la ménagère et, le cas échéant, avec d'autres membres du ménage ; en moyenne chaque ménage a subi 8 heures d'interview. Certains renseignements généraux sur le revenu et l'épargne et sur la consommation de certains produits portaient sur toute l'année 1950 : pour d'autres produits une période plus courte a été choisie : la consommation alimentaire, par exemple, faisait l'objet d'un questionnaire fort détaillé qui portait sur la semaine précédant l'interview. En outre, on a demandé aux ménages de remplir un carnet de comptes pendant la semaine suivante, et d'envoyer le carnet rempli au B.L.S. ; la moitié des ménages ont répondu favorablement à cette demande.

ANGLETERRE 1953

Cette enquête a été faite pendant toute l'année 1953 par le Ministry of Labour and National Service. Son but principal était de fournir les données nécessaires à la révision de l'indice des prix de détail, mais afin d'obtenir également des statistiques plus générales, on a étendu le champ de l'enquête à l'ensemble de la population.

L'échantillon a été déterminé à partir d'un plan de sondage complexe ; en principe, l'échantillon a été obtenu par un tirage aléatoire dans l'ensemble des ménages anglais. L'échantillon initial comportait environ 20.000 ménages ; outre les déchets involontaires, 30 % d'entre eux ont refusé de collaborer. Ce taux élevé de refus est dû au fait que la méthode d'enquête imposait un effort considérable aux ménages qui y participaient.

C.R.E.D.O.C.
BIBLIOTHÈQUE

Les ménages reçurent la visite d'un enquêteur qui recueillait, par interview les données générales concernant la composition et le logement du ménage, ainsi que certaines dépenses peu fréquentes de chaque membre du ménage. Ensuite, les dépenses courantes de chaque membre du ménage devaient être notées dans un carnet de comptes pendant trois semaines consécutives ; le revenu de chaque personne devait également être noté sur un questionnaire individuel. Sous la condition que tous aient collaboré à l'enquête, chaque membre du ménage a reçu $ 1.

AUTRICHE 1954

L'enquête autrichienne a eu lieu d'avril 1954 à mars 1955. Elle a été faite par le "Forschungsstelle zur Aufstellung Volkswirtschaftlicher Bilanzen", organisme officiel qui s'occupe de la comptabilité nationale, et son but principal était de combler certaines lacunes dans la documentation statistique de la consommation.

L'enquête porte néanmoins sur la seule population urbaine de 40 villes autrichiennes. L'échantillon de 9.000 ménages a été déterminé à partir d'un plan de sondage qui tenait compte de la taille des villes et de la profession du chef de ménage ; le taux de sondage a été doublé pour les indépendants et pour les habitants des villes autres que Vienne. Parmi les ménages dont on a sollicité la collaboration, le taux de refus a été de 16 %.

Les ménages ont été visités par des enquêteurs qui se sont efforcés d'obtenir leur concours et qui ont expliqué la méthode de l'enquête ; mais l'ensemble des renseignements a été recueilli sans intervention directe de l'enquêteur par un questionnaire sous forme d'un carnet de comptes que la ménagère devait remplir quotidiennement pendant un mois, 100 $ étaient accordés aux ménages qui avaient collaboré à l'enquête.

N O T E S

(1) - cf. : J. DESABIE : Théorie et pratique des sondages, cours école d'application de l'I.N.S.E.E. Tome I - décembre 1959.

(2) - Cette enquête représente, en fait, l'aboutissement d'une longue série d'expériences menées depuis 1945 par l'Institut National de la Statistique ; nous rendons compte de ce progrès dans la section 4, paragraphe 1 de ce chapitre.

(3) - Une raison non négligeable est que le premier point de vue conduit à des méthodes d'enquête beaucoup plus simples et plus classiques que le second.

(4) - Cette difficulté peut être surmontée. Dans les enquêtes sur l'épargne réalisées par l'Institut de Statistiques de l'Université d'Oxford; on a isolé dans certains ménages plusieurs "unités de revenu" auxquelles ont été indépendamment administré le questionnaire. Notons, d'une part que cela revient à effectuer un sondage en grappes qui pose certains problèmes, et d'autre part que la procédure peut difficilement s'étendre à une enquête de budgets familiaux, où il existe toujours une partie des dépenses communes aux différents groupes ainsi isolés.

(5) - La démonstration des propriétés 2 à 4 est classique et très simple. Cf. G. ROTTIER : Initiation à l'Econométrie. Paris 1961, pp. 39-47.

(6) - Notons que dans certains cas étrangers à notre domaine, cela conduit à préférer un estimateur légèrement biaisé, mais peu dispersé, à un estimateur centré mais imprécis.

(7) - Cité par J. DESABIE, op. cit. p. 158

(8) - En même temps, le calcul pratique de cette erreur devient tellement lourd qu'on ne l'effectue pratiquement jamais.

(9) - Cf. Section 4, paragraphe 2.

(10) - I.N.S.E.E. et C.R.E.D.O.C. : Budgets des Français en 1956 : dépenses et niveaux de vie. DUNOD, Paris, 1960.

(11) - N. TABARD, les variations saisonnières des dépenses des ménages, Consommation, n° 3, 1958.

(12) - Le taux de réponse brut est de 72,3 %. Il est voisir de celui observé dans des enquêtes comparables à l'étranger :

Autriche (1953-1954)	73 %
Etats-Unis (1951)	76,4 %
Grande-Bretagne (1953)	67 %

(13) - C.R.E.D.O.C. - S.E.E.F. "La consommation de 1950 à 1957" Consommation, n° 2, 1958 et "La consommation des particuliers de 1950 à 1958, Consommation, n° 3, 1959.

(14) - Cf. J.S. CRAMER : Comparaison Internationale des Enquêtes de Budgets Familiaux "Revue Economique", n° 2 1961.

C H A P I T R E III

L'EXPLICATION DE LA CONSOMMATION
LES COURBES D'ENGEL

I N T R O D U C T I O N

Les enquêtes sur les budgets familiaux fournissent des informations très riches sur la façon dont les structures de consommation se modifient sous l'influence de divers facteurs sociologiques, économiques, démographiques ou géographiques. Réserve faite des erreurs d'échantillonnnage qui peuvent devenir considérables quand on examine des sous-populations trop restreintes, ces enquêtes fournissent une photographie très fine de la structure des consommations à l'intérieur d'un pays. Leurs utilisations sont donc très variées, qu'on s'intéresse au marché de produits particuliers, aux variations des niveaux de vie d'une section à l'autre de la population ou à des questions de planification et de politique économique.

Les tableaux III-A et III-B donnent un exemple des renseignements qu'on peut ainsi déduire des enquêtes. Le tableau III-A montre que la part des diverses dépenses dans le budget (ou "coefficient budgétaire") varie nettement avec la catégorie socio-professionnelle du chef de ménage ; le tableau III-B montre que les coefficients budgétaires varient aussi, quoique d'une façon moins nette, avec l'importance de la commune habitée par les ménages étudiés (1).

De telles remarques n'ont cependant qu'un caractère descriptif et aucun pouvoir explicatif. Les tableaux III-A et III-B n'indiquent pas en effet l'influence nette des facteurs qu'ils considèrent. Les ménages appartenant à deux catégories sociales du tableau III-A ne se répartissent pas, par exemple, de la même façon entre les divers types d'habitat, les diverses structures familiales, les divers niveaux de revenus, etc... Ce tableau ne mesurerait convenablement l'influence de la catégorie sociale que si l'influence des autres variables était nulle (ce qui est évidemment faut) ou que si l'échantillon avait été choisi selon un plan d'expérience qui élimine cette influence.

Mais même si ces conditions étaient réalisées, des relations empiriques comme celles qu'on peut dégager d'une enquête n'ont de pouvoir explicatif que si elles traduisent une hypothèse théorique préalablement formulée, ainsi que nous l'avons montré au chapitre I.

TABLEAU III - A

Coefficients budgétaires selon la catégorie sociale du chef de ménage

	Expl. agricoles	Ouvriers	Employés	Cadres moyens	Cadres sup. et prof. lib.	Inactifs	Ensemble
Alimentation	55,8	50,5	45,3	37,0	27,8	50,6	46,4
Habillement	10,6	12,4	13,2	13,5	12,5	10,6	11,9
Habitation	16,0	17,0	18,5	19,3	20,4	20,0	18,4
Hygiène et santé	5,7	5,1	6,1	5,4	5,9	6,6	5,8
Transport, vacances	5,7	6,0	7,5	13,6	17,2	4,7	8,5
Culture, loisirs, divers	6,2	9,0	9,4	11,2	16,2	7,5	9,0
TOTAL	100,0	100,0	100,0	100,0	100,0	100,0	100,0

TABLEAU III - B

Coefficients budgétaires selon les catégories de communes

	Communes rurales	Petites villes de province (10-100.000) habitants	Grandes villes de province (+100.000) habitants	Agglomération parisienne	Ensemble
Alimentation	53,3	44,1	43,4	39,2	46,4
Habillement	10,9	12,9	12,6	12,4	11,9
Habitation	16,5	20,0	19,3	19,3	18,4
Hygiène et santé	5,6	5,3	5,8	6,7	5,8
Transport et vacances	6,9	8,0	9,4	11,1	8,5
Culture, loisirs et divers	6,8	9,7	9,5	11,3	9,0
TOTAL	100,0	100,0	100,0	100,0	100,0

Une telle hypothèse a été suggérée, pour la première fois, de façon très empirique, par le statisticien allemand Ernest ENGEL (2) en 1857. Observant des budgets de familles ouvrières, ENGEL avait remarqué que la part de l'alimentation dans le budget de ces familles était en moyenne d'autant plus faible que leur niveau de vie était plus élevé. L'hypothèse qui est à la base de l'analyse économétrique des budgets de famille est une généralisation de l'observation d'ENGEL. Elle consiste à admettre que le niveau de vie (3) des ménages étudiés est le facteur d'explication le plus général et le plus important, les autres facteurs susceptibles d'être introduits n'ayant qu'un rôle second. Ainsi, les observations du tableau III-A traduisent-elles d'abord le fait que les niveaux de vie moyens varient considérablement d'une catégorie sociale à l'autre. Ce n'est qu'après avoir éliminé l'influence de ces différences de niveau de vie qu'on pourrait estimer l'éventuel effet propre des catégories sociales.

Au lieu des variations de la part $\frac{C}{R}$ d'une consommation C dans le budget R, il est plus naturel d'étudier directement la relation entre C et R. La figure 1 du chapitre I a donné l'exemple d'une telle relation ; elle montre comment varie la dépense alimentaire C d'un échantillon de ménages en fonction du niveau de vie R. Cette relation est évidemment stochastique ; elle peut cependant être approximativement représentée par une courbe continue Γ qu'on appelle la courbe d'ENGEL pour les dépenses alimentaires dans la population considérée.

Jusqu'à présent, la démarche suivie est encore empirique (4). Elle cesse de l'être si l'on admet que la relation ainsi étudiée entre le niveau de vie et la consommation d'un produit peut être identifiée avec la relation partielle entre le revenu et la consommation (c'est à dire une relation de demande à prix constants), déduite d'une hypothèse théorique sur le comportement du consommateur individuel.

Cette identification a été proposée par plusieurs auteurs aux environs de 1930, mais la première étude économétrique faite selon cette suggestion est celle d'ALLEN et BOWLEY en 1955. Quelques remarques présentées par ces auteurs dans l'introduction de leur ouvrage en montrent bien l'esprit : "l'objet de cette étude est de dégager la mesure dans laquelle on peut décrire les dépenses de familles, ou de groupes de familles, par des lois ou des équations, de relier les lois ainsi degagées aux postulats de la théorie économique, et de décrire les écarts à la moyenne qui résultent de différences dans les choix des familles étudiées" (5).

Les auteurs, qui ont principalement étudié des budgets rassemblés avant 1914, ont ajusté, par les moindres carrés, des courbes d'ENGEL linéaires, en montrant que cette forme est justifiée si l'on admet que les fonctions de choix sont quadratiques.

Tous les travaux qui ont été faits depuis vingt-cinq ans dans le domaine de l'économétrie des budgets de famille se placent dans la ligne définie par ALLEN et BOWLEY. Les principaux progrès ont consisté à étendre le choix de la forme des courbes d'ENGEL, à étudier avec plus de précision la définition des variables introduites et à analyser l'influence de divers facteurs, qualitatifs ou quantitatifs, dont l'omission risque d'introduire des erreurs systématiques dans l'estimation des paramètres des courbes d'ENGEL.

Mais tous ces travaux reposent sur les mêmes bases logiques et visent à définir le même type de loi.

Nous ne reviendrons pas ici sur les hypothèses théoriques qui fondent l'analyse des courbes d'ENGEL, puisque nous avons déjà examiné ce problème dans notre premier chapitre. Par contre, nous étudierons dans les pages suivantes, les principaux problèmes d'ordre économétrique qu'on doit résoudre pour déterminer correctement les lois d'ENGEL dans une population. Ces problèmes économétriques sont étroitement liés aux propriétés des méthodes d'estimation statistique utilisées. Aussi devons-nous d'abord rappeler un certain nombre de résultats classiques dans ce domaine ; ils permettront d'interpréter plus commodément le reste du chapitre.

1ère PARTIE : PROBLEMES DE METHODE

SECTION 1

LES TECHNIQUES D'ESTIMATION STATISTIQUE

§ 1 - L'estimation dans un modèle linéaire

Nous utiliserons surtout des modèles linéaires des courbes d'ENGEL. La théorie sur laquelle repose l'estimation des paramètres de ces modèles est classique et il n'est pas lieu de l'exposer ici (6). Cependant, la méthode d'estimation par les moindres carrés repose sur un ensemble d'hypothèses qu'il sera utile de rappeler et d'interpréter. C'est, en effet, à partir de ces hypothèses que s'introduisent la plupart des particularités technique de l'étude des courbes d'ENGEL.

Soit :

$$Y_i = \alpha X_i + \beta + U_i \qquad (I\text{-}1)$$

un modèle économétrique linéaire. α et β sont deux paramètres qu'on suppose bien définis dans une population hypothétique infinie. X_i est une variable réelle quelconque, certaine ou aléatoire. Y_i est aléatoire par l'intermédiaire du résidu aléatoire U_i.

On dispose de n observations du couple X Y

(I-2)
$$\left\| \begin{array}{cc} x_1 & y_1 \\ \cdots & \cdots \\ x_i & y_i \\ \cdots & \cdots \\ x_n & y_n \end{array} \right\|$$

On se propose d'estimer α et β respectivement par :

(I-3)
$$\begin{cases} \hat{a} = \dfrac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_1^n (x_i - \bar{x})^2} \\ \hat{b} = \bar{y} - \hat{a}\,\bar{x} \end{cases}$$

$\hat{a}$ et $\hat{b}$ seront des estimateurs convenables si on admet un ensemble d'hypothèses qu'il est commode de hiérarchiser.

A - HYPOTHESE 1

La relation entre X et la part non aléatoire de Y est exactement linéaire.

Cela signifie que le résidu U_i n'incorpore pas d'erreurs dans la spécification du modèle. Nous verrons dans la section 2 l'importance que revêt l'étude de ce point dans le cas des courbes d'ENGEL.

On remarquera que si H_1 n'est pas satisfaite mais si les autres hypothèses le sont, $\hat{a}$ et $\hat{b}$ fournissent cependant la meilleure approximation linéaire du phénomène étudié.

B - HYPOTHESE 2

Les variables x et y sont observées sans erreur de mesure.

HYPOTHESE 3

Le résidu U_i est distribué indépendamment de X_j, que i soit égal ou différent de j. Remarquons que cette hypothèse ne s'introduit que si X est aléatoire, sinon elle se confond avec :

HYPOTHESE 4

L'espérance mathématique de U_i est nulle quel que soit i.

Ces trois hypothèses sont nécessaires pour que les estimateurs $\hat{a}$ et $\hat{b}$ soient centrés.

L'hypothèse 2 est fréquemment mise en doute dans la pratique. Les résultats d'enquêtes de consommation sont en effet affectés d'incertitudes, variables selon le type d'informations recueillies, et souvent assez considérables. L'erreur est particulièrement forte sur les revenus des ménages qu'il serait cependant raisonnable de prendre comme variable explicative.

On sait que lorsqu'une variable explicative est mesurée avec une erreur accidentelle de moyenne nulle, distribuée indépendamment de la variables affectée, il en résulte un biais vers le bas dans l'estimation du coefficient de régression. Ce résultat ne s'applique cependant pas directement à notre cas. En effet, l'erreur de mesure sur les revenus est complexe et comprend des composantes systématiques et accidentelles. Dans ce cas, l'erreur de mesure peut non seulement introduire un biais dans l'estimation, mais affecter la forme apparente de la relation (7).

On argue souvent de cette erreur pour substituer la dépense totale de consommation au revenu comme variable explicative. Nous reviendrons sur cette question (section 3, paragraphe 1). Cette substitution n'est, en effet, pas entièrement justifiée. Outre que la consommation totale peut être affectée d'erreurs accidentelles notables, elle est surtout liée aux consommations partielles et au revenu par un système d'équations simultanées. Or on sait, depuis les travaux de la Cowles Commission (8), que l'hypothèse 3 est nécessairement infirmée dans ce cas.

Cette hypothèse est également mise en cause dans un autre cas, fréquent en pratique. C'est quand le résidu U_i incorpore l'effet de facteurs systématiques influençant la consommation et liés statistiquement au revenu. De tels facteurs peuvent être quantitatifs, comme la taille des familles étudiées ; nous verrons comment en tenir compte dans la section 4. Ces facteurs peuvent aussi être qualitatifs. Nous verrons en particulier que le coefficient de régression est un estimateur biaisé dans le cas où la population est composée de plusieurs sous-populations de revenus moyens différents et dont les courbes d'ENGEL n'ont pas la même ordonnée à l'origine.

Enfin, l'hypothèse 3 est également infirmée lorsque les consommations de divers individus sont interdépendantes, c'est-à-dire lorsqu'un individu détermine sa consommation non seulement en fonction de son revenu, mais aussi de la consommation observée chez d'autres individus au même moment. Il s'agit là d'une hypothèse que nous avons déjà examinée du point de vue de la théorie économique dans le premier chapitre. Nous avons alors reconnu que, lorsqu'elle était vérifiée, la démarche analytique que nous suivons était illégitime.

C - HYPOTHESE 5

Les distributions des résidus pour diverses valeurs de i sont indépendantes.

HYPOTHESE 6

La variance des résidus est indépendante de i.

Ces hypothèses sont nécessaires pour que les estimations soient efficientes. L'hypothèse 5 ne s'introduit pas dans notre domaine. Par contre, l'hypothèse 6 a été prise en considération par certains auteurs, qui en ont tiré augument en faveur d'un ajustement à élasticité constante (cf. section 2, paragraphe 1) et ont proposé des corrections à l'estimateur dans les autres cas (9). Cette question ne nous semble pas d'une importance fondamentale.

D - HYPOTHESE 7

La distribution des U_i est normale.

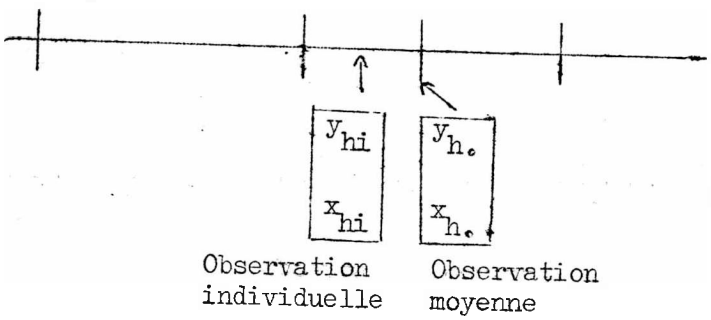
Dans ce cas l'estimation par les moindres carrés se confond avec l'estimation par le maximum de vraisemblance.

Rien ne permet d'affirmer que cette hypothèse soit vérifiée dans le cas de budgets familiaux. Cependant, de nombreux tests d'application commode dans notre domaine (comme les tests d'analyse de la covariance pour examiner l'homogénéité des régressions dans diverses sous-populations) exigent que cette hypothèse soit vérifiée pour pouvoir être appliqués en toute rigueur.

§ 2 - L'estimation sur des données groupées

Dans la pratique, on est souvent amené à grouper les données observées afin d'alléger les calculs. Dans le cas d'une enquête de budgets familiaux portant sur 10.000 ménages, le calcul d'un coefficient de régression exige le calcul de 20.000 carrés et de 10.000 produits xy. Si l'on groupe les ménages en 100 grappes de 100, après les avoir classés dans l'ordre des revenus croissants et si l'on remplace chaque grappe par sa moyenne, on a à calculer 200 additions de 100 nombres en plus, mais on n'a à calculer que 200 carrés et 100 produits. Le gain de temps (et de coût) est considérable.

On peut montrer que cette procédure n'affecte que faiblement les qualités de l'estimation par les moindres carrés.



Nous avons N observations du couple de variables x_{hi}, y_{hi}. Les variables sont groupées dans l'ordre des x croissants en H groupes de n observations (donc nH = N). On remplace les observations d'un groupe par leurs moyennes $x_{h.}$, $y_{h.}$

Sur les données individuelles, toutes les hypothèses énoncées plus haut étaient satisfaites pour qu'on puisse convenablement estimer par les moindres carrés les paramètres du modèle :

(I-4) $$Y_{hi} = \alpha X_{hi} + \beta + u_{hi}$$

En additionnant par rapport à i, et en divisant par n, on trouve :

(I-5) $$y_{h.} = \alpha x_{h.} + \beta + u_{h.}$$

Puisque

(I-6) $$u_{h.} = \frac{1}{n} \sum u_{hi} \quad \text{et que} \quad x_{h.} = \frac{1}{n} \sum x_{hi}$$

On a évidemment

(I-7) $$E\,(x_{h.}\ x_{h'.}) = 0$$ que h et h' soient égaux ou différents.
En effet, cela découle de l'hypothèse

(I-8) $$E\,(u_{hi}\ ;\ x_{h'i'}) = 0$$

L'estimation a* sur les données groupées n'est pas biaisée. Que peut-on dire de sa variance ?

Comme

$$V(u_{h.}) = \frac{\sigma^2}{n} \tag{I-9}$$

On a

$$V(a^*) = \frac{\sigma^2}{n \sum_h (x_{h.} - \bar{x})^2} \tag{I-10}$$

Exprimons V (a*) en fonction de V (â), â étant l'estimation de α à partir du modèle (I-4).

On peut écrire

$$V(\hat{a}) = \frac{\sigma^2}{\sum_h \sum_i (x_{hi} - \bar{x})^2} \tag{I-11}$$

ou, par un artifice classique de calcul

$$\begin{cases} \sum_h \sum_i (x_{hi} - \bar{x})^2 = \sum_h \sum_i (x_{hi} - x_{h.} + x_{h.} - \bar{x})^2 \\ \\ = \sum_h \sum_i (x_{hi} - x_{h.})^2 + \sum_h \sum_i (x_{h.} - \bar{x})^2 \\ \\ = \sum \sum_i (x_{hi} - x_{h.})^2 + n \sum_h (x_{h.} - \bar{x})^2 \end{cases} \tag{I-12}$$

car le double produit s'annule

$$(I\text{-}13) \quad \sum_h \sum_i (x_{hi} - x_{h.})(x_{h.} - \bar{x}) = \sum_h (x_{h.} - \bar{x}) \sum_i (x_{hi} - x_{h.})$$

et

$$(I\text{-}14) \quad \sum_i (x_{hi} - x_{h.}) = 0 \quad \text{quel que soit } h$$

On retrouve l'équation classique de l'analyse de la variance

$$(I\text{-}15) \quad \sum_h \sum_i (x_{hi} - \bar{x})^2 = \sum_h \sum_i (x_{hi} - x_{h.})^2 + h \sum_h (x_{h.} - \bar{x})^2$$

variance totale — variance dans les groupes — variance entre les groupes

on a

$$(I\text{-}16) \quad \frac{1}{V(\hat{a})} = \frac{1}{V(a^*)} + \frac{\sum_h \sum_i (x_{hi} - x_{h.})^2}{\sigma^2}$$

Le second terme est en général faible par rapport au premier dès que le nombre de groupes dépasse une dizaine (10), et la perte d'efficience est donc négligeable. Il faut par contre prendre garde à ne pas utiliser le coefficient de détermination R^2 comme un indicateur de la qualité de l'ajustement quand on travaille sur des données groupées. R^2 mesure en effet le rapport de la variance expliquée à la variance de y entre les groupes, qui est beaucoup plus faible que la variance totale de y (cf. équation I-15). R^2 donne donc une idée trop favorable de la qualité de la liaison.

SECTION 2

LA FORME DES COURBES D'ENGEL

§ 1 - Généralités

Soit C_{ik} la consommation d'un produit k par un individu i de niveau de vie R_i. Les hypothèses introduites dans le chapitre I conduisent à expliquer la consommation C_{ik} par un modèle de la forme :

$$C_{ik} = f(R_i) + u_i \qquad \text{(II-1)}$$

dans lequel $f(R_i)$ est l'équation cherchée de la courbe d'ENGEL pour le produit k. Nous dirons que la population est homogène si le résidu aléatoire U_i satisfait aux conditions suivantes :

- U_i est statistiquement indépendant du revenu R_i

- U_i est statistiquement indépendant des revenus R_j

($j \neq i$) des autres individus, et de leurs consommations C_{jh} de biens quelconques ($h = K$ ou $h \neq k$).

Dans cette section comme dans la suivante, nous admettrons que ces deux conditions sont satisfaites (II). Les résidus U_i traduisent, on le sait, (cf. Chapitre I), quatre ordres d'influences :

- Les variations aléatoires du comportement d'un même individu,

- Les différences de goûts entre individus,

- L'influence des facteurs secondaires non explicitement introduits dans l'analyse, comme les régions, les catégories sociales, etc...

- Les erreurs dues, soit à une mauvaise spécification de la forme de la fonction f, soit à des imperfections dans la mesure de C_{ik}.

Nous étudierons dans cette section les problèmes que pose le choix de la forme de la fonction f avant d'examiner la définition des variables C et R dans la section 3.

On sait que l'estimation des paramètres du modèle (III-16) ne peut se faire qu'une fois qu'on a choisi la forme de la fonction f (R). Dans de nombreux problèmes économétriques, ce choix ne présente pas de difficulté sérieuse ; c'est notamment le cas des analyses de séries chronologiques, où le champ de variation des variables est assez faible. Dans le cas présent, le choix présente au contraire une importance fondamentale qui justifie le nombre des travaux qui y ont été consacrés dans les 10 dernières années. Cela vient de ce que les observations obtenues à partir d'une enquête par sondage correspondent à un très large champ de variation des niveaux de vie, qui peut souvent s'étendre de plus de 1 à 10. Si n'importe quelle forme de courbe peut représenter convenablement les observations voisines de la moyenne, il n'en est évidemment pas de même pour des observations correspondant à des niveaux de vie très faibles ou très élevés. Or, il est important d'avoir une bonne représentation de l'ensemble du phénomène, car on utilise souvent les courbes d'ENGEL pour établir des projections de consommation sous l'hypothèse d'une augmentation importante des niveaux de vie.

Le choix de la fonction ajustée ne peut malheureusement pas découler d'un raisonnement théorique. Un tel raisonnement, nous l'avons vu plus haut, peut seulement suggérer la nature des variables à prendre en considération, sans permettre d'établir, comme l'a montré SAMUELSON (12) aucune restriction opératoire sur la façon dont ces variables peuvent jouer. Le raisonnement d'ALLEN et BOWLEY, selon lequel des courbes d'ENGEL linéaires correspondent à des surfaces de choix quadratiques, n'introduit par exemple aucune restriction sur la forme des courbes d'ENGEL. Il s'applique au contraire dans l'autre sens : si l'on postule l'existence de surfaces de choix identiques, en moyenne, pour tous les consommateurs, et si l'on observe que les courbes d'ENGEL sont linéaires, on peut en déduire que les surfaces de choix sont quadratiques.

Le sens commun peut cependant suggérer quelques hypothèses complémentaires (13). La consommation d'un produit donné ne peut pas être négative. Elle cesse d'être nulle pour une valeur en général non nulle du niveau de vie (Cf. Fig. III-1), correspondant à un seuil moyen d'introduction du bien considéré dans le budget (cela traduit une notion voisine de celle de hiérarchie des choix introduite par R. ROY). Dans une première phase, la croissance de la consommation s'accélère ; après un point d'inflexion, elle se ralentit, pour tendre vers la saturation (15). Dans le cas où l'on considère un produit isolé bien défini, la tendance à la saturation se traduit par l'existence d'une asymptote horizontale ; dans le cas, plus fréquent, où l'on considère des groupes de produits et où l'on raisonne sur les dépenses, le glissement vers des variétés de qualité supérieure, dont nous parlerons plus bas (3 - § 2), peut être tel qu'il n'y ait pas d'asymptote. La notion de saturation correspond alors à une diminution régulière de l'élasticité de la consommation par rapport au niveau de vie quand ce dernier augmente.

On introduit parfois une contrainte supplémentaire : la forme de fonction choisie doit être telle que si on l'ajuste à tous les postes du budget (y compris l'épargne), la somme des lois obtenues doit être identique au revenu. WORSWICK et CHAMPERNOWNE (16) ont montré que cette condition est satisfaite si la fonction choisie est une somme de fonctions arbitraires quelconques, dont au moins l'une d'entre elles est linéaire en R. Cette contrainte d'additivité peut aussi s'exprimer par une relation entre élasticités. En dérivant les deux termes de :

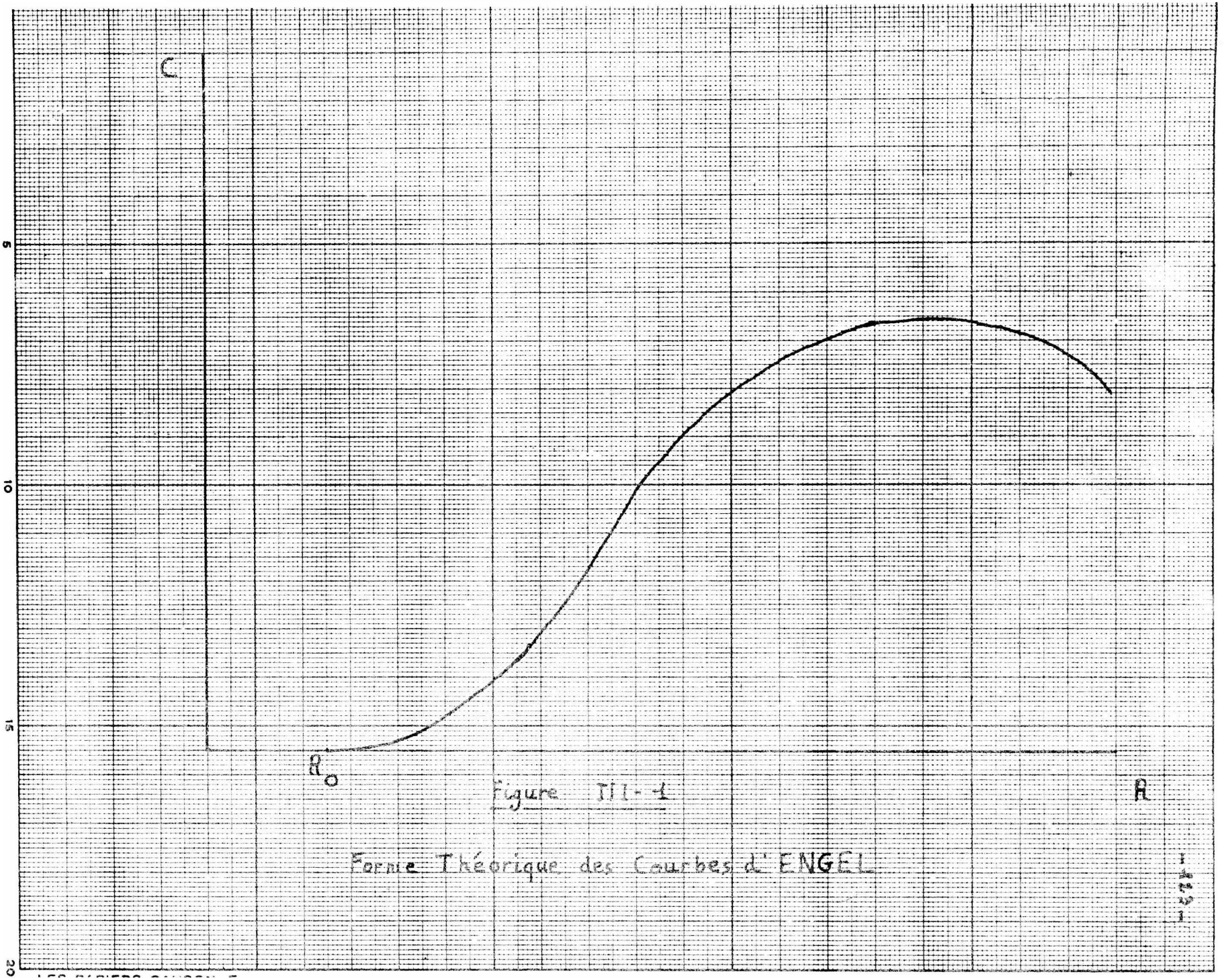


Figure III-1

Forme Théorique des Courbes d'ENGEL

(II-2) $$\sum f_k(R) = R$$

il vient :

(II-3) $$\sum \frac{df_k}{dR} = 1$$

En exprimant les dérivés en termes d'élasticité :

(II-4) $$\frac{df}{dR} = E_R^f \cdot \frac{fk}{R}$$

on obtient :

(II-5) $$\frac{1}{R} \sum f_k E_R^{fk} = 1$$

La moyenne pondérée des élasticités doit être égale à l'unité. Si l'on considère un groupe de produits qui n'épuise pas le budget (par exemple, les diverses composantes de la consommation alimentaire) (II-2) s'écrit :

(II-6) $$\sum f_k(R) = F(R)$$

et (II-5) devient :

(II-7) $$\frac{1}{F} \sum f_k E_R^{fk} = E_R^F$$

La relation (II-7) est beaucoup plus souvent utilisée que (II-5). En effet la contrainte d'additivité, sous la forme directe ou sous la forme (II-5), a peu d'utilité : d'une part, les courbes d'ENGEL n'ont de pouvoir explicatif, nous l'avons vu, que pour des biens de consommation courante et il n'y a pas grand sens à en ajuster pour d'autres consommations. D'autre part, même dans les cas où elles s'appliquent, nous verrons plus bas qu'il est rarement souhaitable d'utiliser la même forme de loi pour tous les produits.

Les remarques précédentes, même si elles sont utiles, ne permettent pas de résoudre le problème. Les conditions résumées dans la figure (III-1) sont, en effet, vérifiées par un grand nombre de fonctions différentes, parmi lesquelles le choix sera largement conditionné par la commodité de l'ajustement. En outre, le champ de variation des données, même s'il s'étend de 1 à 10, n'est pas assez large pour que les trois parties de la courbe (III-1) apparaissent simultanément dans le cas de biens de consommation courante. Il s'agit donc seulement de représenter convenablement un segment de cette courbe, dont la concavité sera habituellement toujours tournée dans le même sens pour un produit donné. Le choix des fonctions ajustées est alors beaucoup plus large.

La solution du problème est donc essentiellement empirique. Cela explique qu'une très grande très grande variété de solutions ait été proposée par les différents auteurs, mais qu'aucune ne soit encore universellement adoptée (17). Nous n'étudierons en détail que trois formes de loi qui nous semblent particulièrement intéressantes et qui traduisent des hypothèses bien définies sur la variation avec R de l'élasticité ou de la propension marginale à consommer le produit étudié : les lois à élasticité constante, semi-logarithmiques et logarithmico-normales. Mais on peut citer aussi l'ajustement de branches d'hyperboles équilatères suggéré par TORNQVIST (18) ; de relations paraboliques (19) ; de divers modèles linéaires plus complexes que ceux que nous étudierons au paragraphe 2.

En général, l'ajustement de ces différents types de fonction aboutit à des élasticités très voisines au point moyen, mais évidemment très différentes quand on s'en écarte. Malheureusement, les échantillons dont on dispose ne comprennent en général pas assez d'observations dans les catégories extrêmes de niveau de vie pour qu'on puisse pousser entièrement l'étude. Aussi convient-il d'utiliser, avec prudence, les résultats correspondant à ces catégories et a fortiori, de s'abstenir d'extrapoler les lois ajustées en dehors du domaine de variation de R pour lequel elles ont été établies.

§ 2 - Modèles linéaires des courbes d'Engel

Le modèle :

(II-8) $$C_{ik} = a_k R_i + b_k + u_{ik}$$

qui correspond à des lois linéaires proposées par ALLEN et BOWLEY, n'est pratiquement plus utilisé. En effet, l'examen graphique des données montre presque toujours une concavité accusée, dans un sens ou dans l'autre.

Le modèle (III-8) correspond à l'hypothèse selon laquelle la propension marginale à consommer a est constante, l'élasticité variant avec R selon la loi :

$$E_k = \frac{a_k R}{a_k R + b_k} \qquad \text{(II-9)}$$

Cette hypothèse est intuitivement peu satisfaisante. En effet, pour les produits alimentaires, $\frac{C_k}{R}$ doit diminuer quand R augmente, conformément à la loi d'ENGEL. Comme pour ces produits, $a_k < 1$, il en résulte que l'ordonnée à l'origine est positive. Dans ce cas, la formule (II-9) montre que l'élasticité des dépenses alimentaires augmente avec R, ce qui est incorrect.

Mais il existe de nombreux autres modèles linéaires que (II-8). En effet un tel modèle doit être linéaire en termes des paramètres à estimer et non des variables C et R. Tout modèle de la forme :

$$\varphi(C_{ik}) = a_k \psi(R_i) + b_k + u_{ik} \qquad \text{(II-10)}$$

satisfait à ces conditions, φ et ψ étant deux fonctions monotones. Plus généralement, un modèle linéaire pourra s'écrire :

$$\varphi(C_{ik}) = \sum a_{jk} \psi_j(R_i) + b_k + u_{ik} \qquad \text{(II-11)}$$

le cas le plus simple d'un tel modèle étant un polynome en R.

Le domaine des modèles linéaires est donc très vaste. Même en se limitant au cas (II-10), on a représenté des courbes d'ENGEL par les différents modèles suivants (20) :

$$
(II\text{-}12)\qquad
\begin{cases}
\log C_{ik} = a_k \log R_i + b_k + u_{ik} \\
C_{ik} = a_k \log R_i + b_k + u_{ik} \\
\log C_{ik} = a_k - \dfrac{b_k}{R_i} + u_{ik} \\
C_{ik} = a_k - \dfrac{b_k}{R_i} \\
\log C_{ik} = a_k - \dfrac{b_k}{R_i} - d_k \log R_i + u_{ik}
\end{cases}
$$

Nous ne considérons que les deux premiers modèles de (III-12), qui sont respectivement logarithmique et semi-logarithmique. Notre expérience rejoint celle de PRAIS et HOUTHAKKER (op.cit.), en les considérant comme particulièrement simples et utiles (Cf. Chap. III, 2ème partie). Dans les quelques cas où ils ne conviennent pas, il nous semble préférable d'adopter le modèle non linéaire que nous décrirons dans le paragraphe 3, et qui est particulièrement souple.

En décrivant les différents modèles, nous indiquerons comment on peut les ajuster graphiquement. Avec les moyens de calcul actuels, les ajustements analytiques de courbes d'ENGEL se font de façon systématique, à partir d'un modèle qui doit être précisé à l'avance. On gagne un temps considérable en effectuant d'abord une exploration graphique rapide de la forme des lois, à partir des données moyennes correspondant à un nombre assez faible de groupes de ménages rangés dans l'ordre des revenus croissants.

A - LOIS A ELASTICITE CONSTANTE

Ce modèle s'écrit :

$$(II\text{-}13) \qquad \begin{cases} \log C_{ik} = a_k \log R_i + b_k u_{ik} \\ \text{ou} \quad C_{ik} = 10^{b_k} R_i^{a_k} 10^{u_{ik}} \end{cases}$$

Il correspond à une loi à élasticité constante, très généralement utilisée dans les travaux d'analyse de la demande. En dérivant (II-13), on voit, en effet, immédiatement que l'élasticité est égale à a.

La figure (III-2) indique la forme de la loi pour quelques valeurs de a.

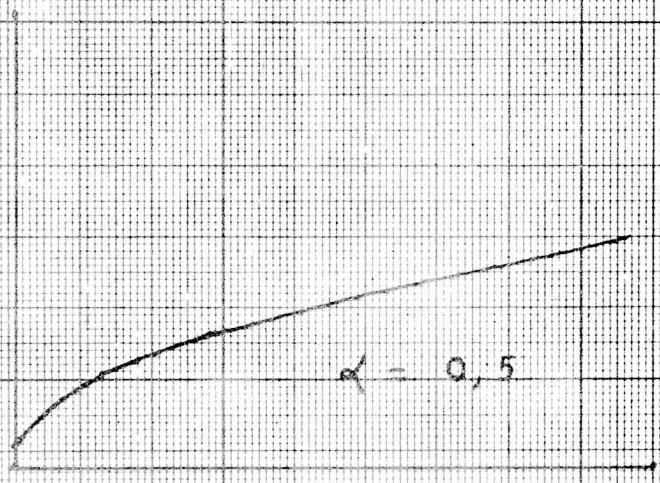
L'ajustement graphique se fait très simplement, en portant les observations sur du papier à graduation doublement logarithmique et en traçant, au jugé, la droite de régression ; l'élasticité a est la pente de cette droite. Deux précautions doivent cependant être prises. D'une part, il est en général nécessaire d'utiliser un papier à deux modules, qui permet de représenter des variations de 1 à 100 au lieu de 1 à 10. D'autre part, quand les données ne sont pas très bien alignées, on gagne de la précision en calculant d'abord les coordonnées du point moyen (qui sont égales à la moyenne des logarithmes des observations, et non aux logarithmes de leurs moyennes, les moyennes étant pondérées par le nombre d'observations dans chaque groupe). On sait que la droite d'ajustement doit passer par ce point moyen.

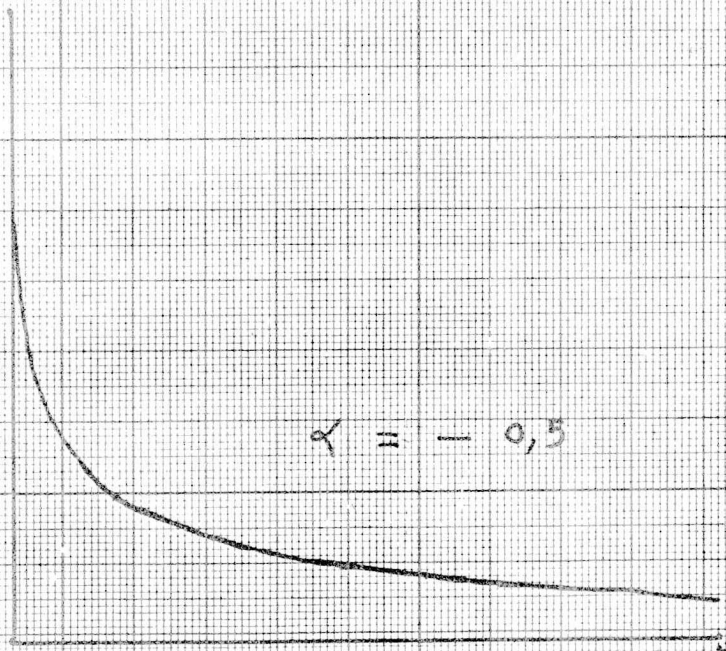
L'expérience nous a montré qu'on jugeait en outre mieux de la forme de la loi quand, au lieu de représenter chaque groupe d'observations par un point, on le représente par un cercle de surface proportionnelle au nombre d'observations du groupe (comme dans la figure III-6).

Figure III-2

LOIS A ELASTICITE CONTANTE

$y = x^{\alpha}$





5 10 15 20

LES PAPIERS CANSON France

L'ajustement analytique se fait sans difficulté en appliquant les formules classiques des moindres carrés. Dans le cas de données groupées, il convient de pondérer chaque moyenne de groupe par le nombre d'observations individuelles qui la compose. Dans le cas de données individuelles, il faut d'abord isoler les ménages ayant une consommation nulle et leur affecter arbitrairement une consommation égale à la plus petite unité disponible (pour éviter d'avoir à prendre le logarithme de zéro).

Les modèles à élasticité constante présentent quelques propriétés intéressantes ; ils traduisent la contrainte que les consommations ne peuvent pas être négatives (dans les modèles où figure la valeur arithmétique de la consommation, cette contrainte affecte non seulement la loi $f_k(R_i)$, mais les distributions des résidus, puisque cette contrainte vaut pour toutes les consommations possibles et pas seulement pour $E(C_{ik} \mid R_i)$. Il serait donc légitime, en toute rigueur, de la traduire dans les méthodes d'estimation). En outre, ces modèles conduisent à des estimations efficientes, car les distributions des résidus U_i y sont plus vraisemblablement homoscédastiques que quand C_i figure par sa valeur arithmétique. Cela vient de ce que l'effet des variables secondaires et des erreurs de mesure sur C_i (effet représenté par u_i) est multiplicatif plutôt qu'additif.

Cependant, ces propriétés ne méritent d'être considérées que si, d'abord, la loi représente correctement l'allure du phénomène. Les résultats que nous avons obtenus, comme ceux de PRAIS et HOUTHAKKER, conduisent à limiter l'application du modèle aux consommations non alimentaires. Malgré cela, les lois à élasticité constante ont une importance analytique particulière : la solution de certains problèmes, que nous étudierons dans les sections 4 et 5 de ce chapitre, est beaucoup plus simple quand on utilise un tel modèle. Cela vient soit de ce qu'il est commode que les variables C et R figurent par leur logarithme, soit de ce que la distribution des résidus U_i doit satisfaire à certaines conditions qui ont plus de chance d'être vérifiées dans ce modèle.

Le groupement des produits et l'addition des consommations individuelles sont assez peu commodes avec une loi à élasticité constante. Si C_1 et C_2 sont deux consommations qui suivent une telle loi, le total $C_1 + C_2$ ne suit en effet pas rigoureusement une loi de même forme. En effet, l'élasticité de $C_1 + C_2$ par rapport à R s'écrit :

$$(II\text{-}14) \qquad E(C_1 + C_2) = a_1 \frac{C_1}{C_1 + C_2} + a_2 \frac{C_2}{C_1 + C_2}$$

Cette élasticité est une moyenne pondérée des élasticités a_1 et a_2, les coefficients de pondération dépendant des parts relatives de C_1 et C_2 dans le budget. Elle n'est constante que si les coefficients budgétaires de C_1 et C_2 sont proportionnels quel que soit R. Cela n'est vérifié en toute rigueur que si les élasticités a_1 et a_2 sont égales, et est approximativement vérifié si ces élasticités sont peu différentes.

Or, on a très souvent à estimer ainsi le coefficient d'élasticité pour un groupe de produits (exemple : élasticité de la dépense totale de viande, connaissant l'élasticité de la consommation des diverses espèces). On calcule habituellement le coefficient d'élasticité en prenant comme pondération les rapports :

$$(II\text{-}15) \qquad \frac{C_1}{C_1 + C_2} \quad \text{et} \quad \frac{C_2}{C_1 + C_2}$$

au point moyen. Mais cela n'est qu'une approximation.

L'agrégation des lois individuelles pour l'ensemble de la population n'est facile que si tous les revenus augmentent du même pourcentage λ. Dans ce cas, la loi peut être appliquée au niveau de vie moyen et à la consommation moyenne de l'ensemble de la population. L'augmentation μ de la consommation est donnée par :

$$(II\text{-}16) \qquad 1 + \mu = (1 + \lambda)^{a_k}$$

$(1 + \lambda)^{a_k}$ ne peut se réduire à son premier terme, $1 + a_k \lambda$, que si λ est faible (inférieur à $\frac{1}{10}$) ou que si a_k est voisin de l'unité. Dans les autres cas, les abaques des figures (III-3) et (III-4) seront commodes pour déterminer $1 + \mu$. (21)

FIGURE III - 3

Valeur de $(1+x)^a$ pour $a > 0$, $x > 0$

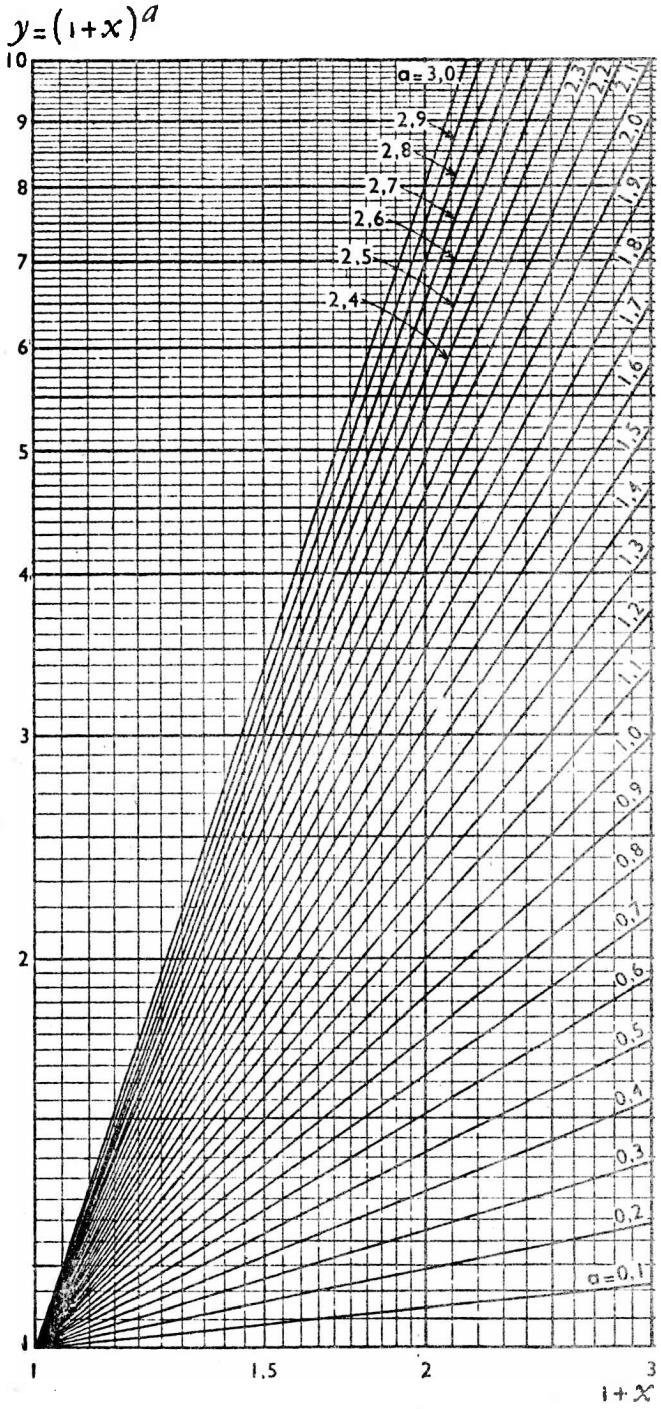
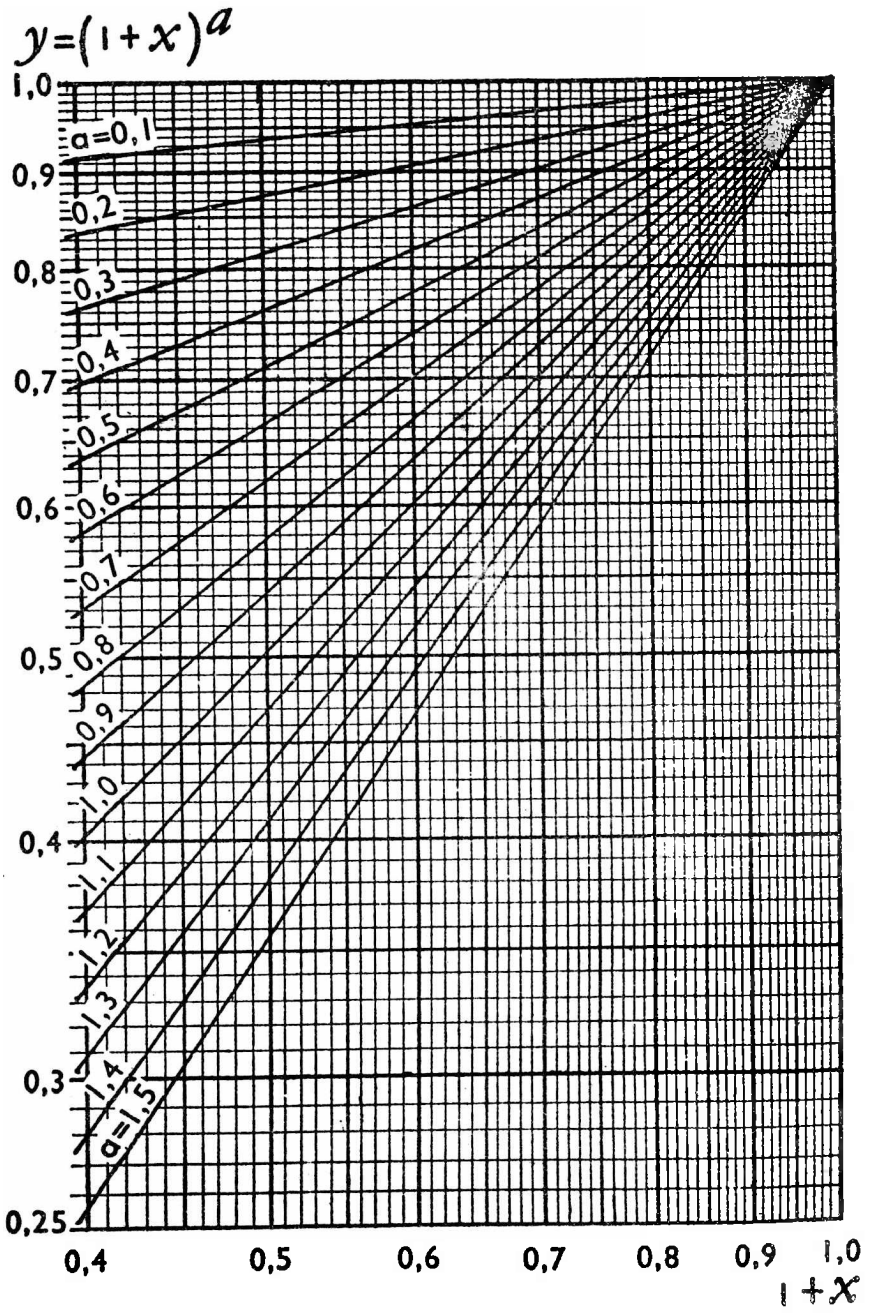


FIGURE III - 4

Valeur de $(1 + x)^a$ pour $a > 0$, $x < 0$



B - LOIS SEMI-LOGARITHMIQUES

Dans le cas des consommations alimentaires, l'examen des observations montre que l'élasticité décroît quand le niveau de vie augmente (Cf. figure III-6). Un modèle simple qui en tient compte est :

$$C_{ik} = a_k \log R_i + b_k + U_{ik} \qquad \text{(II-17)}$$

où $\log R_i$ est le logarithme népérien de R_i.

Ce modèle correspond aux hypothèses suivantes : la propension marginale à consommer, égale à $\frac{a}{R}$, est inversement proportionnelle au niveau de vie, et l'élasticité, égale à $\frac{a_k}{C_k}$, est inversement proportionnelle au niveau de la consommation (22). La figure (III-5) donne un exemple de la variation de l'élasticité avec le niveau de vie. Alors que, dans le modèle à élasticité constante, les paramètres a et b étaient sans dimensions, ils ont ici la même dimension que la consommation C. Ils varient donc avec l'unité de mesure.

La différence entre les modèles logarithmique et semi-logarithmique est simple : si tous les revenus augmentent du même pourcentage, la consommation augmente du même pourcentage pour tous les consommateurs dans le premier cas, et du même montant absolu dans le second cas.

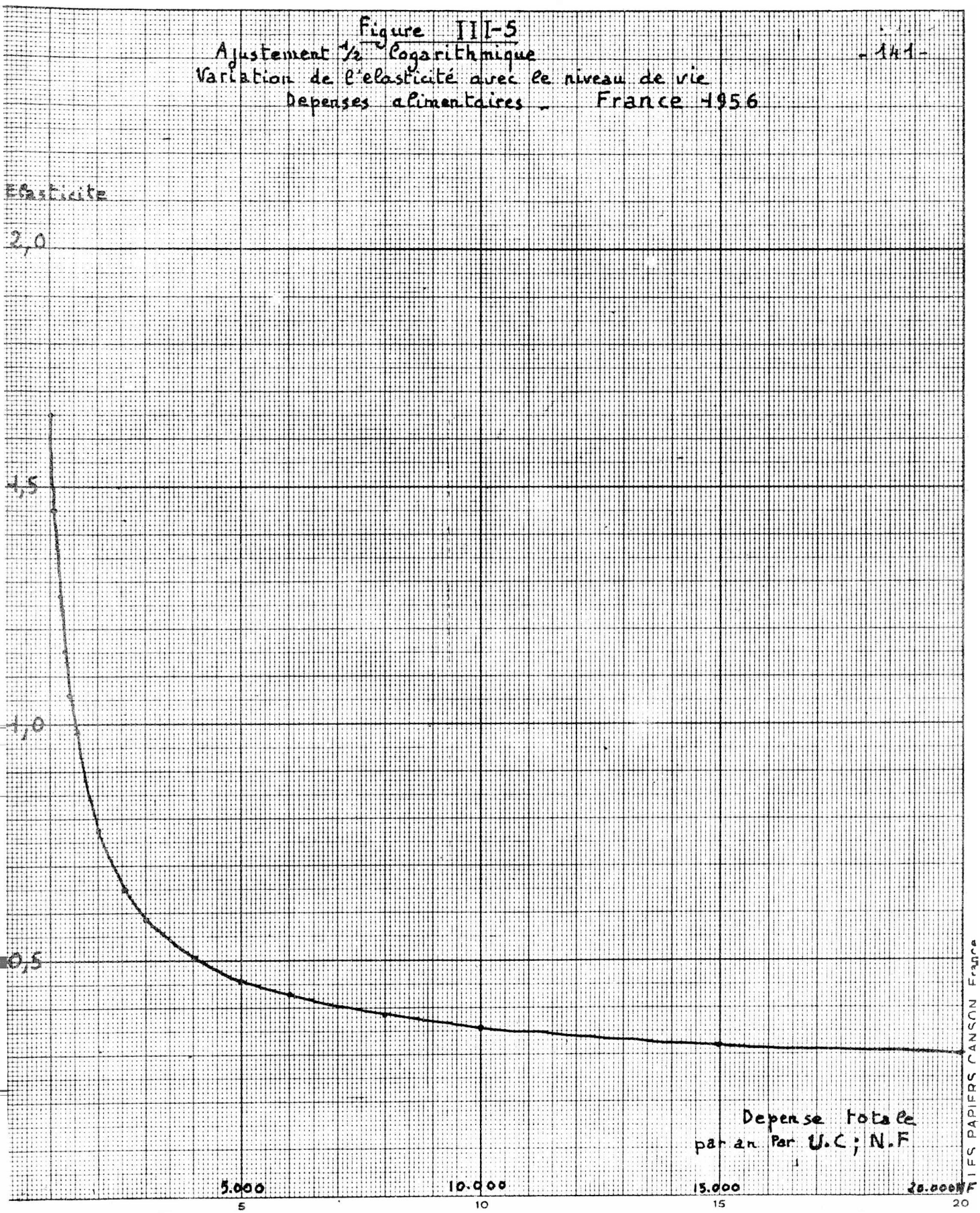
Les ajustements se font de la même façon que pour une loi à élasticité constante, en utilisant du papier semi-logarithmique pour l'ajustement graphique. La valeur de a ne peut cependant pas être lue directement sur le graphique, puisque les échelles ne sont pas les mêmes sur les deux axes de coordonnées. Le moyen le plus rapide est de déterminer a comme la différence de consommation qui correspond à une multiplication par 10 du niveau de vie (23).

Figure III-5

Ajustement ½ Logarithmique

Variation de l'élasticité avec le niveau de vie

Dépenses alimentaires - France 1956



LES PAPIERS CANSON France

La loi semi-logarithmique est particulièrement commode pour les regroupements de produits et les projections. Comme la consommation figure sous forme arithmétique, la loi pour un groupe de produits est la somme des lois pour les différents produits du groupe.

L'agrégation des lois individuelles se fait facilement, comme dans le cas d'une loi à élasticité constante, si l'on admet que tous les revenus augmentent du même pourcentage : comme toutes les consommations augmentent du même montant, il en est de même de la consommation moyenne.

Mais une hypothèse plus générale peut aussi être introduite. Soit $p\ (x)$ la loi de distribution des revenus des consommateurs supposés définis pour $X \geq 0$ et telle que

(II-17) $\int_0^\infty p\ (x)\ dx = 1$. La consommation totale d'un produit donné s'écrit :

(II-18) $$y = N \int_0^\infty yp\ (x)\ dx = N \int_0^\infty (a \log x + b)\ p\ (x)\ dx$$

N étant le nombre total de consommateurs

(II-19) $$y = N \left[a\ E\ (\log x) + b\right]$$

Or,

(II-20) $$E\ (\log x) = \log G\ (x)$$

où $G\ (x)$ est la moyenne géométrique des revenus. La consommation moyenne $\frac{y}{N}$ suit donc la même loi que la consommation individuelle, le revenu individuel étant remplacé par la moyenne géométrique des revenus, quelle que soit la distribution des revenus. Dans le cas où cette distribution est logarithmico-normale, la moyenne géométrique des revenus se confond avec le revenu médian.

C - VERIFICATION DE LA LINEARITE D'UNE LOI

Le choix entre un ajustement à élasticité constante et un ajustement semi-logarithmique, ou tout autre ajustement linéaire après transformation de variables, doit résulter d'une étude sérieuse des données.

Cette étude peut d'abord se faire par une exploration graphique, en utilisant des données groupées. La figure (III-6), montre que, dans le cas étudié, l'examen du graphique suffit nettement à donner la préférence à l'ajustement semi-logarithmique.

Il est cependant utile de pouvoir vérifier ces résultats par des méthodes plus rigoureuses. Plusieurs auteurs ont tenté de le faire en comparant les coefficients de corrélations correspondant aux divers ajustements linéaires retenus. Cette méthode est erronée. En effet, les coefficients de détermination R^2 correspondant aux modèles (II-13) et (II-17) mesurent respectivement le rapport de la variance expliquée par l'ajustement linéaire à la variance des logarithmes des consommations dans le premier cas, et à la variance des consommations elles-mêmes dans le second cas (24). Comme la variance des consommations est plus forte que celle de leurs logarithmes, la comparaison n'a aucune signification.

La méthode convenable consiste à utiliser un des tests habituels de la linéarité d'un ajustement (25). Nous citerons deux de ces tests. Le premier, établi sous des formes voisines par VON NEUMANN (26) et par DURBIN et WATSON (27), consiste à étudier le caractère aléatoire des écarts $\hat{u}_i$ à partir de la distribution du rapport :

(II-21)
$$\frac{\sigma^2}{s^2} = \frac{n}{n-1} \frac{\sum (u_{i+1} - u_i)^2}{\sum u_i^2}$$

FIGURE III-6

RELATION ENTRE LES DÉPENSES ALIMENTAIRES ET LA DÉPENSE TOTALE

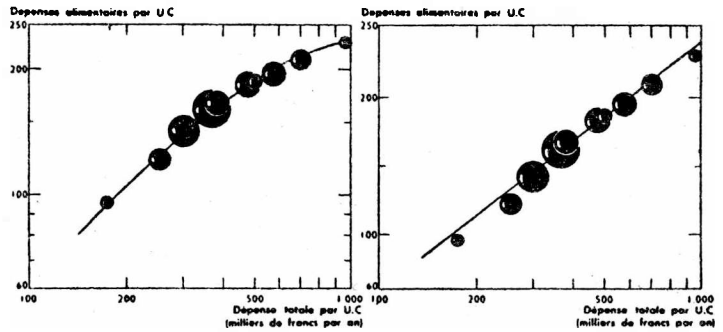
GRAPHIQUE I
Relation logarithmique

$$\log y \doteq a \log x + b$$

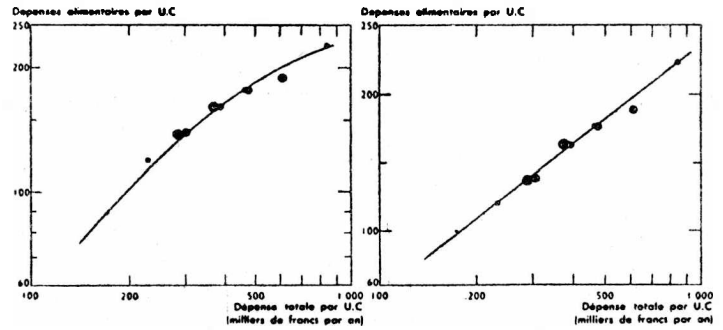
GRAPHIQUE II
Relation semi-logarithmique

$$y = a \log x + b$$

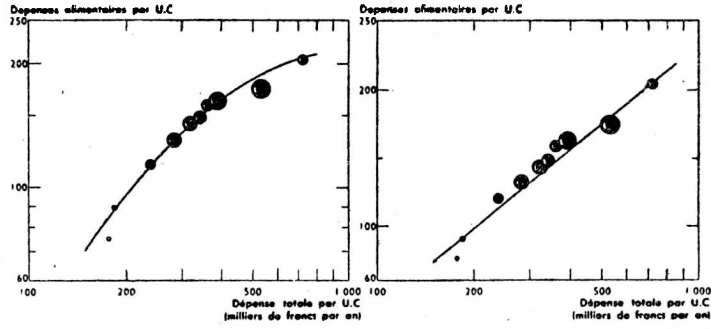
A. — Ménage type 4 : deux adultes, moins de 70 ans



B. — Ménage type 7 : deux adultes et un enfant (6 à 13 ans)



C. — Ménage type 12 : trois adultes



Ce test conduit à des calculs assez longs quand n est grand. Il est surtout utilisé dans un autre contexte, l'étude de l'indépendance des distributions des résidus dans les analyses de séries chronologiques.

Un test non paramétrique d'application plus simple a été décrit par PRAIS et HOUTHAKKER (28) après MOOD (29). Il tient compte uniquement des signes des $\hat{u}_i$, et non de leurs valeurs. Le principe de ce test est illustré par la figure (III-7). Si la relation n'est pas linéaire (cas a), les $\hat{u}_i$ ne changeront de signes que peu de fois. Si, par contre, un ajustement linéaire est convenable (cas b), ils changeront de signes fréquemment.

Pour appliquer ce test, il faut chercher le nombre D de groupes de $\hat{u}_i$ successifs de même signe (30). De tels groupes sont représentés par ① et ② sur la figure (III-7). On utilisera les notations suivantes :

D Nombre de groupes

P Nombre d'écarts positifs

Q Nombre d'écarts négatifs

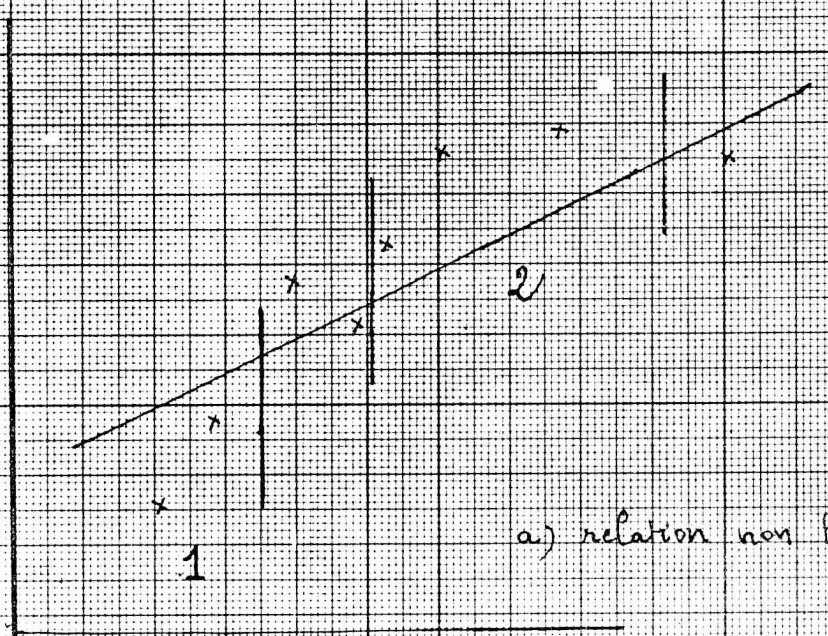
$n = P + Q$

On démontre que, dès que P et Q sont l'un et l'autre supérieurs à une dizaine, pour un ajustement linéaire, D est distribué à peu près normalement avec une moyenne et une variance

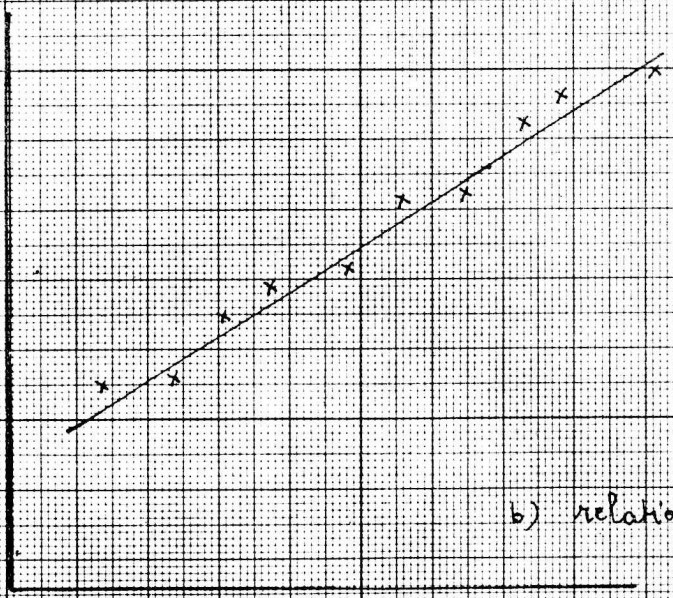
(II-22)
$$E(D) = 1 + 2\,\frac{PQ}{n}$$

$$\sigma_D^2 = \frac{2\,PQ\,(2\,PQ - n)}{n^2\,(n-1)}$$

Figure III - 7



a) relation non linéaire



b) relation linéaire

A partir de ces relations, il est facile de tester la linéarité des données. Il est commode pour cela de considérer leur rapport :

(II-23) $$\lambda = \frac{D - E(D)}{\sigma_D}$$

A titre d'exemple, nous appliquerons ce test à l'ajustement d'une loi semi-logarithmique pour les fruits frais (fig. III-8).

(Voir figure III-8 page suivante)

Dans cet exemple,

$$n = P + Q = 38$$

$$P = 20$$

$$Q = 18$$

$$D = 17$$

$$E(D) = 1 + 2\,\frac{360}{38} = 19,95$$

$$\sigma_D^2 = \frac{720.682}{382.37} = 9,19$$

$$\sigma_D = \sqrt{9,19} = 3,03$$

$$\lambda = \frac{-2,95}{3,03} = -0,975$$

Figure III-8

y
Consommation de fruits frais - Ajustement 1/2 Log.
120
110
100
90
80
70
60
50
40
30
20
10
0
x
800
1000
2000
3000
4000
5000
10000
Consommation totale par UC

On lit dans une table de la loi normale que la probabilité d'observer un écart réduit inférieur ou égal à - 0,975 est supérieure à 15 %. On en conclut donc que l'approximation linéaire ne peut pas être rejetée dans le cas étudié.

On voit que ce test est d'application simple. Il ne peut cependant donner que des résultats peu précis. Comme tous les tests non paramétriques, il est en effet peu efficace, puisqu'il n'utilise qu'une faible partie des informations disponibles. En outre, il ne tient pas compte de la liaison entre les écarts $\hat{u}_i$ qu'introduit l'ajustement d'une droite de régression, alors que cette liaison est prise en considération dans les tables du test de DURBIN-WATSON.

Deux autres difficultés réduisent encore le poids qu'il convient d'attacher aux résultats de ces tests. D'une part, on les applique habituellement à des données groupées, le nombre d'observations variant le plus souvent d'un groupe à l'autre. Cette particularité n'est pas introduite dans la construction des tests. D'autre part, leurs conclusions ne sont valables sans modification que si le niveau de vie R_i est mesuré sans erreur. Or, dès qu'on abandonne les hypothèses simplificatrices sur la distribution des erreurs de mesure que nous avons introduites dans la section 1, on sait qu'il peut en résulter non seulement un biais dans l'estimation des paramètres du modèle, mais une modification de la forme de la courbe. L'absence de linéarité des observations peut donc en partie traduire l'influence des erreurs sur les R_i.

§ 3 - Un exemple de modèle non linéaire des courbes d'ENGEL

Pour certains produits, aucun des deux modèles précédents n'est convenable. L'élasticité diminue avec le niveau de vie, mais moins rapidement que le suppose un ajustement semi-logarithmique. On est, dans ce cas, conduit à utiliser une autre formule.

Plutôt que de chercher un modèle linéaire qui convienne au cas particulier considéré, il est plus commode d'utiliser un modèle non-linéaire qui soit assez souple pour s'adapter à une grande variété de cas.

Le modèle le plus intéressant qui ait été proposé correspond à l'ajustement de la courbe cumulée d'une loi logarithmico-normale. Cette loi a d'abord été appliquée à l'étude de la variation, avec le revenu, du pourcentage de possesseurs d'un type donné de biens durables. Dans ce cas, comme nous l'avons rappelé plus haut (Chap. I Section 5), la loi traduit un schéma probabiliste précis, directement inspiré de l'étude de problèmes de même forme dans le domaine de l'expérimentation biologique.

Par analogie, la même loi a été appliquée par AITCHISON et BROWN (31) à l'ajustement de courbes d'ENGEL. Le choix ne repose plus, dans ce cas, sur un schéma probabiliste précis ; il se justifie seulement par la souplesse de la loi et la connaissance détaillée qu'on a des méthodes d'estimation applicables. Comme les propriétés des distributions logarithmico-normales et les méthodes d'estimation de leurs paramètres ont été étudiées de façon très complète dans un ouvrage d'AITCHISON et BROWN (32), nous présenterons seulement un bref exposé du sujet.

Le modèle suppose l'existence d'un niveau de saturation S_k de la consommation (c'est-à-dire que les courbes d'ENGEL ont une asymptote d'ordonnée k, variant avec le produit considéré). L'hypothèse est d'admettre que le coefficient de saturation $\frac{C_{ik}}{S_k}$ est lié au niveau de vie R_i comme les fractiles d'une distribution logarithmico-normales de paramètres μ et σ^2. On obtient deux modèles différents suivant qu'on admet que les résidus u_i sont additifs ou multiplicatifs :

(II-24) $$C_{ik} = S_k \; P_i \left[\frac{-\mu_k + \log R}{\sigma_k} \right] + u_{ik}$$

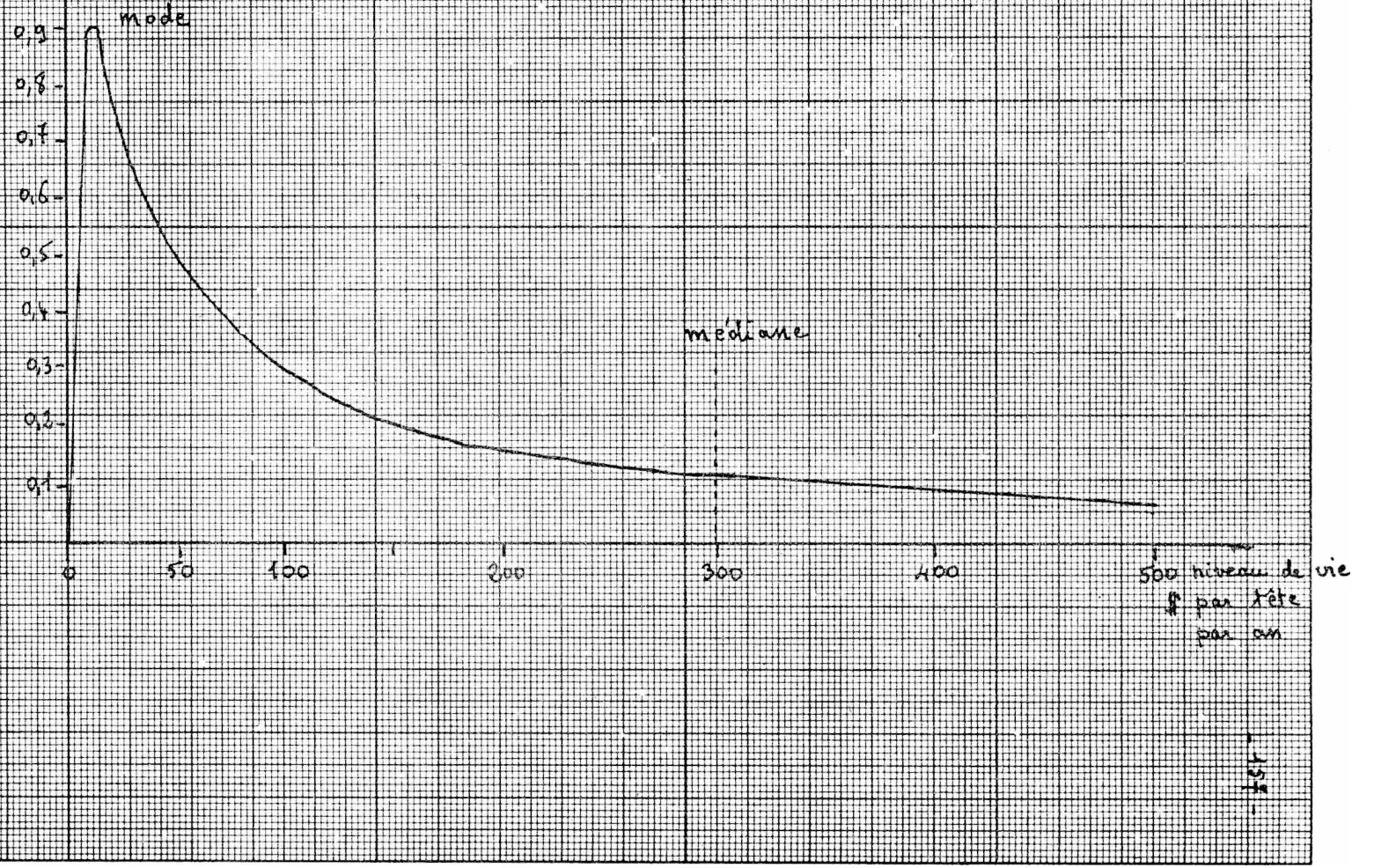
(II-25) $$C_{ik} = S_k \; P_i \left[\frac{-\mu_k + \log R}{\sigma_k} \right] e^{u_{ik}}$$

P_i est l'intégrale de la loi normale standardisée, de moyenne 0 et de variance unité, entre $-\infty$ et $\log R_i$ (33). Ce modèle traduit l'hypothèse que la propension marginale à consommer varie avec le niveau de vie comme les ordonnées de la loi logarithmico-normale de paramètres μ_k et σ_k (figure III-9).

Figure III-9

Propension marginale à consommer en fonction du revenus
(Dépenses alimentaire - JAPON - 1951-55)

Source : GOREUX, op. cit.



La figure (III-10) indique la forme des courbes ajustées pour différents produits. On voit qu'elles correspondent à l'hypothèse représentée dans la figure (III-1), avec la restriction que le seuil R_0 d'introduction des diverses consommations dans le budget est toujours nul... Abandonner cette hypothèse conduirait à remplacer log R par log $(R - R_{0k})$ dans l'équation (II-24). L'estimation simultanée des quatres paramètres serait alors très délicate.

Nous ne considérerons que (II-24). L'autre modèle (qui correspond peut-être mieux à la nature du problème) est étudié par AITCHISON et BROWN (op. cit. pp. 82-84).

Le modèle peut être ajusté graphiquement, par essais successifs, en utilisant du papier à graduations gausso-logarithmiques. Ce papier a une échelle des abscisses graduée en logarithmes et une échelle des ordonnées graduées proportionnellement aux fractiles successifs d'une loi normale : la courbe cumulée d'une loi logarithmico-normale est donc transformée en une droite sur un tel papier.

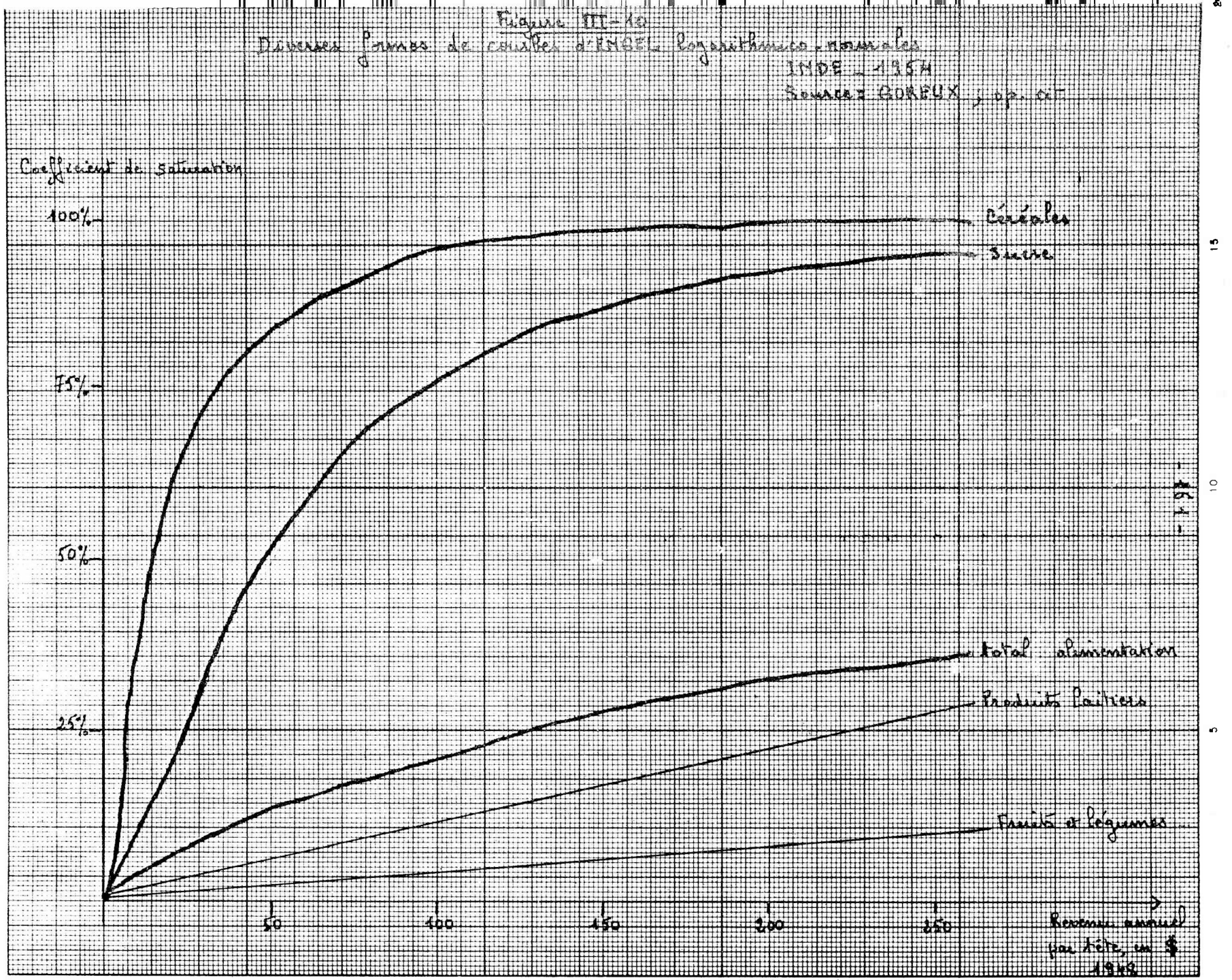
Le graphique ne peut pas être tracé du premier coup, car l'ordonnée k de l'asymptote est inconnue. On doit estimer, à partir d'un graphique ordinaire, un ordre de grandeur S_{ko} de S_k, et établir le graphique sur le papier gausso-logarithmique pour différentes valeurs de S_k voisines de S_{ko}, jusqu'à ce qu'on obtienne des points alignés. Si l'estimation de S_k est trop faible, on obtient en effet une courbe concave vers le haut, et une courbe concave vers le bas si l'estimation de S_k est trop forte..

Figure III-10

Diverses formes de courbes d'ENGEL, logarithmico-normales

INDE - 1956

Source : GOREUX, op. cit.



Les graphiques (III-11 à III-13) donnent un exemple d'ajustement. Il faut remarquer que l'on n'obtient de résultats convenables que si l'on possède des observations assez voisines de la saturation. GOREUX (34) a observé, sur de nombreux exemples, que l'ajustement n'était pratiquement pas possible si l'on ne dépassait pas un coefficient de saturation de 50 % dans le domaine des observations.

Les valeurs de μ et σ se lisent sur le graphique en considérant les abscisses des points pour lesquels le coefficient de saturation est de 0,25 ; 0,50 et 0,75. On a en effet :

$$\text{(II-26)} \qquad \begin{cases} \mu_k = \log_e x_{0,50} \\ \sigma_k = 0,7413 \, (\log_e x_{0,75} - \log_e x_{0,25}) \end{cases}$$

Dans l'étude des résultats d'une enquête britannique, BROWN et AITCHISON (35) avaient observé que σ était très proche de 1 pour la plupart des produits. Cela simplifiait l'ajustement, puisque les observations étaient alignées sur une droite de pente connue. Malheureusement, l'étude d'autres données par VORANGER (36) et GOREUX (37), n'a pas permis de généraliser cette observation.

L'ajustement analytique se fait par le maximum de vraisemblance. Il faut suivre une procédure itérative à partir d'une première estimation S_{ko} de S_k, comme dans l'ajustement graphique. La convergence de l'itération dépend du bon choix de S_{ko}, qui exige lui-même une étude graphique préalable. Les calculs peuvent difficilement être faits sur des machines de bureau et supposent l'utilisation d'un calculateur électronique. Les méthodes d'estimation sont décrites dans AITCHISON et BROWN (38).

FIGURE III-11

FRUITS

Ajustement Log. Normal

$S_K = 100$

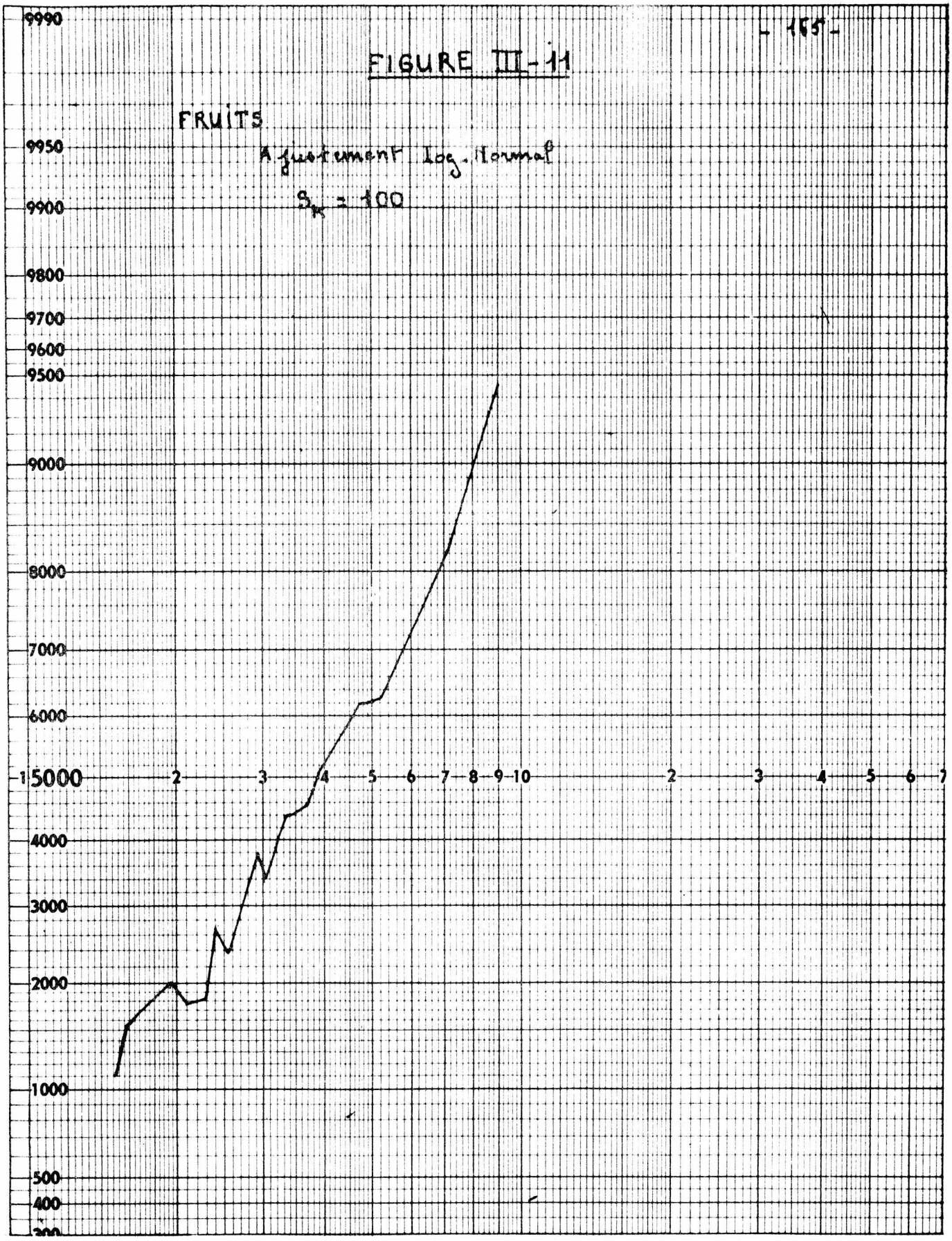


FIGURE III-12

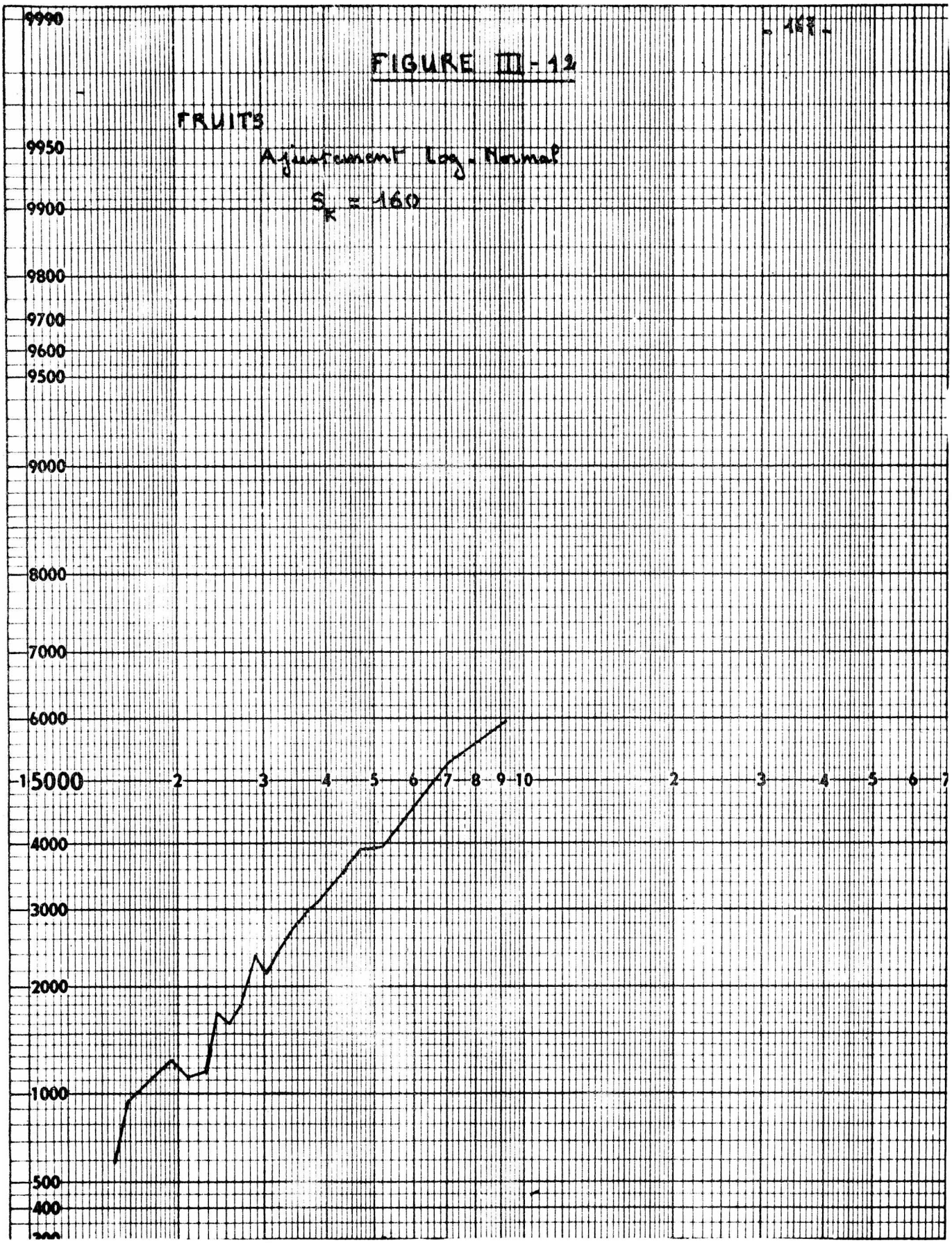
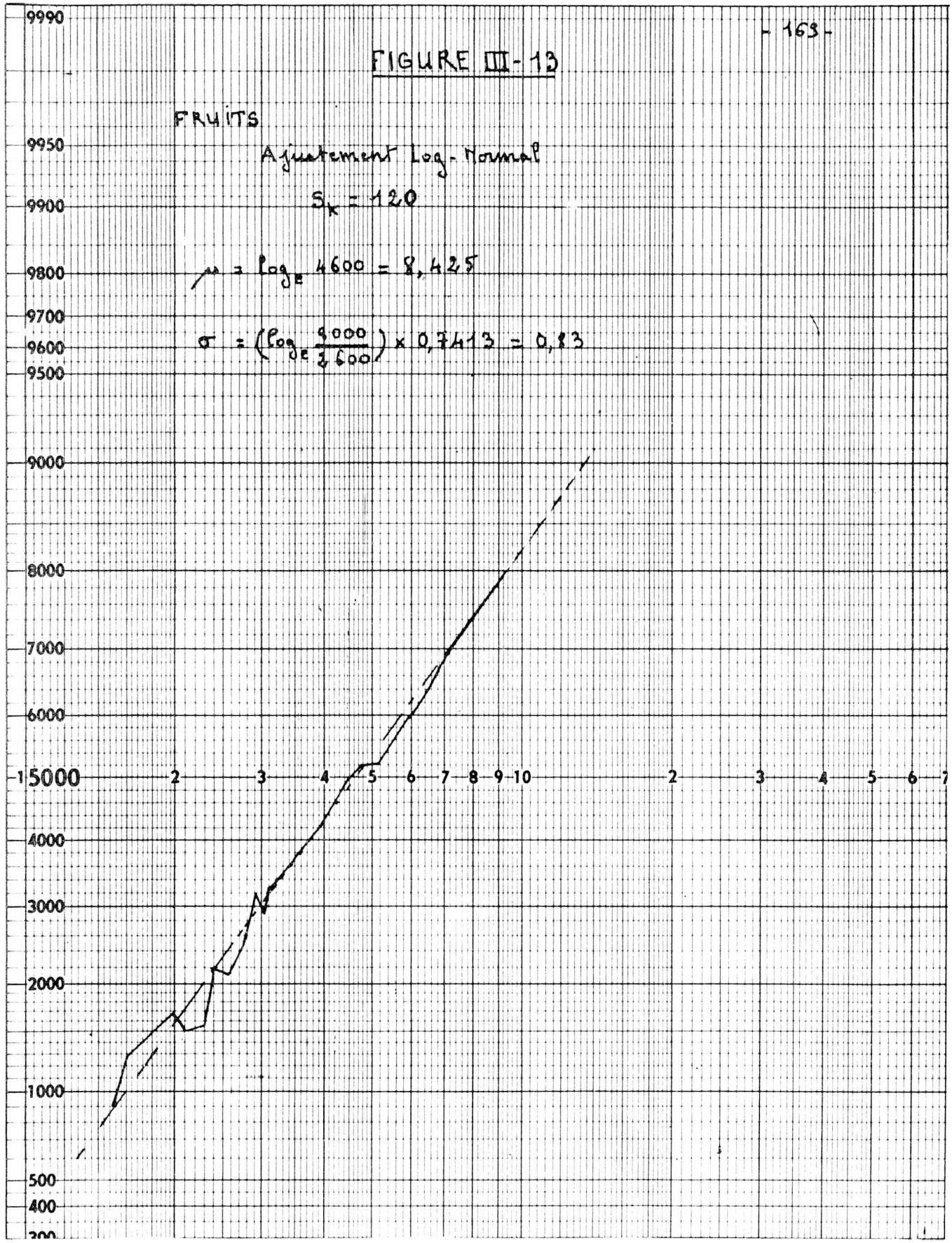


FIGURE III-13

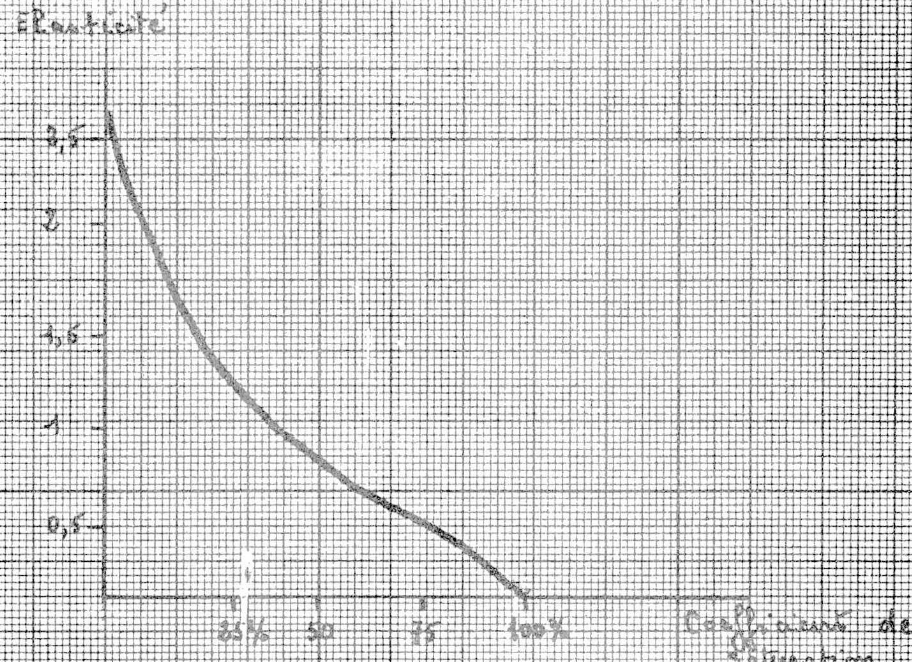


Deux problèmes ne sont pas encore résolus : l'estimation de la variance des paramètres dans le cas de données groupées et l'incidence des erreurs de mesure de R_i sur les résultats.

La figure (III-14) indique comment varie l'élasticité E (f) avec le coefficient de saturation quand $\sigma = 1$. Pour une autre valeur de σ_k, on appliquera la formule :

(II-27) $$E_{\sigma_k}(f) = \frac{E_1(f)}{\sigma}$$

Figure III - 14
Ajustement logarithmico-normal
Variation de l'élasticité avec le coefficient de saturation
pour σ = 1



5
10
15
20

SECTION 3

LA DEFINITION DES VARIABLES INTRODUITES

§ 1 - L'explication par le revenu ou par la dépense totale de consommation

Dans l'étude de la forme des courbes d'ENGEL, nous n'avons pas examiné s'il était légitime d'appliquer les techniques d'estimation habituelles. En particulier, dans le cas de modèles linéaires, nous n'avons pas vérifié si les hypothèses énoncées dans la section 1 de ce chapitre étaient satisfaites. Nous avons seulement éliminé certaines causes particulières de biais en considérant une population homogène.

En même temps, nous n'avons pas précisé la variable explicative, que nous avons appelée "niveau de vie", nous réservant d'en examiner la définition précise jusqu'au présent paragraphe.

En fait, les deux questions sont étroitement liées. La définition de la variable explicative est largement conditionnée à la fois par la nature des données disponibles et par le type de problèmes auquel on veut appliquer les résultats de l'étude (39). Mais les méthodes d'estimation doivent être étudiées en fonction de ce choix. Il s'agit d'un problème particulièrement important qui n'a commencé à être examiné que très récemment.

Pratiquement, le choix pour la définition du niveau de vie est entre le revenu disponible R_i et la consommation totale :

(III-1) $$C_i = \sum_k C_{ik}$$

Le revenu est rarement utilisé, pour des raisons pratiques. En effet, les revenus sont mal saisis dans les enquêtes par sondage. Les enquêtés hésitent à donner des indications précises (et les enquêteurs hésitent à les demander avec assez d'insistance) sur un sujet qui semble particulièrement indiscret. Il en résulte une sous-estimation systématique, même dans le cas des ménages de salariés. Cette sous-estimation provient de l'oubli de revenus annexes (primes, treizième mois, etc...) et d'une forte imprécision dans le relevé des revenus des autres membres de la famille que le chef de ménage, lorsque le questionnaire prévoit de les demander séparément. A ces difficultés générales s'ajoutent des difficultés particulières dans le cas des travailleurs indépendants, et particulièrement des exploitants agricoles pour qui la notion même de revenu est tellement complexe qu'elle semble difficilement saisissable par des enquêtes de consommation. Les revenus sont donc sous-estimés dans une mesure qui varie avec la catégorie socio-professionnelle, mais qui paraît assez systématique à l'intérieur de chaque catégorie.

A ces objections pratiques s'ajoute une objection théorique. FRIEDMAN (40) a en effet suggéré que le comportement du consommateur se définirait en fonction de ce qu'il a appelé le "revenu permanent", qui correspond à la tendance à long terme du niveau de vie. Le revenu instantané en diffère par une composante accidentelle qui est évidemment plus forte, en moyenne, pour les travailleurs indépendants que pour les salariés. NERLOVE (41) a cependant montré que cette notion ne conduisait pas à modifier les méthodes habituelles d'analyse dans le cas des biens de consommation courante.

La plupart du temps, on adopte donc la consommation totale comme variable explicative. Or, si ce choix est légitime, il entraîne des conséquences importantes pour les techniques d'estimation. Deux difficultés se combinent pour aboutir aux mêmes conclusions.

Seule a un sens, d'abord, la consommation totale sur une période assez longue (par exemple l'année). Or les techniques d'enquête aboutissent souvent à relever les consommations sur des périodes assez courtes, soit uniformes (42), soit variables selon la fréquence d'achat des produits (43). Il en résulte que la consommation totale relevée sur ces périodes et étendue à l'année, mesure la consommation totale annuelle avec une erreur accidentelle

qui varie notamment avec l'importance des dépenses exceptionnelles faites par chaque ménage dans la période d'enquête. Le risque d'erreur est d'ailleurs plus fort avec la première méthode d'enquête qu'avec la seconde. Si le revenu est affecté d'une large erreur systématique, les erreurs accidentelles sont vraisemblablement plus fortes que la consommation totale.

Ces erreurs n'affectent pas seulement la précision des ajustements, mais elles biaisent en même temps l'estimation des paramètres.

Un biais résulte aussi d'une autre raison que SUMMERS (44) a mise en lumière : comme la consommation totale C_i est la somme des consommations par produits C_{ik} et que les unes et les autres dépendent stochastiquement du revenu, on doit considérer que C_{ik} et C_i sont des variables endogènes pour le consommateur. Leur relation ne peut donc pas être établie correctement à partir d'un modèle à une seule équation. Il faudrait utiliser le modèle (III-2). Dans ce modèle, les variables sont exprimées en écarts à la moyenne ; R_i désigne le vrai revenu, et C_i et C_{ik} les consommations totales et par produit relevées dans l'enquête (45).

$$\text{(III-2)} \qquad \begin{cases} C_{ik} = \alpha_k \, R_i + u_{ik} \\ C_i = \alpha \, R_i + v_i \\ C_i = \Sigma_k \, C_{ik} \end{cases}$$

La relation entre C_i et C_{ik} s'obtient en éliminant R_i entre les deux premières équations du systèmes :

$$\text{(III-3)} \qquad C_{ik} = \frac{\alpha_k}{\alpha} \, C_i + u_{ik} - \frac{\alpha_k}{\alpha} \, v_i$$

Nous poserons :

$$(III\text{-}4) \qquad \begin{cases} \dfrac{\alpha_k}{\alpha} = \beta_k \\ u_{ik} - \dfrac{\alpha_k}{\alpha} v_i = w_{ik} \end{cases}$$

β_k ne peut pas être estimé par les moindres carrés dans (III-3) puisque v_i et C_i ne sont pas indépendants. On peut calculer le biais :

$$(III\text{-}5) \qquad \hat{b}_k = \frac{\sum_i C_{ik} C_i}{\sum_i C_i^2} = \beta_k + \frac{\sum w_{ik} C_i}{\sum C_i^2}$$

$$(III\text{-}6) \qquad \sum w_{ik} C_i = \sum C_{ik} C_i - \beta_k \sum C_i^2$$

} à partir de (III-3)

En remplaçant C_i et C_{ik} par leurs valeurs en fonction de R_i à partir de (III-2), il vient :

$$(III\text{-}7) \qquad \sum w_{ik} C_i = \sum R_i \ (\alpha u_{ik} - \alpha_k \ u_i) + \sum v_i \ u_{ik} - \beta_k \sum v_i^2$$

L'espérance mathématique du premier terme est nulle, celle des deux autres est :

$$(III\text{-}8) \qquad \text{Cov} \ (C \ U_k) \ - \beta_k \ \sigma_v^2$$

Donc :

$$(III\text{-}9) \qquad E(\hat{b}_i) = \beta_k \left(1 - \frac{\sigma^2_{v_i}}{\sigma^2_{C_i}}\right) + \frac{Cov(v_i\, u_{ik})}{\sigma^2_{C_i}}$$

Le coefficient de β_k est ρ^2_{CR}, carré du coefficient de corrélation entre C et R dans l'ensemble de la population. (III-9) peut s'écrire :

$$(III\text{-}10) \qquad E(\hat{b}_i) = \beta_k \rho^2 + \gamma_k (1 - \rho^2) \text{ où}$$

$$(III\text{-}11) \qquad \gamma_k = \frac{Cov(u_{ik}\, v_i}{\sigma^2_{v_i}}$$

est le coefficient de régression du résidu u_{ik} sur le résidu v_i.

On déduit de III-11 que :

- le biais est nul si $\rho^2 = 1$, c'est-à-dire si $v_i = 0$
- pour un même ρ^2, le biais est vers le bas pour les dépenses courantes (qui sont les plus stables). En effet, pour ces dépenses, Cov $(u_{ik}\, v_i)$ est très probablement négatif.

Le système (III-2) ne permet malheureusement pas d'estimer directement β_k comme rapport de α_k à α. En effet, comme on ne connaît pas le vrai revenu R, on ne peut pas estimer α_k et α.

Deux possibilités se présentent pour éliminer cette difficulté. La première est de travailler sur des données groupées suivant la valeur de R_i, en liant (par les moindres carrés), les valeurs moyennes de C_i et C_{ik} dans chaque groupe (46). Cette méthode se justifie intuitivement : l'erreur dans l'estimation de b_i à partir (III-3) vient de ce qu'un ménage qui a des dépenses exceptionnelles se trouve rangé trop haut dans le classement des ménages. En prenant la moyenne des C_i pour les ménages d'un groupe, on élimine la majeure partie des erreurs accidentelles sur les C_i individuels (47).

Une autre méthode très intéressante et qui peut s'appliquer aux données individuelles a été proposée par LIVIATIAN (48). Elle part de la remarque que R_i, le revenu déclaré par les enquêtés, peut être pris comme variable instrumentale (49), dès qu'on admet (ce qui est légitime) que :

$$\text{(III-12)} \qquad \begin{cases} \text{Cov}\ (C_i, R_i^*) \neq 0 \\ \\ \text{Cov}\ (u_{ik}, R_i^*) = \text{Cov}\ (V_i, R_i^*) = 0 \end{cases}$$

Dans le système (III-2), remplaçons R_i par R_i^*. Montrons que si $\hat{a}_k$ et $\hat{a}$ sont les estimateurs par les moindres carrés de α_k et α dans ces nouvelles équations, $\hat{b}_k = \dfrac{\hat{a}_k}{\hat{a}}$ est estimateur centré de β_k. D'après (III-3) :

$$\text{(III-13)} \qquad \begin{cases} \hat{b}_k = \dfrac{\sum_i C_{ik} R_i^*}{\sum_i C_i R_i^*} = \dfrac{\sum_i (\beta_i C_i + w_{ik}) R_i^*}{\sum C_i R_i^*} = \\ \\ \beta_i + \dfrac{\sum w_{ik} R_i^*}{\sum C_i R_i^*} \end{cases}$$

Sous les hypothèses (III-2), l'espérance mathématique de $\sum w_{ik} R_i^*$ est nulle.

LIVIATIAN a appliqué ces résultats à des données présentées dans PRAIS et HOUTHAKKER (50). R_i^* était dans ce cas le revenu déclaré du chef de ménage. b_k indique l'élasticité calculée par les méthodes habituelles, B_k par la méthode des variables instrumentales (tableau III-C). On voit que l'écart entre les deux estimations est considérable (LIVIATIAN a pu montrer qu'il était significatif dans presque tous les cas) et que son signe est conforme à ce que la formule (III-9) permettait de prévoir.

Les résultats précédents reposent sur des hypothèses qu'il faudrait vérifier par des recherches empiriques plus larges. On doit cependant retenir que l'information que fournissent les revenus déclarés par les ménages doit être utilisés si l'on veut obtenir une estimation correcte des paramètres des courbes d'ENGEL.

TABLEAU III - C

(Source = LIVIATIAN ; op. cit., p. 7)

Produit k	b_k	B_k	$\frac{b_k - B_k}{B_k} \times 100$
- Produits à base de céréales	0,33	0,40	- 17,5
- Produits laitiers	0,26	0,31	- 16,1
- Légumes	0,40	0,47	- 14,9
- Fruits	0,55	0,64	- 14,2
- Poisson	0,57	0,68	- 16,2
- Viande	0,44	0,51	- 13,7
- Loyers	0,49	0,71	- 31,0
- Chauffage, éclairage	0,73	0,81	- 9,9
- Habillement	1,24	1,15	+ 7,8
- Biens durables	1,77	1,66	+ 6,6
- Culture, loisirs	1,05	1,16	- 9,5
- Tabac, alcool	0,61	0,67	- 9,0

§ 2 - L'évolution des dépenses et des quantités consommées

Quel que soit le détail des nomenclatures utilisées pour présenter les résultats d'une enquête de consommation, on ne parvient jamais à isoler des biens économiques homogènes, mais on raisonne toujours sur des groupes de biens plus ou moins étendus. Les résultats de l'enquête britannique de 1937-39, analysée par PRAIS et HOUTHAKKER (op. cit.) sont présentés en 150 postes ; ceux de l'enquête néerlandaise de 1935-36 en 500 postes ; nous avons présenté les résultats de l'enquête française de 1956 en seulement 62 postes (51). Même les plus homogènes de ces postes, comme le pain et le sucre, recouvrent une gamme assez large de variétés de produits.

Un phénomène important a été observé depuis longtemps et étudié avec un soin particulier par PRAIS et HOUTHAKKER (52) en Grande-Bretagne et BLACK (53) aux Etats-Unis : la variation des revenus s'accompagne d'une modification dans la répartition des consommation entre les divers éléments, plus homogènes, qui composent les groupes ainsi définis. Cela correspond à un glissement des achats des consommateurs vers des produits plus coûteux quand le niveau de vie s'élève. Il s'agit là d'une application particulière d'une remarque très générale en matière économique : le progrès, d'un niveau de vie individuel comme d'une économie nationale, ne se traduit pas par une croissance homothétique de tous les éléments de la structure considérée. Il se définit, au contraire, par la substitution de nouvelles activités ou de nouvelles consommations aux anciennes.

En matière de consommation alimentaire, on observe, par exemple, que la ration calorique cesse très tôt d'augmenter avec le revenu. Mais, à ce moment, la structure de la consommation se modifie au bénéfice des protides animales. Ensuite la consommation quantitative se sature, mais on voit augmenter la part de produits coûteux, ayant subi de plus larges transformations par les industries alimentaires ou de conditionnement.

L'application de cette remarque générale à l'analyse des courbes d'ENGEL est simple : les dépenses consacrées à un "produit" (c'est-à-dire, en fait, à un groupe de variétés différentes) augmentent plus vite que la quantité consommée de ce produit. Le prix moyen des produits consommés augmente donc avec le revenu, tous les prix des produits individuels restant constants. Assimilant la qualité d'un produit à son prix, divers auteurs ont appelé, ce phénomène "effet qualité", terme qui est maintenant classique.

On mesure cet effet par une "élasticité qualité" qu'il est facile de définir. Soit d la dépense, p le prix et q la quantité consommée. Ces trois grandeurs sont fonction du revenu R. En prenant les dérivés logarithmiques de :

(III-14) $$d = p\ q$$

il vient :

(III-15) $$\frac{1}{d}\ \frac{dd}{dR} = \frac{1}{p}\ \frac{dp}{dR} + \frac{1}{q}\ \frac{dq}{dR}$$

d'où

(III-16) $$E_d^R = E_p^R + E_q^R$$

L'élasticité qualité, E_p^R, est égale à la différence entre l'élasticité de la dépense et l'élasticité de la quantité consommée par rapport au revenu. Le tableau (III-D) présente quelques résultats obtenus par divers auteurs. Dans le cas de produits alimentaires, la quantité est définie par le poids ou le volume ; dans le cas d'articles d'habillement, par le nombre d'articles achetés.

TABLEAU III - D

Elasticité qualité

Hollande, 1935-1936 (x)	
. Total alimentation	0,28
. Fromage	0,20
. Viande	0,21
. Fruits	0,19
. Légumes	0,20

Angleterre (classe ouvrière) (1937-1938) (x)	
. Farines	0,24
. Lait frais	0,11
. Poisson frais	0,27
. Boeuf et veau	0,19
. Fromage	0,07

France 1955 (xx)

	E_D^R	E_Q^R	E_p^R
. Boeuf	0,51	0,4	0,11
. Veau	0,56	0,5	0,06
. Porc	0,41	0,3	0,11

France 1951 (xxxxx)

. Vin	0,41	0,45	0,11

France 1953 (xxx)

	E_D^R	E_Q^R	E_p^R
. Sous-vêtements féminins	0,7	0,4	0,3
. Chemisiers, corsages	1,2	1,0	0,2
. Chemises d'hommes	0,7	0,4	0,3
. Chaussettes d'hommes	0,9	0,7	0,2
. Cravates	1,8	1,4	0,4
. (xxxx) chaussures d'hommes	0,3	0,2	0,1
. (xxxx) chaussures de femmes	0,5	0,28	0,22

Sources :

(x) - FRAIS et HOUTHAKKER ; op. cit., Economie appliquée, p. 76-77 N°1 1952
(xx) - Estimations personnelles
(xxx) - GOREUX : Perspectives des dépenses d'habillement des français. Etudes et Conjoncture, Oct. 1956
(xxxx) - LISLE et VORANGER : Le marché de la chaussure en France, Annales de Recherches et de Documentation sur la consommation, N°4 - 1957.
(xxxxx) - VORANGER, Consommation N°4 - 1958.

L'utilité de ces résultats est double. D'abord, les courbes d'ENGEL sont utilisées fréquemment pour établir des projections à long terme de la consommation. Observer que l'évolution des dépenses de consommation différera de celle des quantités consommées a des implications considérables pour la politique économique : on a observé aux Etats-Unis, dans les années qui ont suivi la seconde guerre mondiale que les dépenses alimentaires augmentaient assez rapidement avec le revenu, leur élasticité apparente étant supérieure à 0,8, alors que la quantité de produits alimentaires consommés augmentait très peu et que la valeur à la ferme de l'alimentation des américains n'augmentait que médiocrement. Quantifier une telle remarque présente une importance considérable pour fonder une politique de stabilisation des revenus agricoles.

Mais ce n'est pas cette application qui a motivé les travaux de PRAIS et HOUTHAKKER. Leur finalité a été l'analyse de la demande. R. STONE, auprès de qui ils travaillaient, tentait alors d'appliquer systématiquement la méthode des régressions conditionnelles à l'analyse de la demande de produits alimentaires (54).

Cette méthode proposée pendant la guerre par MARSCHAK (55), puis développée par WOLD, STONE (op. cit.), et TOBIN (56), consiste à éviter certaines difficultés dans l'analyse de séries chronologiques en introduisant dans les modèles la valeur de l'élasticité revenu déduite de budgets familiaux. Comme les analyses de séries chronologiques étudient toujours la variation des quantités consommées et que les courbes d'ENGEL étudient plutôt les dépenses de consommation, la démarche de PRAIS et HOUTHAKKER s'introduisait naturellement.

La notion d'élasticité qualité n'épuise cependant pas la description du phénomène qui nous intéresse. Elle ne considère en effet que la variation du prix moyen des produits consommés en fonction du revenu. Or, il est intéressant de considérer d'autres paramètres de la distribution des consommations des diverses variétés. Cette étude a été ébauchée par J. VORANGER dans une note intéressante (57) qui, portant sur la consommation de vin en France en 1951, montre que la dispersion des consommations entre les diverses qualités augmente avec le revenu. Cela vient de ce que les qualités les plus coûteuses ne s'introduisent que quand le revenu augmente, mais que les qualités les moins coûteuses ne disparaissent pas pour autant.

Remarquons que l'ensemble des remarques précédentes doit être interprété avec certaines précautions. L'analyse suppose en effet que le marché soit homogène, afin qu'il existe une relation univoque entre un prix et une variété.

Or, on observe en France des variations assez notables des prix des mêmes produits avec la taille des communes. Comme ces variations sont elles-mêmes en corrélation avec les variations des revenus moyens des consommateurs (qui s'élèvent avec l'importance de l'agglomération), les élasticités qualité peuvent être biaisées pour de nombreux produits. La notion n'est en fait parfaitement claire que pour les produits de marque ayant un marché national.

SECTION 4

L'INFLUENCE D'AUTRES FACTEURS QUE LE NIVEAU DE VIE

§ 1 -- La taille des ménages et les unités de consommation

Les échantillons fournis par les enquêtes de consommation comprennent des ménages de toutes les tailles, depuis les personnes isolées jusqu'aux familles les plus nombreuses. Or, l'effet des variations de la taille des familles ne peut pas être incorporé dans le résidu aléatoire aux courbes d'ENGEL. Il y a à cela deux raisons. D'une part, on observe une corrélation nette entre la taille des familles et leur revenu, comme le montre le tableau III-E. Incorporer la taille des familles dans le terme aléatoire violerait donc l'hypothèse 3 de la section 1 et entraînerait un biais dans l'estimation des paramètres des lois cherchées. D'autre part, la taille des familles est un facteur tellement important qu'il est justifié de le considérer explicitement : pour les produits peu élastiques, les variations des structures familiales ont un effet net plus fort que celui du revenu sur la consommation. Cela peut s'observer sur la figure III-15, où nous avons étudié la variation de la consommation alimentaire par ménage en fonction de la taille du ménage (mesurée par le nombre de personnes) pour une même catégorie de revenus.

Il est donc indispensable de tenir compte explicitement des structures familiales (ou, au moins, de leur composante la plus simple qui est la taille de la famille) dans toute analyse. Les enquêtes qui ne renseignent pas à la fois sur les revenus et la composition des familles sont inutilisables par un économètre.

TABLEAU III - E

RELATION ENTRE LA TAILLE DES MENAGES ET LE REVENU (FRANCE - 1956)

Revenu / Nombre de personnes	0 200	200 300	300 400	400 500	500 600	600 700	700 800	800 1000	1000 1500	1500 & +
1	1 821	686	508	281	131	84	47	79	38	10
2	615	649	743	725	492	459	302	425	370	193
3	54	142	383	490	438	409	279	446	452	228
4	13	23	114	299	308	311	199	284	241	110
5 et plus	7	10	38	141	289	377	376	610	564	305

Ceci étant, il reste à repérer la structure familiale par une variable quantitative. Le plus simple est de prendre, comme dans la figure III-15, le nombre de personnes composant le ménage. Cela se fait parfois (58), mais on utilise plus souvent des "unités de consommation". Les personnes qui composent le ménage étant réparties en plusieurs catégories selon leur âge, leur sexe et éventuellement leur rang dans le ménage, on attribue un poids différent aux individus de chaque catégorie et on fait la somme des résultats obtenus.

Nous citerons deux de ces échelles. La plus simple, utilisée très généralement en France, compte le premier adulte d'un ménage pour 1, les autres adultes pour 0,7 et les enfants de moins de 14 ans pour 0,5 (59). Ainsi, une famille de deux adultes et deux enfants comptera pour 2,7 unités de consommation (au lieu de 4 personnes).

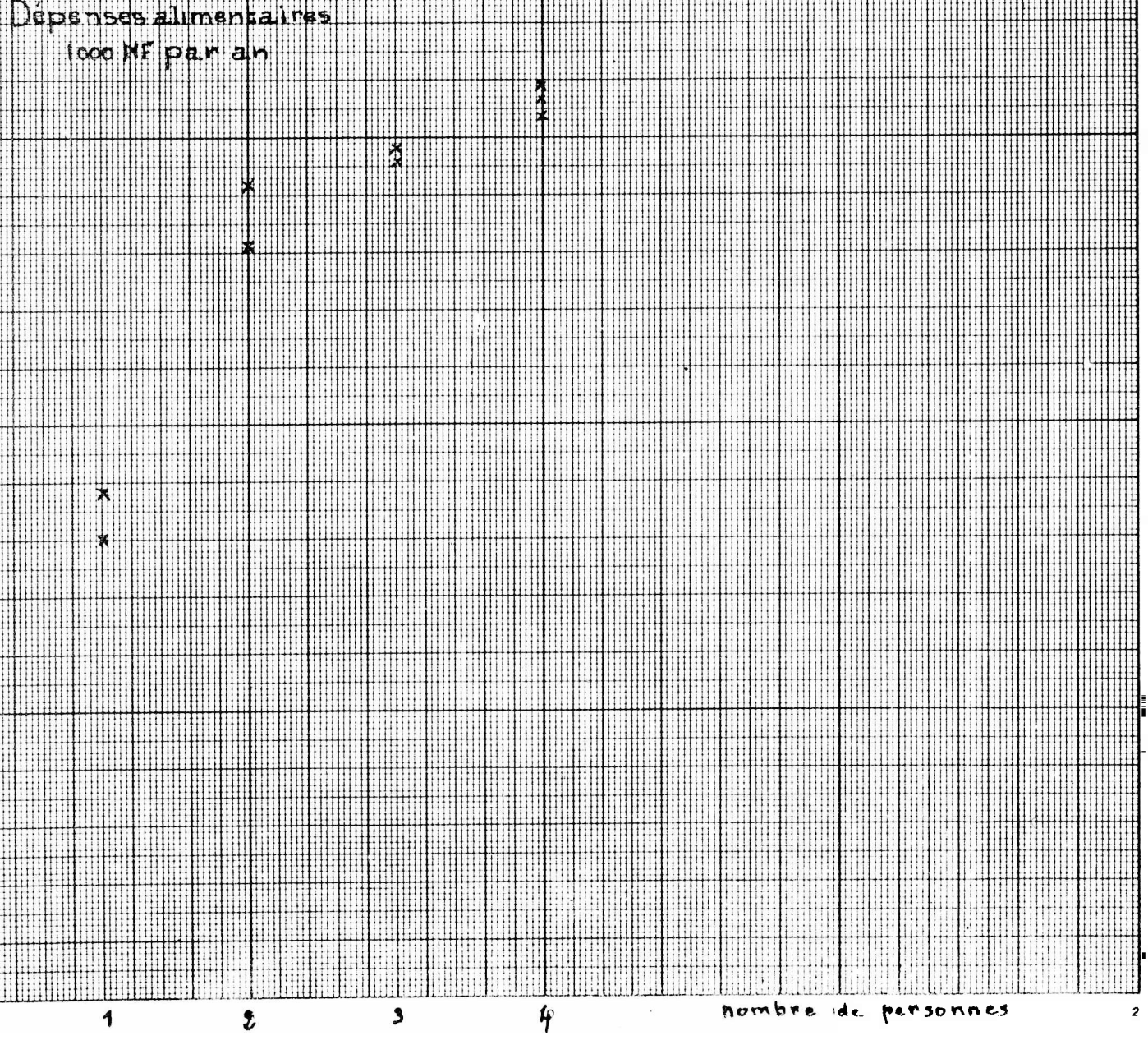
L'échelle suivante, beaucoup plus compliquée, a été utilisée en 1917 aux Pays-Bas :

FIGURE III-15

DEPENSES ALIMENTAIRES EN FONCTION DE LA TAILLE DU MENAGE

POUR UN REVENU DE 6500 NF. PAR AN

(chaque point représente un type de ménage - cf. tableau III-K)



Enfants de	0 à 1 an	0,15	9 - 10 ans	0,65
	1 - 2	0,20	10 - 11	0,70
	2 - 3	0,30	11 - 12	0,75
	3 - 4	0,35	12 - 13	0,80
	4 - 5	0,40	13 - 14	0,85
	5 - 6	0,45	14 - 15	0,90
	6 - 7	0,50	Féminin, 15 et +	0,90
	7 - 8	0,55	Masculin, 15 et +	1
	8 - 9	0,60		

Toutes ces échelles ont des bases entièrement empiriques et leur origine est mal connue. La notion était cependant déjà utilisée par ENGEL, qui avait dénommé les unités de consommation des quets, en hommage à QUETELET. En fait, pour l'utilisation qui est la nôtre, estimer correctement l'influence du revenu sur la consommation, toutes les échelles se valent (60) et la plus simple est la meilleure.

L'utilisation d'unités de consommation est justifiée par le recours au "postulat d'homogénéité". Selon ce postulat, une variation du revenu et de la taille de la famille ont un effet symétrique, l'équation des courbes d'ENGEL devant s'exprimer en termes de consommations et de revenus par unité de consommation.

(IV-1) $$\frac{C_i}{n_i} = f\left(\frac{R_i}{n_i}\right) + u_i$$

Pour examiner ce postulat, nous considérerons uniquement des lois à élasticité constante.

Soit k types de ménages dans la population, chacun étant représenté par la loi

(IV-2) $$\log C_{ik} = a_k \log R_{ik} + b_k$$

Le postulat d'homogénéité exprime que ces k lois peuvent être représentées par une seule relation

$$(IV\text{-}3) \qquad \log \frac{C_{ik}}{n_k} = a \log \frac{R_{ik}}{n_k} + b$$

en termes des variables réduites.

Or, (IV-3) s'écrit

$$(IV\text{-}4) \qquad \log C_{ik} = a \log R_{ik} = (1 - a) \log n_k + b$$

En comparant à (IV-2), on voit qu'une condition nécessaire pour que le postulat soit valable est que les a_k soient les mêmes, aux erreurs d'échantillonnage près, pour tous les types de ménages : les courbes d'ENGEL doivent être parallèles, leurs ordonnées à l'origine étant différentes.

Cette condition est effectivement vérifiée dans la plupart des cas. La vérification n'est cependant pas très efficiente, car l'erreur d'échantillonnage des a_k devient forte quand on considère des sous-populations restreintes. Le tableau III-F donne un exemple des résultats obtenus en même temps que de l'importance des erreurs sur les a_k pour chaque sous-groupe (61).

L'homogénéité des a_k, mise en évidence dans le tableau III-F, fournit une condition nécessaire, mais non suffisante du postulat d'homogénéité. Elle est en effet compatible avec plusieurs modèles différents (62).

$$(IV\text{-}5) \qquad \log C_i = a \log R_i + b \log n_i + c + U_i$$

$$(IV\text{-}6) \qquad \log C_i = a \log R_i + (1 - a) \log n_i + b + U_i$$

$$(IV\text{-}7) \qquad \log \frac{C_i}{n_i} = a \log \frac{R_i}{n_i} + b + U_i$$

TABLEAU III - F

ELASTICITE DE LA DEMANDE D'HABILLEMENT PAR RAPPORT A LA DEPENSE TOTALE

France 1956

Type de ménages	élasticité a_k	σ_{a_k}
Un adulte, moins de 70 ans	1,45	0,13
Un adulte, avec enfants	0,86*	0,16
Un adulte, plus de 70 ans	1,26	0,28
Deux adultes, moins de 70 ans	1,38	0,11
Deux adultes, autres	1,48	0,12
Deux adultes, un enfant de 0-5 ans	1,22	0,11
Deux adultes, un enfant de 6-13 ans	1,19	0,10
Deux adultes, deux enfants 0-5 ans	1,00	0,16
Deux adultes, deux enfants, 6-13 ans	1,31	0,16
Deux adultes, deux enfants autres	1,47	0,20
Deux adultes, trois enfants et plus	1,15	0,16
Trois adultes,	1,36	0,28
Trois adultes, deux enfants et plus	0,95	0,11
Quatre adultes, et plus	1,20	0,09
Quatre adultes et plus, + enfants	1,03*	0,12
ENSEMBLE	1,35	0,04

* Coefficients a_k significativement différents de ceux pour l'ensemble.

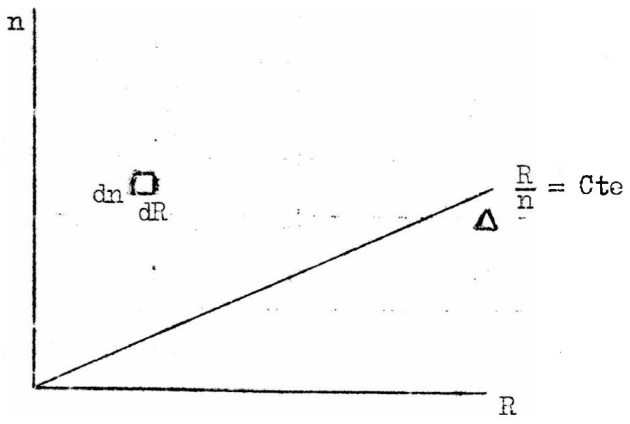
$$(IV\text{-}8) \quad \log \frac{C_i}{n_i} = a \log \frac{R_i}{n_i} + b \log n_i + c + U_i$$

(IV-6) et (IV-7) sont équivalents du point de vue de leur partie non aléatoire. On voit qu'ils représentent un cas particulier de (IV.- 5). Le nuage de points doit être proche d'un plan passant par l'origine dans le second cas, au lieu d'un plan quelconque dans le premier. Cela suppose que, pour tous les types de ménages, on ait :

$$(IV\text{-}9) \quad (1 - a) \log n_k + b_k = Cte$$

Cette condition est rarement vérifiée directement, mais on peut la vérifier indirectement. Il suffit que le a estimé pour l'ensemble de la population à partir du modèle (IV-7) soit égal, aux erreurs d'échantillonnage près, à la valeur commune des a_k dans les types de ménages. Cela est approximativement vérifié dans le tableau (III-F).

Cependant, (IV-6) et (IV-7) ne sont pas équivalents si l'on considère la part stochastique du modèle.



Pour que l'ajustement par les moindres carrés soit convenable, il faut que la moyenne des résidus soit nulle et leur variance constante dans chaque prisme rectangulaire de base dn, dR et perpendiculaire au plan nR dans le cas du modèle (IV-6). Il faut que cette propriété soit vérifiée pour chaque plan $\frac{R}{n}$ = Cte de trace Δ sur le plan (n R) dans le cas du modèle (IV-7).

Le modèle (IV-8) a été introduit par plusieurs auteurs qui utilisent le nombre de personnes au lieu du nombre d'unités de consommation pour définir les variables réduites. Cette réduction exagère la mesure de l'influence de la taille du ménage et exige une correction par un terme en log n, dont le coefficient est en général négatif. Comme la taille du ménage est déjà prise en compte dans la réduction, le coefficient de log n mesure les autres influences liées à la structure familiale : "l'élasticité par rapport à la taille du ménage traduit les économies d'échelle dans (la consommation) des familles les plus grandes et les variations de la consommation par tête liées aux différences dans la composition de la famille par sexe et âge" (63).

Nous donnons, dans le tableau III-G, quelques résultats obtenus par ROCKWELL à partir de l'enquête américaine de 1951.

Le modèle (IV-5) qui serait le plus naturel est rarement utilisé. En effet, la corrélation entre R et n étant assez forte par rapport à celle entre R et C, il n'est pas possible de séparer convenablement les coefficients de log R et log n. Notons cependant que GOREUX (op. cit. p. 38) estime que le coefficient de log n est environ trois fois plus fort, en valeur absolue, quand on utilise des données par ménage que lorsqu'on utilise des données par tête. Or le terme en log n étant un facteur correctif, il est bon qu'il soit aussi faible que possible. Ceci conduit à utiliser des données par unité de consommation plutôt que par tête.

TABLEAU III - G

ELASTICITES PAR RAPPORT AU REVENU ET A LA TAILLE DU MENAGE

Source : ROCKWELL, P. 8-11 et 36-39

Ménages non agricoles de revenu moyen (°)

Modèle : $\log \frac{C}{n} = a \log \frac{R}{n} + b \log n + d$

	a	b
Dépense alimentaire totale	0,21	- 0,26
Produits laitiers	- 0,04	- 0,27
Matières grasses	0,15	- 0,24
Viandes, volailles, poisson	0,31	- 0,25
Légumes frais	0,66	- 0,18
Fruits frais	0,14	- 0,42
Boissons	0,51	- 0,30
Produits sucrés	0,15	- 0,27

(°) - Pour tenir compte de la diminution de l'élasticité avec le niveau de vie, ROCKWELL a ajusté séparément des modèles à élasticité constante pour les ménages de revenu faible, moyen et élevé. Cela revient à représenter les courbes d'ENGEL par une ligne brisée en échelle logarithmique.

C.R.E.D.O.C.
BIBLIOTHÈQUE

§ 2 - Une tentative de détermination statistique des unités de consommation

Les échelles de réduction introduites dans le paragraphe précédent sont empiriques. En les utilisant, on se trouve recourir à une recette, à un tour de main, dont l'expérience montre qu'il donne des résultats convenables sans qu'on puisse vraiment expliquer pourquoi.

Cette situation n'est pas satisfaisante ; il n'est donc pas étonnant qu'on se soit préoccupé de déterminer la valeur des unités de consommation à partir d'une analyse du comportement observé des ménages de différentes tailles. La tentative la plus intéressante a été faite par PRAIS et HOUTHAKKER, à partir d'une hypothèse proposée par ALLEN (64).

L'idée sur laquelle repose la méthode est voisine de celle introduite par HICKS (65) pour analyser l'influence d'une variation des prix sur la consommation. L'addition d'une personne à une famille dont le revenu nominal ne change pas entraîne un double effet. D'une part, le niveau de vie de cette famille diminue (c'est l'équivalent de l'effet de revenu de HICKS) ; d'autre part, même si l'on compensait cette diminution de revenu réel, la consommation des divers produits se trouverait quand même modifiée. En effet, le changement de la composition de la famille entraîne une modification de ses goûts et donc une substitution entre produits.

Pour distinguer ces deux effets, ALLEN a proposé d'utiliser des échelles différentes pour réduire le revenu et la consommation d'un produit. La première échelle, dite échelle de niveau de vie est la même pour tous les produits, l'autre, dite échelle spécifique, varie d'un produit à l'autre. Bien entendu, l'analyse des ménages en catégories homogènes selon l'âge et le sexe est la même pour les deux échelles.

Le modèle proposé s'écrit, j étant un produit et i une catégorie de personnes :

$$\text{(IV-10)} \qquad \log \frac{C_i}{\sum_i \lambda_{ik} n_i} = a_j \log \frac{R}{\sum \lambda_i n_i} + b_j$$

Le problème traité par les auteurs est d'estimer simultanément l'élasticité a_j et les unités spécifiques λ_{ik} en se donnant l'échelle des niveaux de vie λ_i^j. Bien que le modèle ne soit pas linéaire en termes des paramètres, leur estimation par le maximum de vraisemblance ne pose pas de problèmes particuliers (66). Nous donnons dans le tableau III-H quelques exemples de résultats obtenus par PRAIS (67) et par VORANGER (68).

Le problème est seulement à moitié résolu. Il resterait à déterminer également l'échelle de niveau de vie à partir des données. PRAIS et HOUTHAKKER (69) ont montré que, si une échelle spécifique était estimée pour tous les postes composant le budget, l'échelle de niveau de vie devrait être une moyenne pondérée des échelles spécifiques, avec des poids voisins des coefficients budgétaires des différents produits. Une première estimation de l'échelle de niveau de vie étant donnée (comme dans le tableau III-H), on devrait pouvoir corriger cette estimation par itération. Aucun résultat n'a été publié à ce sujet et il semble bien que les tentatives effectuées dans ce sens aient abouti à un échec. Le problème posé dans ce paragraphe n'a pas encore été résolu complètement.

Cela n'est pas étonnant. Il existe de nombreux postes du budget pour lesquels, non seulement les unités de consommation, mais les courbes d'ENGEL elles-mêmes sont dépourvues de signification. Cela semble exclure la possibilité de définir une échelle de niveau de vie par la méthode suggérée.

TABLEAU III - H

ECHELLES SPECIFIQUES D'UNITE DE CONSOMMATION

A - Royaume-Uni 1937-38 (Ouvriers) PRAIS

Individus	Echelle de niveau de vie	Echelles spédifiques			Légumes
		Total alimentation	Viande	Produits laitiers	
Homme	1	1	1	1	1
Femme	1	0,86	0,68	0,96	0,89
Macs. 14 - 17 ans	1	0,82	0,54	0,85	1,05
Fem. 14 - 17 ans	1	0,65	0,47	0,69	0,73
Enfant 10 - 13 ans	1	0,69	0,48	0,69	0,85
Enfant 5 - 9 ans	1	0,55	0,31	0,54	0,77
Enfant 1 - 4 ans	1	0,51	0,26	0,58	0,74
Enfant 0 - 1 an	1	0,32	0,05	0,75	0,33

B - France 1951 (VORANGER)

	Niveau de vie	Total alimentation	Viande
Hommes 20 - 59 ans	1	1	1
Hommes 60 ans et plus	0,80	0,81	0,63
Femmes 20 - 59 ans	0,80	0,77	0,69
Femmes 60 ans et plus	0,70	0,76	0,52
Enfants 5 - 4 ans	0,40	0,36	0,16
Enfants 5 - 13 ans	0,50	0,47	0,32
Enfants 14 - 19 ans	0,70	0,66	0,69

§ 3 - Variables qualitatives affectant les courbes d'ENGEL

D'autres facteurs que les structures familiales peuvent introduire des hétérogénéités dans les lois de consommation. On a, par exemple, observé dans tous les pays que les lois n'étaient pas les mêmes pour les ménages agricoles et non agricoles (70). Comme il existe une différence notable entre les revenus agricoles et non agricoles, on obtient une estimation biaisée des lois cherchées en étudiant l'ensemble de la population.

Nous examinerons concrètement l'influence de certains facteurs qualitatifs dans le chapitre suivant. Le sujet ne justifie pas, contrairement à l'étude de la taille des ménages, de longs développements théoriques. Les seuls problèmes réels sont liés aux tests statistiques permettant de vérifier si différentes sous-populations sont homogènes (71).

L'interprétation des résultats obtenus peut cependant être assez délicate. Considérons plusieurs sous-populations caractérisées par l'indice k. Soit :

(IV-11) $$\text{Log } C = a_k \log R + b_k$$

l'équation des courbes d'ENGEL pour un produit donné dans la strate k. Trois possibilités peuvent se présenter :

a. La population est homogène :

Aux erreurs d'échantillonnage près,

$$a_k = a \qquad b_k = b \quad \text{quelque soit } k.$$

b. Les a_k sont homogènes, mais pas les b_k.

Les courbes d'ENGEL sont parallèles mais décalées (comme dans le tableau III-E). On peut, dans ce cas, définir sans ambiguité une élasticité pour l'ensemble de la population, mais on en obtiendra une estimation biaisée à partir des données d'ensemble. Il convient de l'estimer comme la valeur commune des élasticités estimées dans les strates.

Dans le cas où il n'y a que deux strates, on peut cependant traiter l'ensemble des données à partir du modèle :

$$\log C = a \log R + b x + d \qquad \text{(IV-12)}$$

où x est une variable auxiliaire égale à 1 pour une strate et nulle pour l'autre.

c. Enfin, les a_k et les b_k peuvent être hétérogènes. Dans ce cas, il n'y a aucun sens, en toute rigueur, à définir une élasticité moyenne pour l'ensemble de la population. On peut cependant le faire de façon approximative pour des applications pratiques dans le cas où les a_k ne sont pas trop différents et où ce sont les différences entre les b_k qui pèsent le plus. On est alors assez près du cas b).

Il n'est pas toujours facile d'interpréter l'hétérogénéité des a_k. En effet, les revenus moyens des différentes strates ne sont habituellement pas les mêmes et l'hétérogénéité observée peut traduire une erreur dans la spécification de la forme des courbes d'ENGEL autant que l'influence du facteur qualitatif considéré. On ne pourrait éliminer cette difficulté qu'en disposant d'un échantillon choisi de telle façon que ce facteur et le revenu n'y soient pas en corrélation l'un avec l'autre (72).

2ème PARTIE : RESULTATS DEDUITS DE L'ENQUETE FRANCAISE DE 1956

SECTION 1

LES COURBES D'ENGEL POUR L'ENSEMBLE DE LA POPULATION NON AGRICOLE

§ 1 - Introduction

Nous présenterons dans cette partie quelques résultats déduits de l'analyse de l'enquête que nous avons décrite dans le chapitre II. L'étude que nous avons pu faire jusqu'à présent de cette enquête n'a exploré que quelques unes des pistes reconnues dans la partie précédente. Les résultats de l'exploitation mécanographique de l'enquête n'ont, en effet, été disponibles qu'au milieu de 1959 et la première tâche a consisté à les interprêter selon des méthodes classiques et éprouvées, afin de dégager des résultats pouvant aider à la préparation de la politique économique du gouvernement ; en particulier, ces résultats ont été utilisés pour établir les perspectives de développement économique préparatoires au quatrième Plan de modernisation et d'équipement.

Nous avons largement utilisé des méthodes graphiques dans les premières phases de cette analyse. Bien que ces méthodes ne donnent pas de résultats numériques précis et ne permettent pas de jugements en probabilité, leur utilité est considérable. Elles permettent, en effet, une exploration rapide assez sûre de certaines questions traitées dans la première partie, comme la forme des courbes d'ENGEL et la légitimité de la réduction en unités de consommation. Elles sont aussi très utiles pour orienter les recherches ultérieures, en suggérant les questions les plus pertinentes auxquelles une analyse statistique permettra de répondre. Ces méthodes permettent donc une économie notable de temps et d'argent dans le déroulement des recherches.

Nous considérerons uniquement la population non-agricole dans les pages qui suivent. Nous avons étudié ailleurs le comportement de la population agricole, définie comme l'ensemble des ménages dont le chef est exploitant ou salarié agricole (73). Les résultats de cette étude ont été beaucoup moins satisfaisants que ceux que nous présentons ici, car les paramètres des lois obtenues sont affectés d'une erreur d'échantillonnage qui est en général assez forte. Cela traduit surtout l'incidence de l'autoconsommation de produits alimentaires provenant de l'exploitation qui varie considérablement d'un ménage à l'autre sans être nettement liée au revenu. Les premiers résultats de travaux en cours semblent indiquer que la consommation alimentaire de la population agricole ne pourrait être analysée correctement qu'en introduisant, outre le revenu, la région et des facteurs liés à la structure de l'exploitation qu'une enquête de budgets familiaux ne permet pas de repérer.

§ 2 - La méthode d'analyse (74)

On a trié l'échantillon à la fois selon les structures familiales et selon les revenus déclarés par les ménages ou estimés par les enquêteurs. 16 types de ménages ont été définis de façon à ce que le nombre d'unités de consommation par ménage varie peu à l'intérieur de chaque type et que le nombre de ménages de chaque type varie assez peu. En général, dix catégories de revenu ont été retenues ; la première correspond aux revenus inférieurs à 200 000 Frs par an, la dernière aux revenus supérieurs ou égaux à 1 500 000 Frs par an. Dans trois cas particuliers, les ménages d'une personne (types 1 et 3) et les ménages de deux personnes dont au moins une de plus de 70 ans (type 4), une trop grande partie de l'échantillon se trouvait dans la catégorie la plus basse ; on a, dans ce cas, divisé cette catégorie en deux par une limite placée à 100 000 Frs par an. Le tableau III-K définit les types de ménages et les catégories de revenu adoptées.

Les ménages ont été ainsi répartis en 163 cellules. Pour chaque cellule, on a obtenu par tabulation la somme des consommations de tous les ménages de la cellule pour chaque rubrique du questionnaire, et les résultats obtenus ont été perforés dans des cartes récapitulatives. Ce regroupement a permis de faire les calculs ultérieurs à partir d'environ 6 000 cartes récapitulatives seulement au lieu de près de 350 000 cartes originales.

TABLEAU III - K

DEFINITION DES CRITERES DE TRIS

Types de ménages

1 un adulte de moins de 70 ans
2 un adulte de moins de 70 ans avec enfants
3 un adulte de plus de 70 ans
4 deux adultes de moins de 70 ans
5 deux adultes, autres
6 deux adultes, un enfant de 0 à 5 ans
7 deux adultes, un enfant de 6 à 13 ans
8 deux adultes, deux enfants de 0 à 5 ans
9 deux adultes, deux enfants de 6 à 13 ans
10 deux adultes, un enfant de 0 - 5 ans, un enfant 6 - 13 ans
11 deux adultes, trois enfants et plus
12 trois adultes
13 trois adultes, un enfant
14 trois adultes, deux enfants et plus
15 quatre adultes et plus, sans enfants
16 quatre adultes et plus, avec enfants

Catégories de revenu (borne supérieure exclue)

1	0 — 0 - 100 000 F.	6	600 000 - 700 000 F.
	1 — 100 000 - 200 000 F.	7	700 000 - 800 000 F.
	2 — 200 000 - 300 000 F.	8	800 000 - 1 000 000 F.
	3 — 300 000 - 400 000 F.	9	1 000 000 - 1 500 000 F.
	4 — 4000000 - 500 000 F.	10	1 500 000 F. et plus
	5 — 500 000 - 600 000 F.		

Les cartes récapitulatives ont été introduites dans un calculateur électronique afin d'obtenir, pour chaque cellule, la valeur de la consommation totale par unité de consommation, prise comme mesure du niveau de vie, et de la consommation par unité de consommation de chacun des groupes de produits définis dans les tableaux III-N et III-P. Ces données ont servi de base à l'ajustement analytique des lois décrites sur ces tableaux, en utilisant le même calculateur électronique. L'ajustement a donc été fait sur des données groupées de la façon que nous avons considérée comme correcte dans la troisième section de la partie précédente. Au cours du calcul, les observations correspondant à chaque cellule ont été pondérées par le nombre de ménages compris dans cette cellule.

Pour pouvoir faire systématiquement les ajustements analytiques, on a exploré graphiquement en détail la forme des lois en établissant des graphiques analogues à la figure III-6 de la partie précédente. Cette exploration a conduit à choisir un ajustement semi-logarithmique pour les consommations alimentaires et les dépenses de chaussures et de tabac, et un ajustement à élasticité constante dans les autres cas.

On a également étudié graphiquement la validité de la réduction en unités de consommation. Pour les principaux produits, on a porté les observations correspondant à chaque type de ménages sur un papier calque à graduation doublement logarithmique, et on a superposé les graphiques obtenus après les avoir décalés d'une mesure correspondant à une division par le nombre moyen d'unités de consommation dans le type de ménages correspondant. On a vérifié que le nuage d'ensemble obtenu était très homogène.

§ 3 - Résultats

Nous présentons les résultats dans les tableaux III-N et III-P. Le tableau III-N correspond aux ajustements semi-logarithmiques. Il indique, pour chaque groupe de produits, l'équation ajustée, l'écart-type du coefficient de log x, la dépense moyenne par unité de consommation et la valeur de l'élasticité par rapport au revenu pour trois valeurs du niveau de vie (75) dont l'une correspond à la moyenne pour l'ensemble de la population étudiée.

Le tableau III-P correspond aux ajustements à élasticité constante. Il donne, pour chaque groupe de produits, la dépense moyenne par unité de consommation, le coefficient d'élasticité et son écart-type. Nous commenterons successivement les résultats donnés dans ces deux tableaux.

a - Ajustements semi-logarithmiques

Cet ajustement a été choisi pour les dépenses alimentaires analysées en 35 groupes, les repas à la cantine et au restaurant, les consommations prises au café, l'autoconsommation évaluée aux prix de détail, et les achats de chaussures et de tabac.

Après une discussion sur les paramètres a et b, nous considérerons les variations de l'élasticité en fonction du niveau de vie pour les principaux produits ou groupes de produits.

On remarque sur le tableau III-N que les paramètres a et b sont de signe contraire ; ce résultat pouvait être prévu. En effet, a et b ne peuvent pas être tous les deux négatifs, car il doit exister des valeurs positives de la consommation pour au moins une partie du domaine de variation de R observé ; le fait que a et b ne sont pas tous les deux positifs signifie simplement que la consommation d'un produit à élasticité positive n'apparaît qu'à partir d'un seuil de revenu non nul.

Deux cas sont donc possibles, suivant que a est positif ou négatif, b ayant toujours le signe contraire de a. Deux remarques sont valables dans les deux cas. D'une part, l'élasticité décroît toujours avec le niveau de vie, puisqu'elle est égale à :

$$E = \frac{a \log_{10} e}{C}$$

L'élasticité est inversement proportionnelle au niveau de la consommation ; sa variation avec R est représentée par une branche d'hyperbole équilatère en coordonnées semi-logarithmiques. D'autre part, on ne peut pas utiliser directement la valeur de a pour comparer l'influence du revenu sur la consommation de différents produits, comme dans le cas d'une loi à élasticité constante. En effet, a qui dépend des unités choisies, n'a pas de signification. Par contre, le rapport $-\frac{b}{a}$ varie comme l'élasticité pour une valeur donnée du niveau de vie. On en déduit que les hyperboles en log R correspondant à des produits différents ne se coupent pas.

Le cas où a est positif est le plus général. La consommation et l'élasticité varient avec log R d'une façon qu'indique la figure III-16 : le produit considéré s'introduit dans le budget pour un revenu égal à $-\frac{b}{a}$. Ce seuil est d'autant plus élevé que le produit est plus sensible au revenu. Les dépenses de restaurant s'introduisent pour une dépense totale de l'ordre de 220 000 Frs par unité de consommation, celles de volaille ou de fruits vers 150 000 Frs, alors que les dépenses de vin apparaissent à partir de 72 000 Frs et celles de porc vers 30 000 Frs (en fait, tous les groupes de ménages étudiés ont des niveaux de vie supérieurs aux deux derniers seuils).

L'élasticité est positive et décroît d'autant plus rapidement avec le revenu qu'elle est plus élevée. Sauf dans le cas de produits fréquemment consommés dont l'élasticité varie assez peu avec le revenu et est toujours faible (produits comme le boeuf, le vin, le porc, le beurre, etc...), la valeur de l'élasticité est peu significative pour des valeurs extrêmes du niveau de vie, auxquelles correspondent peu d'observations. Les lois ajustées s'appliquent surtout pour des valeurs du niveau de vie comprises entre 200 000 et 600 000 Frs.

Le cas où a est négatif correspond aux produits pauvres qui représentent environ 20 % de la consommation alimentaire : autoconsommation, achats de lard et de saindoux, de pain, de margarine et de sucre (pour ces deux derniers produits l'ajustements est peu précis). Le figure III-17 représente la variation de la consommation et de l'élasticité en fonction du logarithme de la dépense totale.

La consommation décroît quand le revenu augmente et s'annule pour une valeur du niveau de vie égale à $-\frac{b}{a}$.

FIGURE III-16

$a > 0$

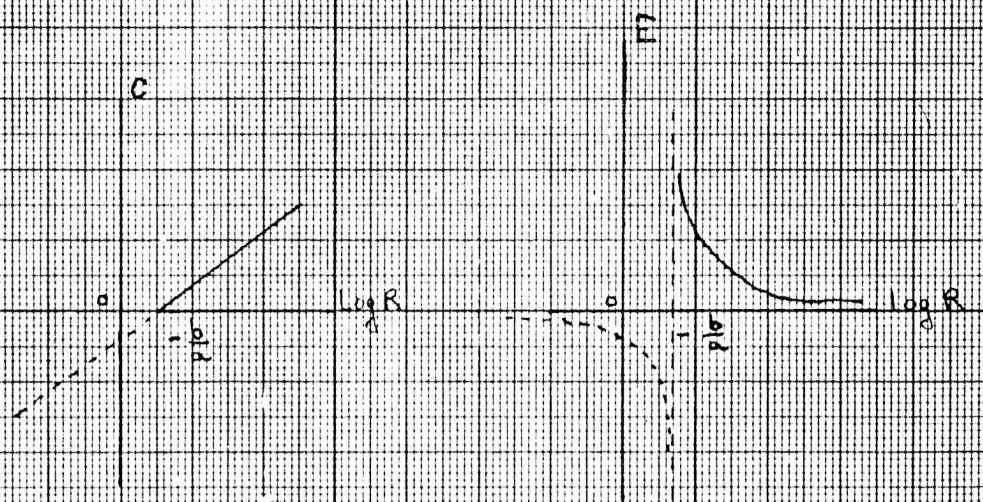
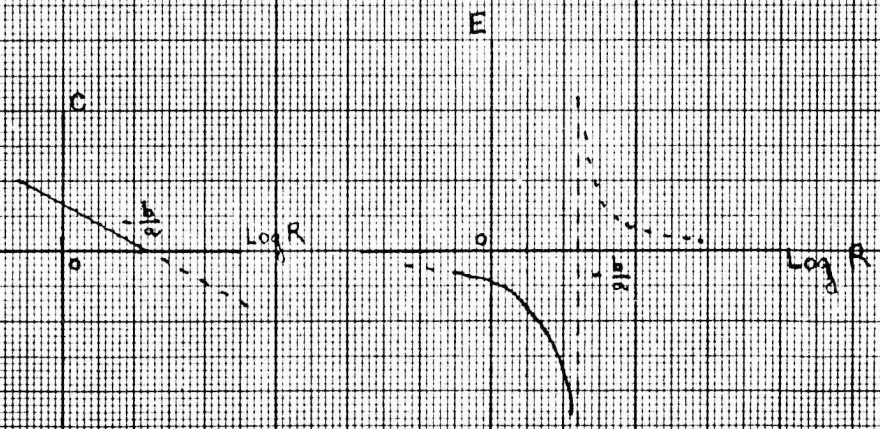


FIGURE III-17

$a < 0$



Cette valeur est d'autant plus faible que le produit est plus pauvre. En fait, les produits les plus pauvres ne disparaissent du budget que pour des valeurs élevées de la dépense totale par unité de consommation : 1 300 000 Frs pour l'autoconsommation, environ 1 500 000 Frs pour les achats de margarine. Le point d'abscisse $-\frac{b}{a}$ est donc pratiquement toujours à l'extérieur du domaine d'observation.

L'élasticité est négative et décroît avec le niveau de vie. Mais elle est toujours faible en valeur absolue et varie donc assez peu. Elle est comprise entre - 0,10 et - 0,13 pour le pain, - 0,13 et - 0,15 pour la margarine, - 0,25 et 0,45 pour le lard. On n'observe de variation notable que pour l'élasticité de l'autoconsommation qui varie de -0,30 à - 1,50.

La figure III-18 montre la variation de l'élasticité pour quelques produits. Elle représente une application concrète des figures III-16 et III-17. Les courbes tracées sur cette figure dont les segments d'une branche d'hyperbole équilatère, limitée aux valeurs utiles de log R.

Si l'on range les différents produits en fonction de l'élasticité de leur consommation au point moyen et si l'on calcule la part de la consommation alimentaire prise par différents groupes de produits d'élasticité voisine, on voit (Tableau III-L) que si les élasticités s'étendent entre - 0,7 et 2,5, les produits dont l'élasticité est comprise entre 0 et 1 représentent plus de la moitié de la consommation alimentaire. Sur la figure III-18, nous avons représenté un exemple de chacun des groupes ainsi définis (le quatrième groupe est représenté par l'ensemble des achats alimentaires). On situera facilement les autres produits sur cette figure par interpolation. Dans le cas du total des achats alimentaires, l'élasticité est inférieure à l'unité dès que le niveau de vie dépasse 150 000 Frs par an par unité de consommation. Peu de familles ont un niveau de vie plus faible : seulement quelque 5 % de l'échantillon, principalement des ménages d'un ou deux adultes de plus de 70 ans. Sauf pour ces ménages à niveau de vie très faible, la première loi d'ENGEL qui dit que l'élasticité des dépenses alimentaires est inférieure à l'unité est donc vérifiée. Pour le reste de l'échantillon, l'élasticité des dépenses est comprise entre 0,76 et 0,40.

FIGURE III - 18

Loi Semi-logarithmique

Variation de E_2^* avec R

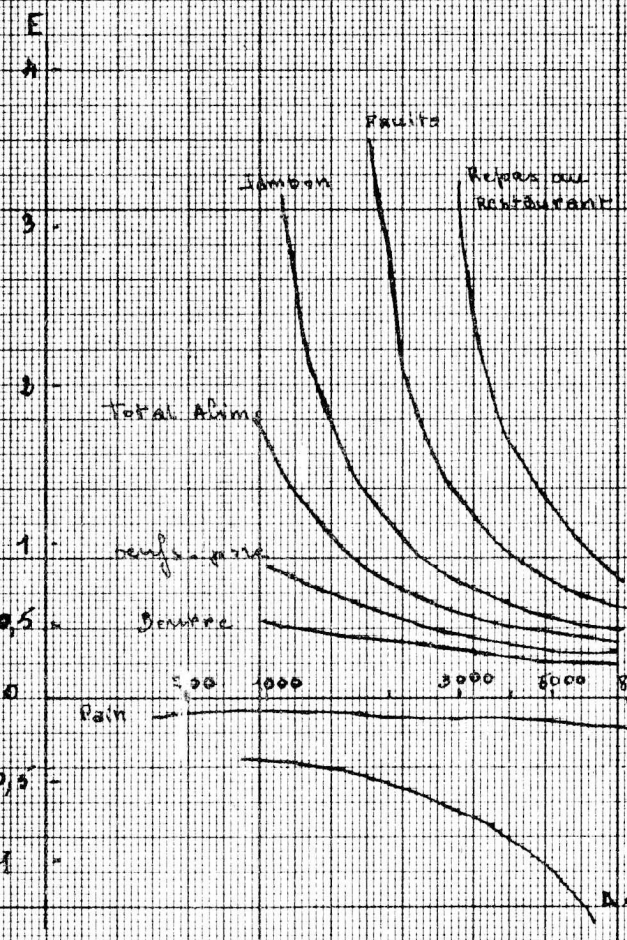


TABLEAU III - L

Produits	% de la dépense alimentaire	Elasticité au point moyen
- Repas au restaurant et à la cantine, consommations prises au café, achats d'apéritifs et de liqueurs	6 %	de 1,8 à 2,5
- Volaille, lapin, gibier, mouton, pâtisserie, fruits, légumes verts eaux minérales, jus de fruits ...	18 %	de 1 à 1,39
- Conserves de fruits et légumes, triperie, jambon-charcuterie, cheval, poissons, vin	17 %	de 0,66 à 0,99
- Fromages, veau, chocolat-confiserie, boeuf	18 %	de 0,51 à 0,59
- Oeufs, porc, pommes de terre, produits d'épicerie, beurre, légumes secs, huile, café-thé-chicorée, farine	21 %	de 0 à 0,43
- Lait, sucre, pain, cidre-bière, margarine, lard-saindoux	12 %	de - 0,34 à 0
- Auto-consommation	8 %	- 0,72

Un modèle semi-logarithmique a aussi été adopté pour les achats de tabac et de chaussures. Le tabac se situe entre le second et le troisième groupe du tableau III-L, l'élasticité au point moyen étant voisine de l'unité. Les achats de chaussures dont l'élasticité est plus faible se situent dans le troisième groupe.

b - Ajustements à élasticité constante

Ces ajustements ont été choisis pour les dépenses non alimentaires, sauf les achats de tabac et de chaussures qui ont été considérés en a) et les achats de véhicules individuels pour lesquels, comme nous l'avons montré dans le chapitre I, la notion de courbe d'ENGEL n'a aucune signification théorique. Les ajustements présentés dans le tableau III-P correspondent à 88 % des dépenses non alimentaires. En dehors du cas limite des achats de véhicules individuels, il existe plusieurs catégories de consommations non-alimentaires pour lesquelles les courbes d'ENGEL ont peu de pouvoir explicatif ; elles n'ont alors d'utilité que pour décrire la structure des consommations dans la population étudiée. C'est le cas des loyers et charges, des dépenses d'équipement ménager, des dépenses de santé, des achats de biens durables, de culture et loisirs et de certaines consommations contractuelles figurant parmi les dépenses diverses.

Même en ne considérant que les dépenses courantes, l'intervalle de variation des coefficients d'élasticité observés est très large, puisqu'il s'étend de 0,4 pour les achats de produits d'entretien à 3,15 pour les dépenses de vacances. Le tableau III-M répartit les consommations non alimentaires en cinq groupes selon la valeur de l'élasticité. Plus de la moitié des dépenses correspond à des produits dont l'élasticité est comprise entre 1 et 1,6.

TABLEAU III - M

Dépenses	% par rapport à l'ensemble étudié	Elasticité
- Produits d'entretien, combustibles, dépenses médicales nettes, accessoires d'habillement	23 %	de 0,4 à 0,9
- Eau-gaz-électricité, tabac, petit équipement ménager, lingerie-bonneterie, loyer	22 %	de 1 à 1,4
- Vêtements principaux, réparation et entretien de l'habillement, dépenses courantes pour les loisirs, meubles et dépenses diverses	31 %	de 1,5 à 1,8
- Hygiène et soins personnels, transports collectifs, services domestiques, gros équipement ménager, achats d'équipement pour les loisirs	14 %	de 2 à 2,5
- Vacances, dépenses pour l'utilisation des véhicules	10 %	de 3,2 à 3,7

Certains des résultats obtenus doivent être interprêtés avec prudence. C'est le cas des consommations médicales, dont d'autres études ont montré qu'elles ne pouvaient être que faiblement expliquées par l'influence du revenu (76). L'élasticité des dépenses de vacances est estimée avec un fort biais quand on considère l'ensemble de la population non agricole, car elles varient fortement avec l'urbanisation à revenu constant (77). Enfin, l'élasticité des dépenses d'utilisation des véhicules ne mesure pas l'influence du revenu sur l'utilisation de leurs véhicules par les possesseurs, puisque l'étude a été faite sans distinguer les ménages possédant ou ne possédant pas de véhicules. Le coefficient indiqué mesure donc à la fois l'augmentation de la circulation chez les possesseurs de véhicules et l'accroissement du nombre relatif de possesseurs quand le niveau de vie augmente (78).

A. — AJUSTEMENTS SEMI-LOGARITHMIQUES

	PRODUITS	LOI AJUSTÉE * $y = a \log x + b$ — y = Dépense annuelle pour l'achat d'un produit ou d'un groupe de produit. x = Dépense totale annuelle. Centaines de francs par unité de consommation	ÉCART-TYPE DE a	DÉPENSE MOYENNE — Centaines de francs par an, par U. C.	ÉLASTICITÉ POUR UNE DÉPENSE TOTALE ANNUELLE PAR U. C. DE : 180 000 F	333 000 F (moyenne de l'ajustement)	600 000 F
A	**DÉPENSES ALIMENTAIRES**						
11	**Produits à base de céréales**	$y =$ 75,1 log x — 90,2	7,8	175	0,21	0,19	0,17
111	Pain	$y =$ — 26,0 log x + 187,1	5,9	98	— 0,11	— 0,12	— 0,13
112	Farine ordinaire, farines composées, riz, pâtes	$y =$ 1,8 log x + 31,2	2,4 *	37	0,02	0,02	0,02
113	Biscottes, biscuits, pâtisserie	$y =$ 99,9 log x — 310,6	5,6	40	2,96	1,05	0,65
12	**Légumes**	$y =$ 234,8 log x — 703,5	8,6	123	1,68	0,83	0,55
121	Pommes de terre	$y =$ 26,9 log x — 65,9	3,4	29	0,54	0,41	0,33
122	Légumes verts et racines	$y =$ 185,3 log x — 574,0	7,0	77	2,75	1,03	0,64
123	Légumes secs	$y =$ 4,0 log x — 7,0	1,0	7	0,29	0,25	0,22
124	Conserves de légumes	$y =$ 18,6 log x — 56,6	1,7	9	2,06	0,91	0,59
13	**Fruits**	$y =$ 232,7 log x — 727,2	6,8	90	3,34	1,10	0,66
131	Agrumes, bananes	$y =$ 87,4 log x — 270,6	3,5	37	2,70	1,02	0,63
132	Autres fruits frais	$y =$ 113,1 log x — 356,1	4,5	41	4,05	1,17	0,69
133	Fruits secs, conserve de fruits, confiture, miel, compotes	$y =$ 28,1 log x — 86,7	2,4	12	2,54	0,99	0,63

Tableau III-N (suite)

14	**Viandes**	$y =$	497,9 log x — 1 412,5	23,2	349	1,03	0,64	0,46
141	Bœuf	$y =$	149,0 log x — 398,9	9,3	129	0,75	0,51	0,39
142	Veau	$y =$	85,9 log x — 236	6,5	67	0,85	0,56	0,42
143	Mouton	$y =$	78,4 log x — 251,6	3,9	24	—	1,39	0,76
144	Cheval	$y =$	18,2 log x — 53,2	3,4	12	1,32	0,73	0,51
145	Porc frais	$y =$	32,0 log x — 78,9	5,2	36	0,55	0,41	0,33
146	Jambon, charcuterie	$y =$	109,5 log x — 318,5	5,6	69	1,25	0,72	0,50
147	Triperie	$y =$	16,7 log x — 50,7	1,5	8	1,92	0,89	0,58
148	Conserves de viande	$y =$	12,0 log x — 38,0	3,7	4	5	1,24	0,71
15	**Volailles, œufs, poissons**	$y =$	237,2 log x — 716,1	11,2	121	1,84	0,87	0,57
151	Volailles, lapins, gibiers	$y =$	127,5 log x — 409,0	7,0	41	9,13	1,39	0,76
152	Œufs	$y =$	28,8 log x — 72,0	1,9	29	0,58	0,43	0,34
153	Poissons, conserves de poisson	$y =$	80,9 log x — 235,0	4,7	51	1,24	0,70	0,49
16	**Lait, fromage**	$y =$	97,3 log x — 214,1	7,1	130	0,41	0,33	0,28
161	Lait frais, lait en poudre	$y = -$	0,3 log x + 58,3	5,4 *	59	— 0,003	— 0,003	— 0,003
162	Fromages, petits suisses	$y =$	97,6 log x — 272,4	4,2	71	0,93	0,59	0,44
17	**Corps gras**	$y =$	73,3 log x — 116,7	8,3	138	0,26	0,23	0,19
171	Beurre	$y =$	61,3 log x — 128,6	4,7	86	0,38	0,31	0,26
172	Margarine, graisses végétales	$y = -$	1,9 log x + 12,2	0,8	6	— 0,13	— 0,14	— 0,15
173	Huile	$y =$	19,3 log x — 26,5	4,6	39	0,23	0,20	0,18
174	Lard, saindoux, autres graisses	$y = -$	5,5 log x + 26,2	1,1	7	— 0,28	— 0,34	— 0,43
18	**Autres produits alimentaires**	$y =$	29,3 log x — 37,9	4,2	65	0,21	0,19	0,17
181	Sucre	$y = -$	6,7 log x + 54,0	2,6	29	— 0,09	— 0,09	— 0,10
182	Chocolat, cacao, confiserie, crèmes glacées	$y =$	23,6 log x — 63,4	3,4	21	0,77	0,52	0,39
183	Autres produits	$y =$	12,5 log x — 28,5	1,6	15	0,45	0,35	0,29

* L'écart entre la loi pour 13, 14, 19 et la somme des lois pour les composantes de ces postes vient de ce que les ajustements ont été faits séparément pour les postes à deux et à trois chiffres.

A. — AJUSTEMENTS SEMI-LOGARITHMIQUES *(suite)*

	PRODUITS	LOI AJUSTÉE $y = a \log x + b$; y = dépense annuelle pour l'achat d'un produit ou d'un groupe de produit. x = Dépense totale annuelle. Centaines de francs par unité de consommation	ÉCART-TYPE DE a	DÉPENSE MOYENNE Centaines de francs par an, par U. C.	ÉLASTICITÉ POUR UNE DÉPENSE TOTALE ANNUELLE PAR U. C. DE : 180 000 F	333 000 F (moyenne de l'ajustement)	600 000 F
19	**Boissons et stimulants**	$y = 276{,}1 \log x - 768{,}1$	15,8	198	0,92	0,59	0,43
191	Café, thé, chicorée, malt	$y = 25{,}0 \log x - 27{,}6$	6,0	56	0,20	0,18	0,16
192	Eaux minérales, jus de fruits	$y = 32{,}2 \log x - 99{,}2$	1,8	14	2,57	1	0,63
193	Vins	$y = 156{,}7 \log x - 448{,}1$	12,1	103	1,09	0,66	0,47
194	Cidre, bière	$y = 3{,}0 \log x + 0{,}8$	2,1 *	12	— 0,12	— 0,13	— 0,11
195	Apéritifs, liqueurs, alcools	$y = 62{,}8 \log x - 206{,}2$	4,2	14	—	1,82	0,87
	TOTAL ACHATS	$y = 1\,749{,}6 \log x - 4\,772{,}4$	64,5	1 388	0,82	0,55	0,41
2	**Autoconsommation et autofournitures**	$y = - 197{,}9 \log x + 815{,}2$	32,5	119	— 0,50	— 0,72	— 1,27
3	**Repas et consommations pris hors du domicile**	$y = 568{,}3 \log x - 1\,898{,}3$	60,8	80	—	2,41	0,97
301	Restaurant, cantine	$y = 500{,}8 \log x - 1\,676{,}3$	55,4	67	—	2,51	1,01
302	Consommations au café	$y = 67{,}5 \log x - 222{,}0$	6,8	13	—	1,88	0,89
4	TOTAL ALIMENTATION	$y = 2\,120{,}1 \log x - 5\,855{,}5$	54,6	1 587	0,88	0,57	0,43
B	DÉPENSES NON ALIMENTAIRES						
514	Chaussures	$y = 135{,}8 \log x - 413{,}7$	6,4	66	2,09	0,92	0,59
571	Tabac, allumettes	$y = 114{,}3 \log x - 353{,}3$	9,2	48	2,66	1,01	0,63

* L'astérisque concerne des ajustements imprécis.

Tableau III-P

B. — AJUSTEMENTS LOGARITHMIQUES

	CONSOMMATION NON ALIMENTAIRE	DÉPENSE MOYENNE PAR U. C. (centaines de francs par an)	COEFFICIENT D'ÉLASTICITÉ a	ÉCART-TYPE DE a
51	**Habillement** (1)	425,9	1,35	0,041
511	Vêtements principaux et de travail	209,3	1,50	0,059
512	Lingerie bonneterie façon	63,0	1,36	0,125
513	Accessoires, laine, tissus de confection, aiguilles, fil, boutons	37,8	0,92	0,061
515	Réparation entretien { blanchisserie . teinturerie ... ressemelages .	49,8	1,62	0,063
52	**Loyers et charges** (2)	214,6	1,46	0,064
521	Loyers (locataires)	117,2	1,35	0,092
524	Services domestiques, résidences secondaires	25,7	2,43	0,37
53	**Équipement du logement**	168,9	1,64	0,069
531	Meubles, literie, linge de maison	59,9	1,70	0,129
532	Équipement ménager	45,5	2,38	0,154
533	Petit équipement réparations	63,5	1,33	0,068
54	**Fourniture et énergie**	284,5	0,60	0,041
541	Eau, gaz, électricité, butane, glace	79,9	1,03	0,037
542	Combustibles solides et liquides	146,6	0,40	0,088
543	Produits d'entretien	58,0	0,43	0,034
55	**Hygiène et soins**	203,3	0,98	0,045
551	Hygiène et soins personnels	54,8	1,95	0,065
552	Dépense de santé (brutes)	214,9	0,68	0,054
554	Dépenses de santé (nettes)	148,5	0,74	0,063
56	**Transports et vacances** (3)	318,8	2,87	0,105
563	Utilisation de véhicules individuels	112,8	3,74	0,219
564	Transports collectifs et communications	45,4	1,97	0,064
565	Vacances	61,8	3,15	0,114
57	**Cultures et loisirs**	216,9	1,65	0,055
571	Tabacs, allumettes	47,7	1,09	0,096
572	Culture et loisirs (biens d'équipement)	62,7	2,50	0,134
573	Culture et loisirs (biens non durables et services)	106,5	1,60	0,074
58	**Dépenses diverses nettes**	112,0	1,83	0,070
5	TOTAL NON ALIMENTAIRE	1944,9	1,37	0,016

(1) Y compris les dépenses de chaussures.

(2) Y compris les dépenses de loyers concernant les ménages propriétaires de leur résidence principale ou logés gratuitement.

(3) Y compris l'achat des véhicules.

SECTION 2

L'EFFET DE L'URBANISATION SUR LA CONSOMMATION

§ 1 - La méthode d'analyse

Une observation intuitive que nous venons d'évoquer en parlant des dépenses de vacances suggère que la sociologie des consommateurs est différente dans les communes rurales et dans les villes. Cela conduit à étudier systématiquement l'effet de l'urbanisation sur les paramètres des courbes d'ENGEL. Cette étude avait été suggérée par le tableau III-B du début de ce chapitre qui montrait les différences entre les budgets des consommateurs habitant des agglomérations de taille différente. Nous n'avions pas pu alors conclure à un effet propre à l'urbanisation, puisque les données de ce tableau n'étaient pas corrigées des différences de revenu.

Pour séparer l'influence du revenu de celle de la taille de l'agglomération, la méthode la plus directe est d'ajuster séparément des courbes d'ENGEL pour chaque type d'agglomération et d'examiner si les paramètres des lois obtenues pour un même produit sont significativement différents.

L'étude a été faite en divisant l'échantillon suivant l'importance de la commune de résidence des familles enquêtées d'une façon que décrit le tableau III-Q. A l'intérieur des cinq catégories de communes retenues, on a mené l'analyse d'une façon voisine de celle de la section précédente, sous réserve de deux différences mineures. D'une part, on a trié l'échantillon selon sept types de ménages au lieu de seize, afin de ne pas obtenir de sous-échantillons trop peu nombreux : d'autre part, on a pris comme variable explicative le total des dépenses de consommation au lieu de la valeur de la consommation totale, afin d'éviter le calcul coûteux de l'autoconsommation des produits des jardins familiaux. Le tableau III-Q montre que la différence entre la consommation totale et la somme des dépenses de consommation, si elle varie avec la catégorie de communes, est toujours faible.

Nous présentons les résultats obtenus dans le tableau III-R. Nous les commenterons dans le paragraphe suivant, en nous limitant aux consommations alimentaires. Le tableau montre en effet que les différences observées pour les consommations non alimentaires ne sont pas en général significatives (79).

TABLEAU III - Q

Valeur moyenne, 1000 F par an par UC

Catégorie de communes	Nombre de ménages	Dépenses de consommation	Auto-consommation	Consommation totale
1. Communes rurales	5 393	262	25	287
2. Province, moins de 10.000 h.	2 162	303	18	321
3. Province, 10.000 à moins de 100.000 h.	3 350	342,5	8,5	351
4. Province, 100.000 h. et plus	2 849	367,5	5,5	373
5. Agglomération parisienne	3 678	420	2,0	422
Σ ENSEMBLE	17 432	341	12	353

TABLEAU III - R

I - CONSOMMATIONS ALIMENTAIRES

Loi ajustée : $C = a \log R + b$

Unités : a, b, s_a, $\bar{C}$ en centaines de francs par an par unité de consommation

a_m = moyenne pondérée des a_k

E_m = mpyenne pondérée des E_k

E_0, a_0 = ajustement sur l'ensemble des observations, sans stratification.

signes o et * : voir pages 239

C.C.		a	b	s_a	$\bar{C}$	Valeur de E pour R = 3140
I.- Achats alimentaires (total)						
1	o*	1 809,6	- 4 920,8	71,6	1 206,5	0,56
2		1 865,1	- 5 384,7	99,6	1 426,3	0,57
3		1 966,9	- 5 410,1	86,1	1 491,7	0,58
4		2 092,2	- 5 821,7	94,7	1 586,9	0,61
5	o	2 504,9	- 7 256,0	120,6	1 851,9	0,72
Ens.	a_m	2 016				0,60 = E_m
	a_0	2 197,5		43,4	1 486,9	0,64 = E_0
II.- Produits à base de céréales						
1		88,3	- 120,8	9,6	178,3	0,20
2		89,5	- 128,2	16,6	182,1	0,21
3	*	104,5	- 192,1	14,5	174,5	0,26
4	*	122,1	- 253,1	15,4	179,2	0,30
5	*	108,7	- 237,9	13,5	157,2	0,33
Ens.	a_m	95				0,25 = E_m
	a_0	59,6		6,0	173,7	0,15 = E_0

TABLEAU III - R (suite 1)

C.C.	a	b	s_a	$\bar{c}$	Valeur de E pour R = 3140
12.- Légumes					
1 *	172,0	- 512,6	12,3	69,6	0,84
2	175,9	- 500,0	23,0	109,8	0,66
3 *	176,7	- 489,1	14,9	130,9	0,60
4 *	169,0	- 441,6	19,9	157,0	0,49
5 *	161,9	- 413,9	19,8	174,7	0,46
Ens. a_m	170,0				0,61 = E_m
Ens. a_0	235,2		7,3	122,8	0,83 = E_0
13.- Fruits					
1 o*	138,1	- 417,3	7,8	50,2	0,92
2 *	159,1	- 477,0	13,1	74,3	0,87
3 *	186,4	- 561,1	9,1	93,0	0,89
4 o	231,6	- 707,2	17,2	113,0	0,98
5 *	194,9	- 564,0	12,2	144,5	0,72
Ens. a_m	171,0				0,85 = E_m
Ens. a_0	230,2		5,4	91,6	1,09 = E_0
14.- Viandes					
1	466,8	- 1 298,8	34,6	281,9	0,61
2	422,7	- 1 142,8	45,1	322,4	0,55
3	392,7	- 1 040,1	41,1	338,1	0,51
4	470,0	- 1 313,4	41,9	350,7	0,62
5	457,5	1 236,8	53,3	426,6	0,55
Ens. a_m	432,0				0,57 = E_m
Ens. a_0			16,0	339,5	0,61 = E_0

TABLEAU III - R (suite 2)

C.C.	a	b	s_a	$\bar{C}$	Valeur de E pour R = 3140
15.- Volailles, Poissons, Oeufs					
1 *	177,7	- 519,1	12,8	82,6	0,75
2	183,3	- 523,8	24,1	111,5	0,68
3	230,5	- 678,3	21,5	130,7	0,78
4	195,1	- 553,6	22,7	137,2	0,66
5	234,1	- 699,0	23,7	152,1	0,85
Ens. a_m	199,0				0,74 = E_m
Ens. a_0	226,1		8,4	119,0	0,83 = E_0
16.- Lait, Fromages					
1	108,1	- 255,9	10,1	110,0	0,38
2	96,4	- 205,7	14,6	128,5	0,32
3 *	63,8	- 93,2	12,3	130,5	0,21
4	93,4	- 192,0	12,7	138,7	0,30
5	72,3	- 118,9	14,2	144,0	0,23
Ens. a_m	87,0				0,30 = E_m
Ens. a_0	99,2		4,8	128,1	0,34 = E_0
17.- Corps gras					
1	89,8	- 166,3	18,1	137,7	0,26
2	83,4	- 139,8	23,0	149,4	0,24
3	73,3	- 110,8	15,9	146,6	0,22
4	93,8	- 193,7	16,0	138,5	0,30
5	87,6	- 180,5	13,2	137,8	0,30
Ens. a_m	82,0				0,26 = E_m
Ens. a_0	64,7		6,6	141,0	0,20 = E_m

TABLEAU III - R (suite 3)

C.C.	a	b	s_a	$\bar{C}$	Valeur de E pour R = 3140
18.- Produits divers					
1	49,5	- 99,0	20,6	68,6	0,29
2	39,9	- 67,1	9,9	71,3	0,24
3	31,92	- 48,6	8,6	63,4	0,22
4 *	48,3	- 105,0	8,1	66,2	0,33
5 *	50,6	- 125,9	6,9	58,0	0,43
Ens. a_m	44,2				0,30 = E_m
Ens. a_0	23,9		3,5	65,3	0,16 = E_0
19.- Boissons					
1	240,8	- 639,4	26,2	176,0	0,52
2	267,7	- 718,7	39,8	209,3	0,53
3	244,3	- 656,0	28,6	201,2	0,54
4	204,8	- 522,5	32,0	202,9	0,46
5	299,8	- 845,9	30,2	244,0	0,64
Ens. a_m	248,0				0,54 = E_m
Ens. a_0	252,4		12,8	203,7	0,54 = E_0

TABLEAU III - R (suite 4)

II - CONSOMMATIONS AUTRES QUE LES ACHATS DE PRODUITS ALIMENTAIRES

Loi ajustée : log. C = a log R + b

Unité : $\bar{C}$ en centaines de francs par an par unité de consommation

	Elasticité E	s_E	$\bar{C}$
51.- Habillement			
1	1,19	0,10	320
2	1,34	0,08	380
3	1,28	0,07	456
4	1,39	0,07	467
5	1,32	0,06	587
Ens. E_m	1,29		
Ens. E_0	1,25	0,03	431
52.- Loyers et charges			
1	1,19	0,19	135
2	1,46	0,16	156
3	1,39	0,13	231
4	1,23	0,17	239
5	1,33	0,10	344
Ens. E_m	1,30		
Ens. E_0	1,35	0,06	215
53.- Equipement du logement			
1	1,62	0,10	120
2	1,61	0,19	161
3	1,67	0,15	183
4	1,59	0,12	193
5	1,57	0,13	219
Ens. E_m	1,61		
Ens. E_0	1,48	0,05	169

TABLEAU III - R (suite 5)

	Elasticité E	s_E	$\bar{C}$
54.- Fournitures, énergie			
1	0,61	0,08	233
2	0,71	0,11	275
3	0,62	0,07	291
4	0,43	0,07	293
5	0,45	0,06	63
Ens. E_m	0,56		
Ens. E_0	0,58	0,03	286
55.- Hygiène et soins			
1	0,75	0,10	155
2	1,05	0,17	164
3	0,79	0,08	184
4	1,18	0,11	217
5	0,90	0,07	317
Ens. E_m	0,90		
Ens. E_0	0,96	0,04	204
56.- Transports, vacances			
1 *	3,31	0,22	224
2	3,10	0,22	255
3	2,69	0,19	276
4	2,61	0,15	351
5	2,24	0,10	530
Ens. E_m	2,80		
Ens. E_0	2,57	0,08	320

TABLEAU III - R (suite 6)

	Elasticité E	s_E	$\bar{C}$
57.- Culture, loisirs			
1	1,45	0,12	142
2	1,63	0,11	185
3	1,62	0,09	226
4	1,77	0,11	240
5	1,51	0,08	334
Ens. E_m	1,54		
Ens. E_0	1,57	0,04	218
58.- Divers			
1	1,25	0,22	68
2	1,09	0,37	94
3	1,62	0,20	118
4	1,64	0,13	117
5	1,71	0,19	161
Ens. E_m	1,47		
Ens. E_0	1,49	0,08	108
51.- Total des dépenses non alimentaires			
1	1,31	0,02	1 397
2	1,39	0,05	1 670
3	1,34	0,04	1 965
4	1,32	0,02	2 118
5	1,28	0,02	2 856
Ens. E_m	1,32		
Ens. E_0	1,30	0,01	5 614

TABLEAU III - R (suite 7)

	Elasticité E	s_E	$\bar{C}$
31.- Repas pris au restaurant et à la cantine			
1	2,77	0,39	40
2	3,64	0,40	50
3	2,05	0,37	70
4	2,89	0,41	87
5	2,25	0,28	192
Ens. E_m	2,53		
Ens. E_0	2,68	0,15	86
32.- Consommations prises au café			
1	1,94	0,19	12
2 *	2,66	0,26	18
3	2,01	0,31	13
4	1,97	0,24	17
5	2,19	0,26	21
Ens. E_m	2,11		
Ens. E_0	1,74	0,10	16
3.- Consommations et repas pris à l'extérieur			
1	2,45	0,29	52
2	3,63	0,32	68
3	2,32	0,36	83
4	2,77	0,32	103
5	2,26	0,28	213
Ens. E_m	2,56		
Ens. E_0	2,52	0,13	102

§ 2 - L'étude des consommations alimentaires

Comme dans la section précédente, nous avons ajusté une loi semi-logarithmique dans le cas des consommations alimentaires. Le tableau III-R donne, pour chaque catégorie de communes et chaque produit, les paramètres a_k et b_k des lois ajustées et l'erreur d'échantillonnage s_{ak} sur a_k. Pour faciliter la comparaison rapide des lois en termes plus habituels que les paramètres d'une relation semi-logarithmique, nous avons calculé l'élasticité des consommations dans chaque sous-population pour un même niveau de vie correspondant à une dépense totale de consommation de 314 000 Frs par an par unité de consommation (80).

Afin d'étudier les différences entre catégories de communes, nous avons estimé de deux façons le coefficient de log x pour l'ensemble de la population. D'une part, nous l'avons calculé sur l'ensemble des observations sans distinguer les catégories de communes. On obtient ainsi les coefficients a_o (81). D'autre part, nous avons calculé pour chaque produit un coefficient a_m qui est la moyenne pondérée des a_k, les poids étant proportionnels au nombre de ménages dans chaque catégorie. Dans le tableau III-R, nous avons marqué au signe * les a_k qui diffèrent de a_m de plus de :

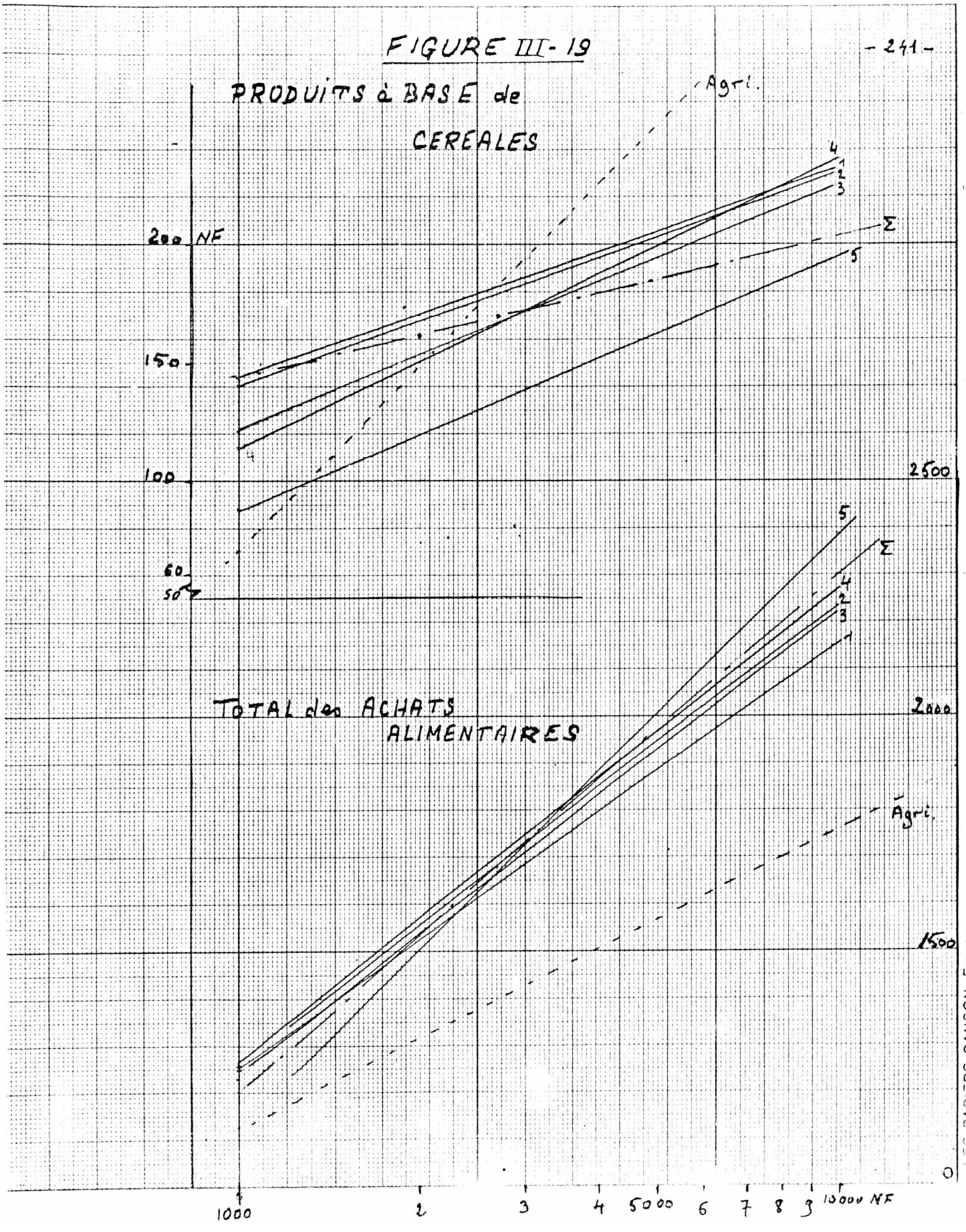
$$2 (s_{a_k} + s_{a_o})$$

c'est-à-dire qu'on pourrait considérer comme significativement différents de s_o si les distributions d'échantillonnage des paramètres étaient normales. Nous avons, de même, repéré par le signe ° les a_k qui diffèrent de a_m de plus de $2\, s_{a_k}$. L'examen des signes * et ° permet de distinguer trois grandes catégories parmi les consommations alimentaires.

a - Lois homogènes pour l'ensemble de la population

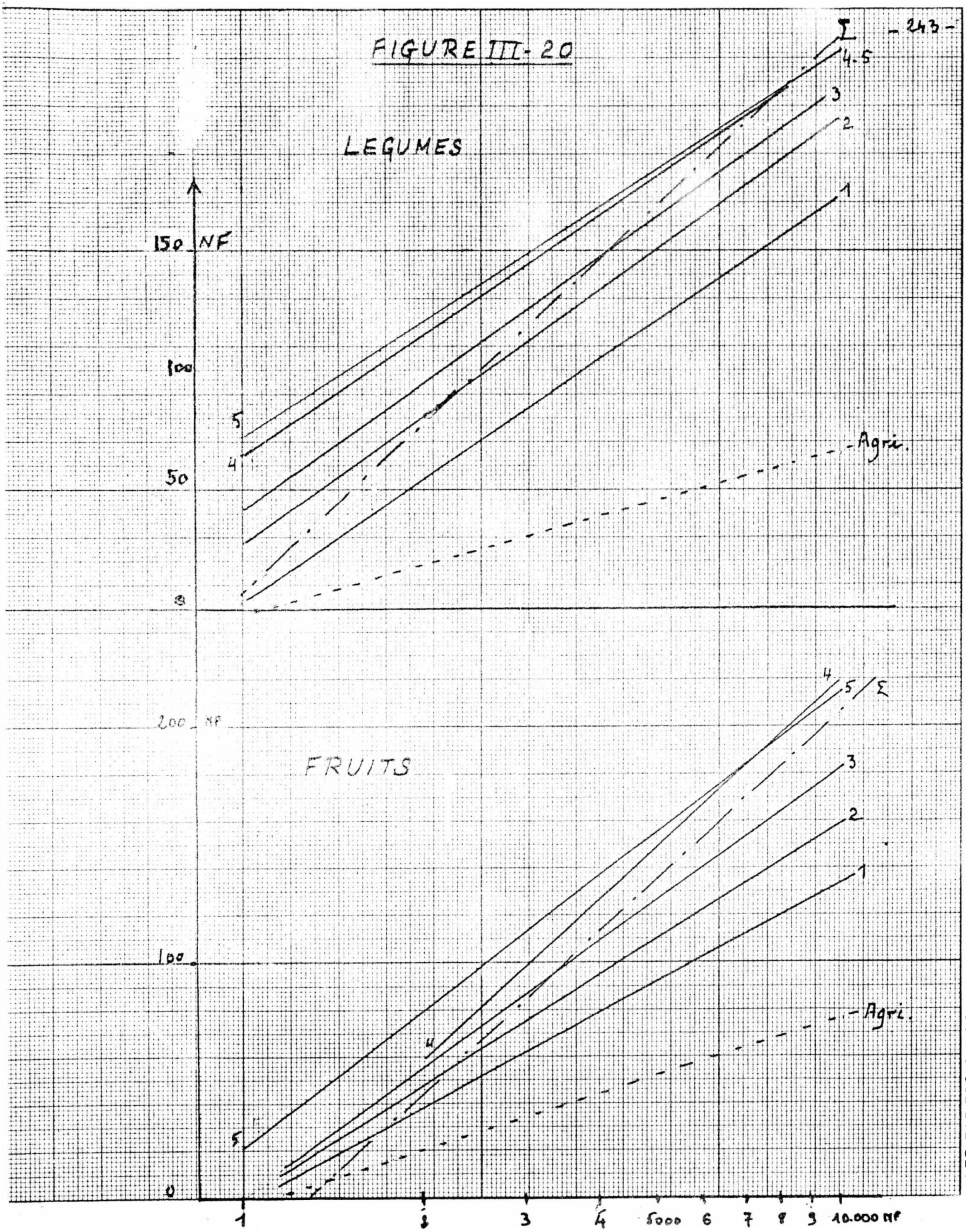
La catégorie de communes n'a pas d'effet significatif sur les paramètres des courbes d'ENGEL dans le cas des consommations de viande de boucherie, de charcuterie et triperie, de corps gras, et de boissons. Au point moyen, l'ensemble de ces consommations représente environ la moitié de la dépense alimentaire.

FIGURE III-19



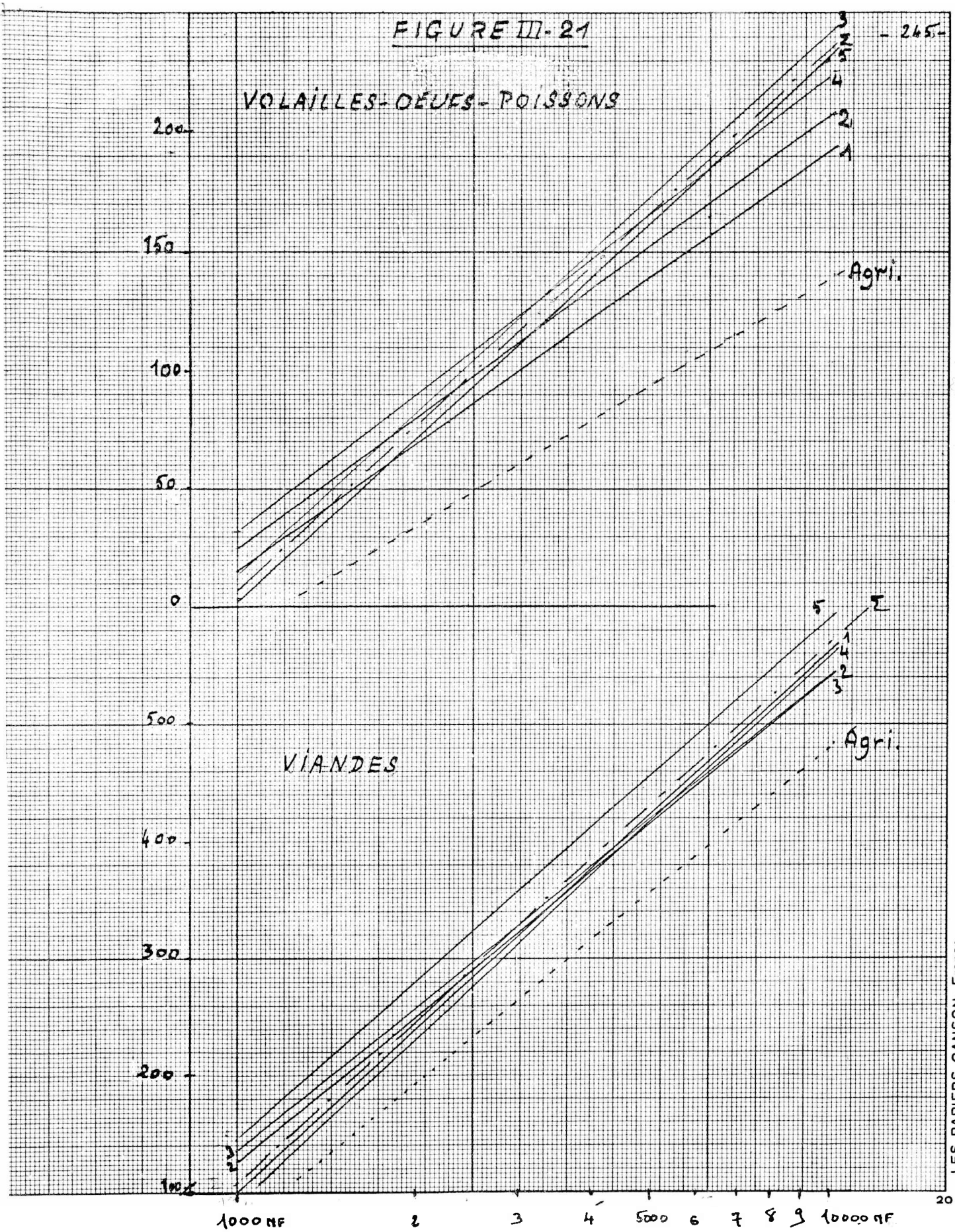
LES PAPIERS CANSON France

FIGURE III - 20



LES BARIBES CANSON FRANCE

FIGURE III-21



LES PAPIERS CANSON France

FIGURE III-22

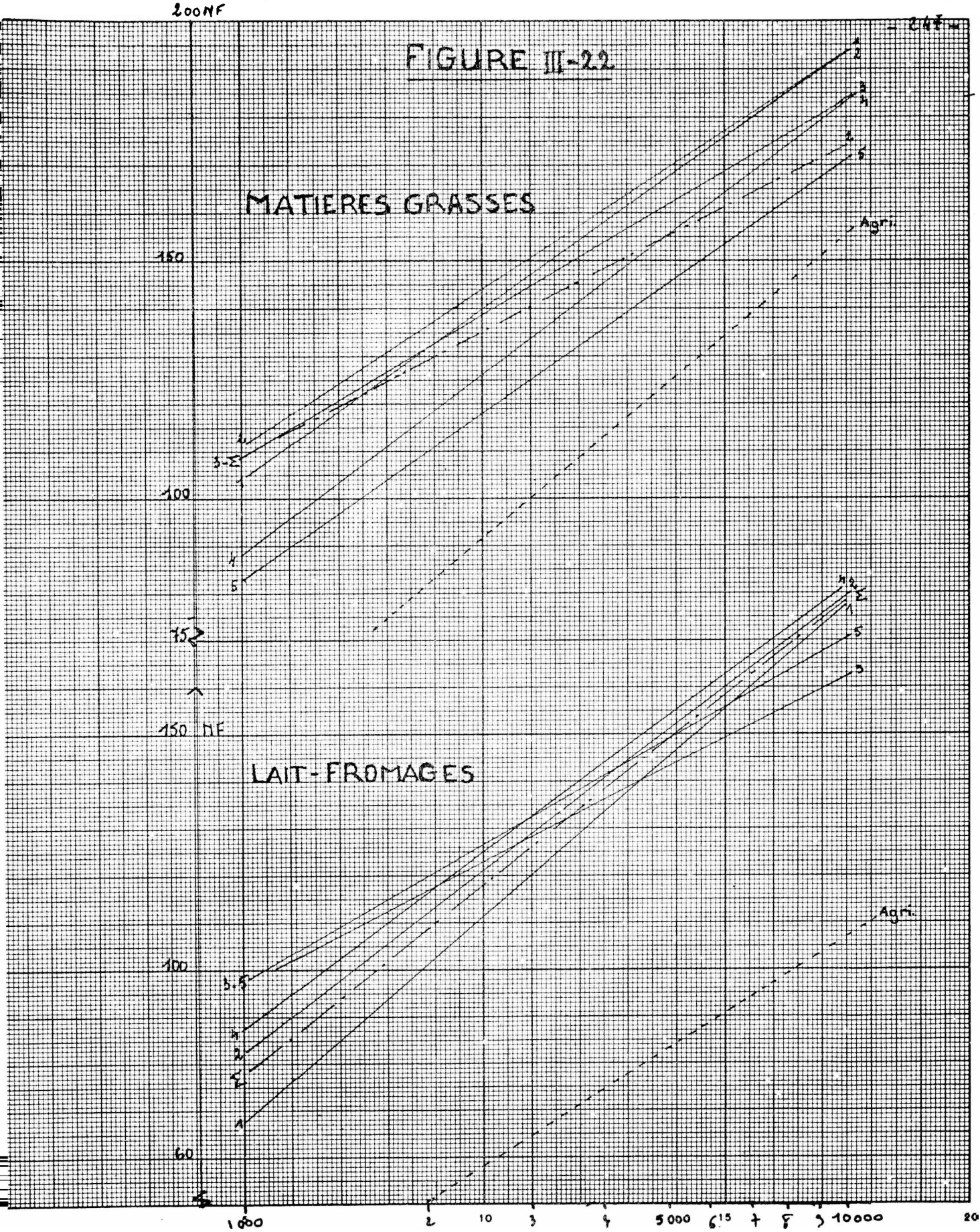
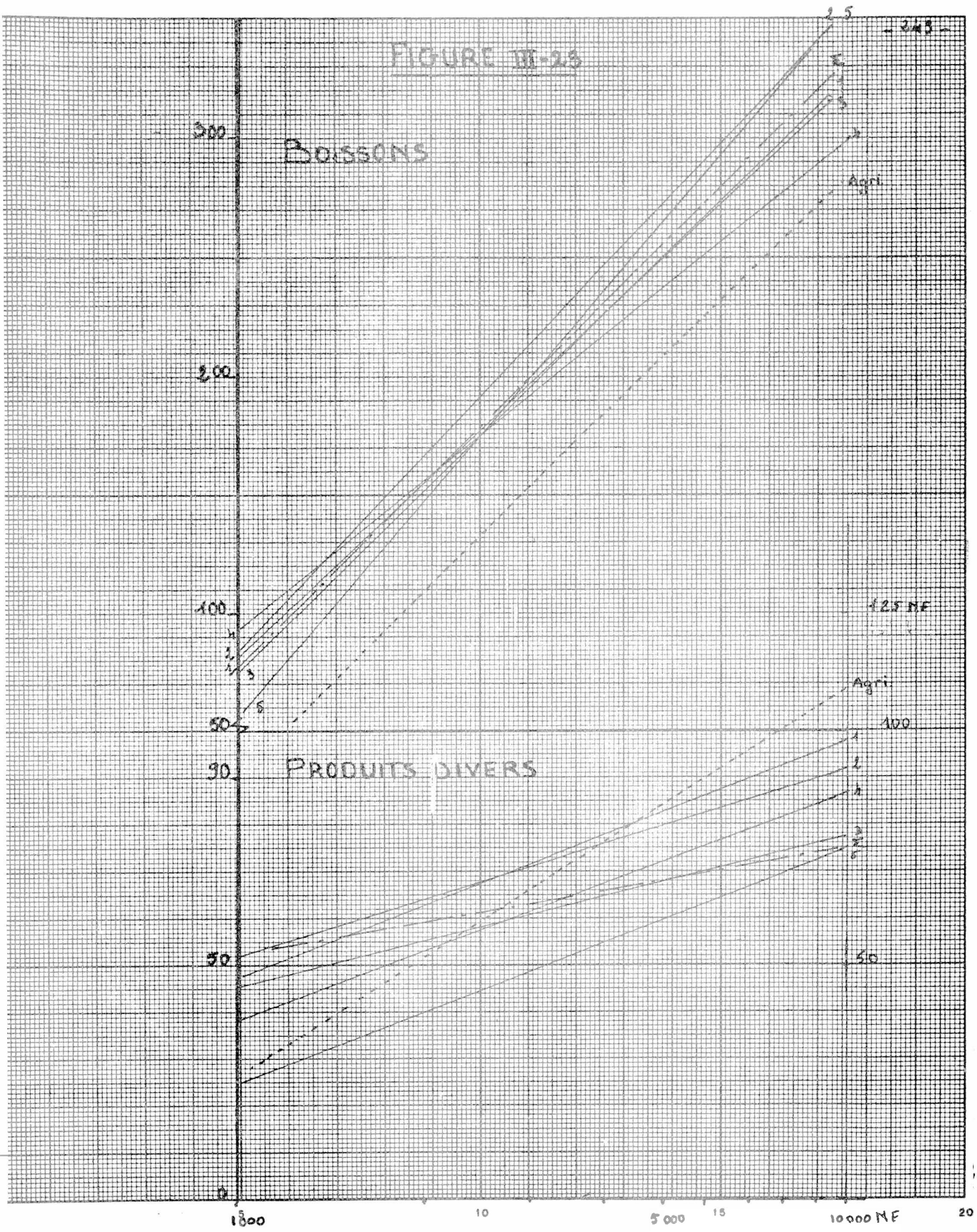


FIGURE III-23



Dans le cas de ces produits, la loi estimée dans la section précédente sur les données relatives à l'ensemble de la population non agricole est correcte.

Les lois relatives aux produits de ce premier groupe sont représentées sur les figures III-21, III-22 et III-23. Nous avons également porté sur ces figures la loi correspondant aux achats des mêmes produits par la population agricole (82). Sans calcul, on voit sur ces figures que le comportement de la population agricole est très différent de celui de la population non agricole. Cela traduit en partie l'incidence de l'autoconsommation de produits provenant de l'exploitation sur les habitudes d'achat de produits alimentaires par les exploitants agricoles. Cette explication n'est cependant pas suffisante.

b - Pentes homogènes pour l'ensemble de la population

Une situation intéressante s'observe pour cinq groupes de produits dont la consommation au point moyen représente près de 45 % de la dépense alimentaire des ménages étudiés. Il s'agit des produits à base de céréales, des légumes, des volailles, oeufs et poissons, des produits laitiers (lait et fromages) et des "produits divers" (sucre, produits sucrés et produits d'épicerie).

Pour ces produits, certains des a_k sont significativement différents de a_o ; par contre, les a_k ne sont pas significativement différents les uns des autres, ni de a_m. On se trouve dans une situation que nous avons déjà rencontrée en étudiant l'influence des structures familiales sur la consommation, dans la première partie de ce chapitre (section 4). Les courbes d'ENGEL correspondant aux différentes catégories de communes sont parallèles, mais leurs ordonnées à l'origine ne sont pas les mêmes. A revenu donné, la valeur de la consommation dépend de la catégorie de communes, mais la propension marginale à consommer le produit étudié n'en dépend pas. Comme les revenus moyens varient avec la catégorie de communes, on obtient une estimation biaisée de la valeur commune des a_k en ajustant une loi sur l'ensemble

des observations sans distinguer la catégorie de communes. Cela revient en effet à introduire l'effet de la catégorie de communes dans le terme aléatoire du modèle, alors qu'elle est en corrélation avec la variable explicative. L'influence du revenu, à urbanisation constante, peut être définie sans ambiguïté pour l'ensemble de la population ; elle ne peut être estimée correctement qu'en triant d'abord la population selon les catégories de communes.

L'effet de la catégorie de communes sur l'ordonnée à l'origine des lois peut se présenter de deux façons différentes selon que la consommation, pour une même valeur du revenu, augmente ou diminue avec la taille des agglomérations. Les explications possibles des différences observées dépendent aussi des produits. Le groupe que nous avons défini doit donc être étudié avec un peu plus de détails. Dans le cas des produits à base de céréales et des produits divers, dont la composante principale est le sucre, l'ordonnée à l'origine des courbes d'ENGEL et donc la consommation à revenu constant, diminuent quand on passe des communes rurales aux grandes villes. Comme il s'agit de produits qui sont toujours achetés et qui ne peuvent pas être autoconsommés, les différences dans les lois ajustées traduisent des différences effectives dans les niveaux de consommation. La diminution de la consommation à revenu constant quand la taille des communes augmente entraîne un biais vers le bas dans l'estimation a_o de a sur l'ensemble des observations.

Comme ces deux groupes de produits sont caractérisés par une forte élasticité qualité, l'augmentation de la consommation avec le revenu à l'intérieur de chaque catégorie de communes, traduit une modification dans la nature des produits achetés sans augmentation notable des consommations quantitatives de sucre ou de céréales. Par contre, la variation de la consommation d'une catégorie de communes à l'autre pour un même revenu recouvre à la fois une modification qualitative et une modification quantitative, sans que les données permettent de distinguer les deux effets.

D'une part, le type d'activité des habitants des communes rurales rend plus élevés leurs besoins alimentaires moyens exprimés en termes de calories : leur consommation de pain et de sucre est donc normalement plus forte que celle des consommateurs urbains. D'autre part, des raisons sociologiques font que la consommation de certains produits comme le pain de fantaisie, les biscottes et biscuits et la confiserie augmentent avec l'urbanisation en même temps qu'avec le revenu (83).

Dans le cas des légumes et des volailles, oeufs et poissons, la valeur des achats augmente par contre avec la taille des agglomérations à revenu constant. L'estimation a_o de a à partir de l'ensemble des observations est donc biaisée vers le haut.

L'interprétation des différences de consommation doit, pour ces produits, tenir compte de l'importance de la production des jardins et basses-cours familiales. Nous l'avons estimée dans le cas des légumes (84). Partant de l'observation que l'autoconsommation de légumes par les ménages non agricoles habitant une même catégorie de communes n'est pas liée significativement au revenu, nous avons établi la figure III-24 qui se déduit de la figure III-20 en décalant chaque courbe d'une différence d'ordonnées égale à la valeur moyenne de l'autoconsommation de légumes dans la catégorie de communes à laquelle elle se rapporte. Les lois obtenues sont voisines les unes des autres et on peut conclure que les différences dans l'autoconsommation suffisent à expliquer celles qu'on observe entre les habitudes d'achats. Bien que nous n'ayons pas pu disposer de données précises, des indications antérieurement publiées montrent que le même résultat est vraisemblable dans le cas des produits de basse-cour (85).

Comme pour le premier groupe de produits, les graphiques montrent que les lois relatives à la population agricole sont, par contre, très différentes de celles valables pour les différentes populations non agricoles.

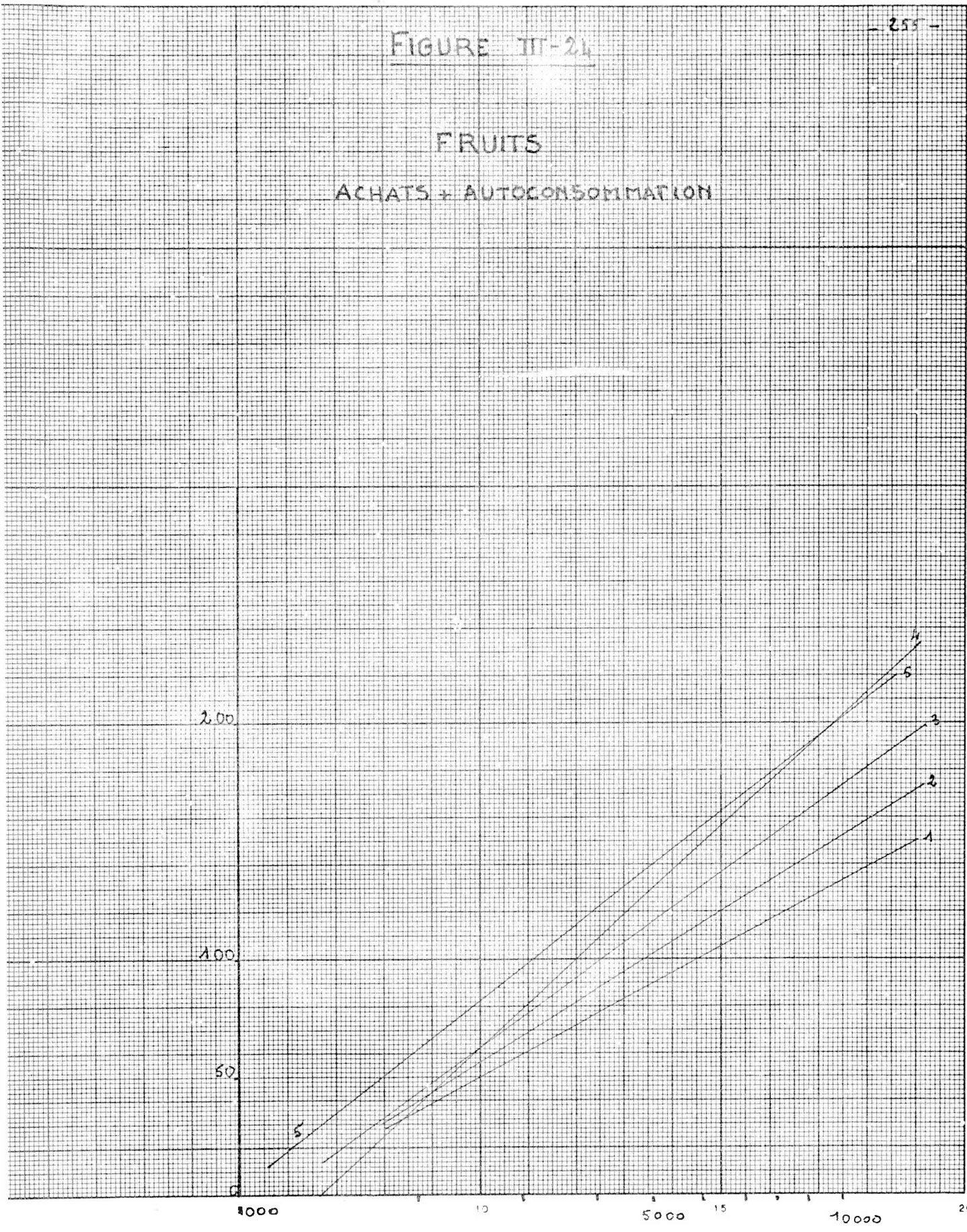
c - Lois entièrement hétérogènes

Les fruits représentent une consommation très particulière : c'est le seul groupe de produits pour lequel à la fois les a_k et les b_k varient significativement d'une catégorie à l'autre. Il n'y a pas de sens, en toute rigueur, à définir dans ce cas un coefficient a_m mesurant l'influence moyenne du revenu sur la consommation pour l'ensemble des catégories de communes.

FIGURE III-24

FRUITS

ACHATS + AUTOCONSOMMATION



IFG PARIS CANSON FRANCE

Une partie des différences dans les habitudes d'achats correspond, comme pour les légumes, à l'autoconsommation. Cependant, la valeur des fruits produits dans les jardins familiaux est beaucoup plus faible que celle des légumes (86) et la figure III-25, établie de la même façon que la figure III-24, montre qu'il subsiste d'importantes différences entre les lois obtenues en considérant l'autoconsommation.

Alors que les consommations de produits à base de céréales comme le pain, et de sucre diminuent avec l'urbanisation, la consommation de fruits augmente avec elle.

§ 3 - L'influence combinée de l'urbanisation et du revenu

Les résultats du paragraphe précédent sont susceptibles de deux interprétations selon le problème qu'on étudie.

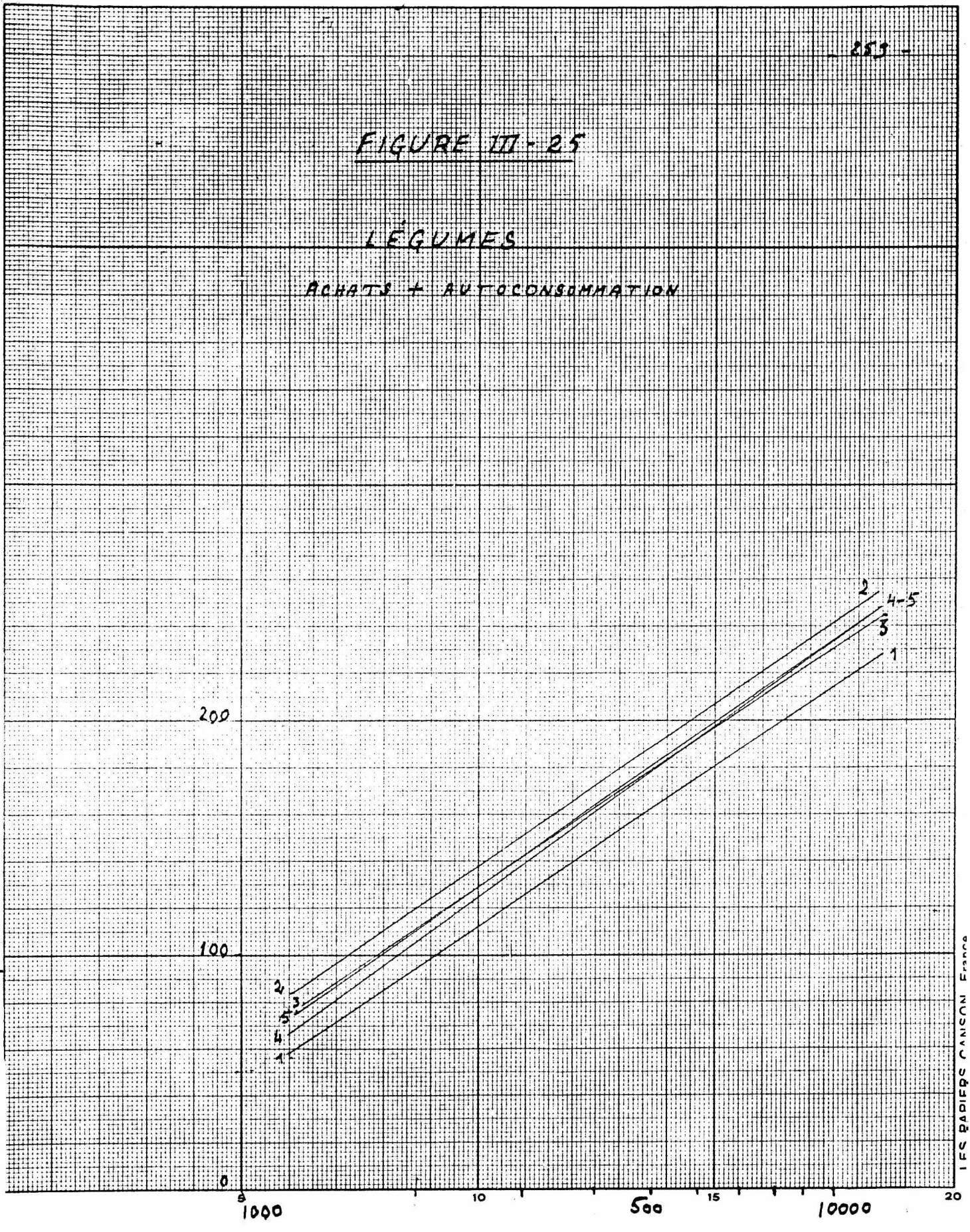
On peut d'abord étudier l'effet net du revenu sur la consommation en admettant que la répartition de la population entre les catégories de communes ne se modifie pas. Pour les produits des deux premiers groupes que nous avons distingués, l'effet du revenu peut être défini rigoureusement pour l'ensemble de la population, mais dans le groupe b) il ne peut être estimé correctement que si l'on analyse d'abord la population selon les catégories de communes. Dans le cas des fruits, on ne peut par contre pas définir en toute rigueur une mesure de l'influence nette du revenu sur la consommation pour l'ensemble de la population, puisque les a_k sont différents.

Mais on peut considérer un autre problème qui se présente quand on établit une perspective de consommation à long terme. Sur une période de plusieurs années, la tendance à l'urbanisation modifie la répartition de la population entre les catégories de communes en même temps que les revenus augmentent. On rencontre alors une même situation pour les produits des deux derniers groupes, puisque pour chacun, une tendance à l'urbanisation modifie le niveau de la consommation même si le revenu ne varie pas. Il ne s'agit plus d'estimer correctement l'influence nette du revenu, mais l'influence combinée des variations des revenus et de l'urbanisation.

FIGURE III - 25

LÉGUMES

ACHATS + AUTOCONSOMMATION



LES PAPIERS CANSON France

Le problème peut être étudié simplement moyennant quelques hypothèses sur la distribution des revenus. Nous admettrons que nous nous trouvons dans l'un des deux cas où l'agrégation stricte des lois semi-logarithmiques est possible :

- ou bien, tous les revenus augmentent du même pourcentage dans chaque sous-population,
- ou bien, la moyenne géométrique des revenus augmente du même pourcentage dans chaque sous-population.

Nous admettrons en outre que l'augmentation prévue du revenu moyen (87) est du même pourcentage dans chaque catégorie de communes (88).

Nous utiliserons les notations suivantes :

C_i Consommation moyenne par unité de consommation dans la strate i à l'origine de la projection

n_i Nombre d'unités de consommation dans la strate i à l'origine de la projection

C et n Mêmes variables pour l'ensemble de la population

C'_i , C' , n'_i , n' mêmes variables au terme de la projection

a_i et b_i paramètres des courbes d'ENGEL dans la strate i

Si k est l'augmentation logarithmique des revenus moyens, que nous avons supposée être la même dans chaque strate, l'augmentation unitaire de la consommation dans la strate i s'écrit :

$$C'_i - C_i = k\, a_i$$

La consommation totale du produit considéré devient :

$$C'_i = n'_i \ (C_i + k\, a_i)$$

pour la strate i, et

$$C' = \sum_i n'_i \ C_i + k \sum_i n'_i \ a_i$$

pour l'ensemble de la population.

La consommation unitaire pour l'ensemble de la population s'écrit :

$$C' - C = \left[\frac{\sum C_i \, n'_i}{\sum n'_i} - \frac{\sum C_i \, n_i}{\sum n_i} \right] + k \ \frac{\sum n'_i \ a_i}{\sum n'_i}$$

L'augmentation de la consommation unitaire est donc la somme de deux termes indépendants. Le premier mesure l'influence des modifications de la structure de la population à revenu constant ; le second l'effet des variations des revenus dans une population ayant la structure constante de l'année terminale.

Le plus souvent, comme les modifications de la structure de la population sont assez lentes, $\frac{\sum n'_i \ a_i}{\sum n'_i}$ sera peu différent de $\frac{\sum n_i \ a_i}{\sum n_i}$ qui est le coefficient a_m que nous avons défini dans le paragraphe précédent.

§ 4 - L'effet d'autres facteurs qualitatifs

L'urbanisation ne représente qu'un des facteurs qualitatifs susceptibles d'introduire des hétérogénéïtés dans les lois d'ENGEL. On pourrait étudier de la même façon, par exemple, l'influence de la région ou de la catégorie sociale du chef de ménage.

L'exploitation mécanographique de l'enquête n'a pas permis d'examiner l'influence de la région ; nous avons par contre examiné celle de la catégorie sociale. Nous ne présenterons pas les résultats que nous avons obtenus, car nous avons abouti à un échec.

D'une part, les ajustements sont médiocres et les différences observées sont rarement significatives (89). D'autre part, les quelques différences significatives n'ont pas pu être interprêtées simplement.

L'échec a deux raisons probables. Les effectifs des sous-populations étudiées varient d'abord de près de un à dix, alors qu'ils ne variaient guère que du simple au double dans le cas des catégories de communes. C'est un exemple de l'inefficience d'un échantillon tiré avec une fraction de sondage constante : on a beaucoup trop de ménages d'ouvriers ou d'inactifs et pas assez de ménages de cadres supérieurs ou de travailleurs indépendants pour qu'on puisse les étudier séparément.

Ensuite, les revenus moyens varient considérablement d'une catégorie à l'autre ; les différences observées entre les lois de comportement peuvent donc traduire une erreur dans la forme des lois ajustées autant qu'un effet propre à la catégorie sociale.

Enfin, les études antérieures suggèrent que l'effet de la catégorie sociale devrait surtout être pris en considération pour les consommations non alimentaires. Or il s'agit de produits pour lesquels les lois d'ENGEL, lorsqu'elles conservent un sens, sont beaucoup moins bien définies que pour les consommations alimentaires. L'examen d'une question importante et apparemment simple dépasse donc les limites qu'imposent à la fois les techniques actuelles d'enquête et les méthodes habituelles d'analyse de leurs résultats. Le développement de l'étude des courbes d'ENGEL depuis quinze ans pose finalement autant de problèmes nouveaux qu'il en résout d'anciens.

NOTES DU CHAPITRE III

(1) - Plusieurs études descriptives, portant sur l'ensemble des budgets de diverses sous-populations ou la consommation de produits particuliers ont été publiées à partir de l'enquête française de 1956 ; notamment :
G. ROTTIER et E. SALEMBIEN : Les budgets familiaux en 1956 ; "Consommation" N° 1 - 1958.
H. FAURE et Cl. SEIBEL : L'étude régionale de la consommation "Consommation", N° 1 - 1959.
G. ROTTIER : La distribution des revenus non-agricoles, Consommation, N° 1 - 1958.
M. GUILLOT : Le crédit à la consommation dans les budgets familiaux, Consommation, N° 4 - 1958.
H. FAURE et E.A. LISLE : Les dépenses d'automobile des ménages français.

(2) - E. ENGEL : Die Productions - und Consumptions-verhältnisse des Königreichs sachsen 1857.
Ernest ENGEL : Die Lebenkosten Belgischer arbeiter familien früher und jetzt, 1895.
Reproduits l'un et l'autre dans le bulletin de l'Institut International de Statistique, Vol. IX, appendice.

(3) - Ce qu'on entend par niveau de vie sera examiné plus bas (section 3), par. 1).

(4) - Un intéressant article de STIGLER a montré l'antériorité de l'étude empirique sur la réflexion théorique dans notre domaine : G.J. STIGLER : The early history of empirical studies of consumer behavior ; Journal of political economy, 1954.

(5) - ALLEN, R.G.D. et BOWLEY, A.L. : Family Expenditure, Londres, 1935, p. 1.

(6) - Des exposés élémentaires et assez complets se trouvent dans E. MALINVAUD : Méthodes statistiques de l'économétrie, Vol 1 - Chap. I - G. ROTTIER : Initiation à l'Econométrie, Chap. IV.

(7) - Cf. M.G. KENDALL : Regression, structure and the Functional Relationship Biometrica, 1951, p. II.

(8) - Cf. W.C. HOOD et T.J. KOOPMANS : Studies in Econometric Method, 1953, spé. chap. II, V et VI.

(9) - Cf. PRAIS et HOUTHAKKER, The analysis of Family budgets, Chap. V, paragraphe 3, pp. 55 seq.

(10) - Si l'on suppose que, avec un nombre donné d'observations, tous les x_{hi} sont égaux à $x_{h.}$, le terme correcteur est nul.

(11) - Concrètement, cela signifie que nous considérerons d'une part une population sociologiquement et géographiquement bien définie, comme peut l'être, en première approximation, la population urbaine française ; d'autre part composée de familles de même taille et structure. Anticipant sur les conclusions de la section 4, nous donnerons certains résultats pour une population comprenant des familles de tailles différentes, en raisonnant sur des données par tête (ou par unité de consommation ; cf. 4) cela suffit en général à éliminer l'influence perturbatrice de la taille du ménage.

(12) - P.A. SAMUELSON ; Foundations of Economic Analysis ; Harvard, 1947, Introd. et chap. III.

(13) - Cf. H.S. HOUTHAKKER : la forme des courbes d'ENGEL - cahier du séminaire d'économétrie, vol. II, 1953, C.N.R.S.
PRAIS, S.J. HOUTHAKKER, H.S. ; the Analysis of family budgets, 1955, chap. 2 et 7.

(15) - La forme des courbes d'ENGEL pour de faibles valeurs du revenu a été examinée par Y. MAINGUY, à partir d'enquêtes limitées faites auprès de volontaires en France en 1943. Ces observations mettent en doute le fait que l'abscisse correspondant à l'introduction d'une consommation dans le budget soit autre que nulle pour des grands groupes de dépenses : alimentation, etc... Pour de tels groupes les courbes passeraient par l'origine dans tous les cas. Cela ne met évidemment pas en cause la possibilité d'observer une abscisse à l'origine positive pour des produits particuliers. Y. MAINGUY : la Consommation incompressible, Economie Appliqu'e : N° 1 - 1948.

- Détermination statistique de la consommation incompressible ap. cahiers de l'I.S.E.A., série B ; annexe 4, 1948 - Nouvelle contribution à l'étude du comportement du consommateur, cahiers de l'I.S.E.A., Série B, annexe 5, 1950.

(16) - G.D.N. WORSWICK et D.G. CHAMPERNOWNE : A note on the adding - up criterion - Review of Economic Studies, vol. 22, 1954.

(17) - Il est surprenant de remarquer que la seule loi qui n'ait pratiquement jamais été expérimentée est la logistique, que suggérerait assez naturellement la figure III-1. La difficulté d'en estimer analytiquement les paramètres est probablement l'explication de cet oubli.

(18) - Cité WOLD et JUREEN : Demand Analysis, p. 107 et 345.

(19) - Cf. GOREUX : L'élasticité de la demande de produits alimentaires par rapport au revenu. Document ronéoté, diffusion restreinte, AGRI/WP-7/2, Commission Economique pour l'Europe, 22 Juin 1959, IIème partie, 2, section 4.

(20) - Les quatre premiers modèles ont été étudiés avec quelque détail par PRAIS et HOUTHAKKER (op. cit.) Comme exemples de (II-11), on peut citer :

$$C_i = a_k R_i^2 + b_k R_i + d_k + u_{ik}$$

$$C_i = a_k (\log R_i)^2 + b_k \log R_i + d_k + u_{ik}$$

GOREUX op. cit.

(21) - Cf. H. FAURE : Abaques pour l'utilisation des élasticités, Consommation N° 1 - 1960.

(22) - Ces formules deviennent respectivement :

$\frac{a_k M}{R}$ et $a_k \frac{M}{C_k}$ avec $M = \log_{10} e$ quand on utilise des logarithmes décimaux

au lieu de logarithmes népériens pour effectuer les calculs.

(23) - Quand on utilise des logarithmes décimaux. Par k si on utilise des logarithmes à base k.

(24) - En outre, dans tous les cas, le coefficient de corrélation a peu de signification pour des données groupées, puisqu'il ne tient compte que de la variance des consommations entre les groupes et non dans les groupes.

(25) - Ce problème est étudié dans MALINVAUD : Méthodes Statistiques de l'Econométrie, tome II, pp. 41-44.

(26) - Cf. EZKIEL et FOX, Methods of correlation and Regression Analysis, chap. 20, pp. 337-343.

(27) - Cf. R. STONE, The measurement of consumers Expenditure and Behaviour in the United Kingdom, pp. 287 seq.

(28) - Op. cit. pp. 51-55

(29) - MOOD, A.M. - Introduction to the Theory of Statistics, New-York, 1950, pp. 391-394.

(30) - Le nombre D de groupes est supérieur d'une unité au nombre de changements de signes des $\hat{u}_i$.

(31) - AITCHISON et BROWN : A Synthesis of ENGEL Curve theory - Review of Economic Studies, 1954.

(32) - The log-normal distribution, Cambridge, 1957.

(33) - Des logarithmes népériens sont utilisés dans ces formules.

(34) - L'élasticité de la dépense alimentaire par rapport au revenu, op. cit.

(35) - A Synthesis of ENGEL Curve Theory, op. cit.

(36) - VORANGER : L'élasticité des dépenses des ménages, Annales du C.R.E.D.O.C. N° 2 - 1956.

(37) - L'élasticité de la dépense alimentaire par rapport au revenu, op. cit.

(38) - The log-normal distribution, pp. 79-82.

(39) - La projection des dépenses de consommation par produit est souvent établie à partir d'hypothèses sur l'augmentation de la consommation totale et non des revenus disponibles des ménages. Cela conduit à rechercher une estimation correcte de la liaison entre les consommations totales et la consommation partielle.

(40) - A Theory of the Consumption function, op. cit.

(41) - NERLOVE ; M. : The implications of FRIEDMAN'S, permanent income hypothesis for demand Analysis;Agricultural Economics Research.

(42) - 3 semaines pour tous les produits dans l'enquête Britannique de 1953.

(43) - Comme dans l'enquête française de 1956.

(44) - R. SUMMERS : A note on least squares bias in household expenditure analysis - Econometrica, janvier 1959 - pp. 121-126.
Cet article a appelé dans le même numéro une réponse de PRAIS (pp. 127-129) dont les arguments nous ont paru assez faibles.

(45) - Comme nous nous intéressons seulement à l'examen d'un biais au voisinage du point moyen, nous utiliserons des relations linéaires afin de simplifier l'exposé. Les résultats s'étendraient facilement à d'autres formes de relations.

(46) - Cf. A MADANSKY : the fitting of straight lines when both variables are subject to errors. J. Amer. Stat. Assoc., 1959.
LIVIATIAN, N. : Errors in variables and ENGEL Curve Analysis - Communication au Congrès de Naples de la Société d'Econométrie, 1960, pages 26-27

(47) - Nous avons utilisé ce mode de groupement dans l'exploitation de l'enquête de 1956 (cf. Chap. IV), mais sans nous rendre compte de ses avantages du présent point de vue. La nature des données l'imposait alors. Il faut remarquer que ce mode de groupement allège notablement les opérations d'exploitation de l'enquête : il suffit de calculer la consommation totale pour un petit nombre de groupes de ménages au lieu de le faire pour chaque questionnaire séparément.

(48) - LIVIATIAN, op. cit., pp. 4-6.

(49) - La méthode des variables instrumentales est décrite par MALINVAUD : méthodes statistiques de l'économétrie, tome II, pp. 18 seq.

(50) - The analysis of family budgets, op. cit., p. 101.

(51) - Le questionnaire d'enquête et les états mécanographiques permettent cependant une analyse plus détaillée.

(52) - PRAIS, S.J. et HOUTHAKKER, H.S. : les variations de qualité dans les budgets de famille. Economie Appliquée, Janvier-Mars 1952.

(53) - G. BLACK : Variations in prices paid for food by income level - Journal of Farm Economics, 1952.

(54) - R. STONE et alia : the measurement of consumers' expenditure and behaviour in the United - Kingdom, vol. 1 Cambridge, 1954.

(55) - MARSCHAK, J. : Money Illusion and demand Analysis, Review of Economic statistics, 1943 ; cité par L. KLEIN, Econometrics, 1953, p. 241.

(56) - TOBIN : Statistical demand Function for food in the U.S.A. - Journal of the Royal Statistical Society, 1950.

(57) - VORANGER J. : Le facteur qualité dans l'analyse de la demande ; Consommation, N° 4 - 1958.

(58) - Cf. GOREUX, op. cit.
et G.R. ROCKWELL, J.R. : Income and household size ; their effects on food consumption - Marketing Report N° 340 ; U.S. département of Agriculture, 1959.

(59) - Un avantage de cette échelle est d'assurer une liaison biunivoque entre le nombre d'unités de consommation et les structures familiales les plus courantes (l'ambiguïté ne s'introduit qu'à partir des familles de un adulte et 7 enfants, qui ont la même mesure que six adultes) - Cette particularité est commode puisqu'elle permet de trier selon le type de familles en n'utilisant que deux colonnes d'une carte perforée.

(60) - Notons que les échelles d'unités de consommation sont l'occasion de malentendus constants entre nutritionistes, cherchant à estimer les besoins alimentaires d'individus, les sociologues, cherchant à comparer des niveaux de vie, et les économètres, qui visent seulement à tourner une difficulté statistique.

(61) - Des tableaux analogues, mais pour l'alimentation, se trouvent dans : BROWN ; the Consumption and income. Econometrica, 1954.

(62) - Les modèles (IV-5) à (IV-7) ont été comparés par GUILBAUD : "La réduction des dépenses à l'unité de consommation", ap. cahiers de l'I.S.E.A., série B, annexe 4, pp. 56-63.

(63) - G.R. ROCKWELL, op. cit., p. 40
GOREUX, op. cit., pp. 37-38, indique à tort que le coefficient de log n mesure uniquement l'économie d'échelle.

(64) - R.G.D. ALLEN : Expenditure patterns of families of different sizes (ap. Memorial Schultz, 1942)
HOUTHAKKER : the econometrics of family budgets, J.R. S.S. 1952, p. 9
PRAIS et HOUTHAKKER : Family Expenditure, chap. 9

(65) - HICKS, J.R. : Value and Capital, chap. 2

(66) - On peut simplifier la méthode et se ramener à un système linéaire en estimant préalablement l'élasticité revenu a_k par une analyse dans les types de ménages.

(67) - Op. cit. p. 803.

(68) - "L'élasticité des dépenses des ménages...", op. cit., pp. 47-48

(69) - Une solution rigoureuse a seulement pu être trouvée dans le cas où l'élasticité du bien considéré par rapport au revenu est nulle : cf. QUENOUILLE, M.H. : An Application of least squares to family budget surveys, Econometrica, 1950.

(70) - Cf. p. ex, G. STUVEL and S.F. JAMES : Household Expendidure on food in Holland ; Journal Royal Statist. Society, 1950.

(71) - Cf. GOREUX, op. cit. pp. 32-35.

(72) - Cf. ROTTIER : Quelques problèmes posés par l'analyse des grandes enquêtes de consommation. Communication au Congrès Européen de la Société d'Econométrie, Amsterdam, Septembre 1959.

(73) - G. ROTTIER : Consommation et niveau de vie de la population agricole ; Consommation, N° IV - 1959.

(74) - Ce paragraphe résume un texte déjà publié : G. ROTTIER Niveau de vie et consommation de la population non agricole ; Consommation N° III-1959.

(75) - Pour alléger la rédaction, nous désignerons indifféremment dans la suite par les termes "niveau de vie", revenu ou dépense totale par unité de consommation, la variable explicative utilisée qui est la valeur moyenne dans chaque cellule de la consommation totale par unité de consommation. Cette consommation totale est la somme de la dépense totale de consommation et de la valeur, estimée aux prix de détail, de l'autoconsommation et des autofournitures. Puisqu'on n'étudie que la population non agricole, l'autoconsommation porte presque exclusivement sur les produits des jardins familiaux. L'autofourniture correspond surtout à la consommation de produits prélevés sur leurs stocks par des détaillants.

(76) - Cf. G. ROSCH : les dépenses médicales en 1956 ; Consommation N° III-1958.

(77) - Cf. Cl. SEIBEL : Les dépenses de vacances des Français ; Consommation N° IV-1958.

(78) - L'élasticité des dépenses d'utilisation en fonction du revenu chez les propriétaires d'automobiles a été étudiée par H. FAURE et E.A. LISLE : Les dépenses d'automobile des ménages français ; Consommation, N° II-1959.

(79) - Il n'y a que deux exceptions peu importantes : les dépenses de transport et vacances pour la catégorie I et les consommations prises au café pour la catégorie 2.

(80) - Egale à la moyenne géométrique des niveaux de vie pour l'ensemble de la population non agricole.

(81) - Ce coefficient a_o est très voisin du coefficient a qui figure pour le même produit dans le tableau III-D. La différence traduit seulement le fait qu'on a modifié la définition de la variable explicative.

(82) - Ces lois sont publiées dans G. ROTTIER : Consommation et niveau de vie de la population agricole, op. cit.

(83) - L'importance du facteur qualitatif peut être observée en partie en comparant les consommations moyennes de pain et de sucre dans les différentes sous-populations. Comme il s'agit de produits dont l'élasticité par rapport au revenu est négligeable et dont l'élasticité-qualité est faible, la comparaison peut aider à interpréter les différences observées dans le niveau des courbes d'ENGEL :

Consommation moyenne par U.C. 100 Frs par an

	Catégories de communes		
	I	4	5
Pain	110	95	78
Sucre	33	26	24

(84) - Les données proviennent de "La consommation de fruits et légumes" ; Consommation N° IV-1959, p. 53.

(85) - Cf. Claude SEIBEL : Les jardins et élevages familiaux en France ; Consommation, N° IV-1958.

(86) - L'autoconsommation moyenne de fruits par la population non agricole des communes rurales avait une valeur moyenne de 4 000 Frs par ménage et celle de légumes 15 000 Frs.

(87) - Il s'agit d'une moyenne arithmétique dans le premier cas et d'une moyenne géométrique dans le second.

(88) - Par contre, l'augmentation du revenu moyen de l'ensemble de la population ne sera pas la même, puisque l'urbanisation modifiera la répartition de la population entre des groupes dont les revenus moyens sont différents.

(89) - Elles ne sont jamais significatives dans le cas des consommations non alimentaires et elles le sont dans seulement 12 cas sur 120 dans le cas des consommations alimentaires.

C O N C L U S I O N

Dans les pages qui précèdent, nous avons examiné quelques aspects d'un problème très limité d'analyse économique inductive : la recherche de la mesure la plus convenable de l'influence du revenu des consommateurs individuels sur leurs achats de biens de consommation courante. Nous sommes partis d'une hypothèse théorique très simple et très générale qui nous a aidé à formuler un modèle permettant l'étude des résultats d'une enquête de consommation. Comme dans toute analyse économétrique moderne, le modèle choisi a été exprimé sous une forme telle que les résultats obtenus soient valables, en probabilité, pour définir la structure des comportements dans une population hypothétique infinie, dont la population française de 1956, d'où a été tiré l'ensemble des ménages étudiés, n'est elle-même qu'une image partielle.

Ce dessein très humble et limité est en fait apparu trop ambitieux. Dans un domaine aussi simple et largement exploré, de nombreuses questions primordiales n'ont pas encore pu recevoir de solution complète, ni même parfois d'esquisse de solution généralement acceptée.

On qualifie souvent l'économétrie de dialogue entre la théorie d'un phénomène et les observations, l'instrument de ce dialogue étant fourni par les techniques de l'induction statistique : ces techniques permettent à la fois la vérification et la spécification numérique de relations que la théorie économique permet de poser en hypothèses.

Il s'agit là cependant plus d'un programme que d'une réalité. Le caractère souvent peu opératoire de la théorie économique, l'insuffisante spécificité de techniques statistiques que l'économètre emprunte fréquemment à d'autres domaines, l'imprécision ou l'inadéquation des sources statistiques analysées imposent encore de sérieuses limites à sa réalisation.

Ces limites se sont présentées à chaque étape de notre exposé. Il peut être bon maintenant d'en résumer les principaux aspects en les ordonnant autour de trois thèmes : la distinction entre la vérification d'une hypothèse et l'estimation des paramètres d'une loi - la nature et la signification des hypothèses introduites - la relation entre les données de l'expérience et les modèles l'interprêtant.

Vérification et estimation

Au stade de l'exploitation de l'enquête où nous nous trouvons, nous avons été contraints de travailler uniquement sur des données groupées. Nous avons montré que cela affectait de façon négligeable les paramètres des lois ajustées. Mais nous avons en même temps remarqué que les coefficients de détermination de ces lois n'avaient aucun sens, puisqu'ils ne tenaient compte que de la variance des consommations entre groupes et non de leur variance à l'intérieur des groupes. Comme le coefficient du revenu dans l'équation des courbes d'ENGEL est significatif dans presque tous les cas, il est certes légitime d'admettre que l'hypothèse théorique que nous avons introduite au premier chapitre est vérifiée : le revenu influence effectivement la consommation.

Cependant cette hypothèse est implicitement plus forte : la plupart des applications concrètes des courbes d'ENGEL supposent non seulement que le revenu influence la consommation, mais qu'il soit en outre le, ou au moins un des facteurs dominants qui l'expliquent.

Or des résultats récents sont troublants de ce point de vue. Le développement des moyens de calcul modernes permet, en effet, d'ajuster sans grande difficulté des courbes d'ENGEL à partir de données individuelles ; dans tous les cas, les ajustements obtenus sont très médiocres. Nous donnerons en exemple les estimations présentées par ROCKWELL (1) et par LIVIATIAN (2), de coefficients de détermination calculés à partir de données individuelles et de données groupées (3)

---

(1) - op. cit. p. 50.53

(2) - op. cit.

(3) - Dans une seconde étape de l'analyse de l'enquête française de 1956, nous avons fait engager des travaux dans la même direction que les auteurs cités. Les résultats, encore trop provisoires pour être publiés, semblent conduire à une conclusion voisine.

LIEN ENTRE LA CONSOMMATION ET LE REVENU

Valeur du coefficient de détermination R^2

1° - ROCKWELL, p. 53

Source : enquête de 1955, Etats-Unis

Variable explicative :	Revenu disponible	
Données :	Groupées	Individuelles
Total alimentaire	0,98	0,25
Lait et fromage	0,93	0,06
Viandes	0,98	0,21
Légumes frais	0,92	0,03

2° - LIVIATIAN, p. 20

Source : enquête de 1956-57, Israël

Variable explicative :	Revenu		Consommation totale	
Données :	Groupées	Individuelles	Groupées	Individuelles
Alimentation	0,87	0,22	0,87	0,35
Habillement	0,72	0,09	0,88	0,31
Biens durables	0,72	0,02	0,91	0,41
Hygiène et soins	0,79	0,08	0,89	0,14
Dépenses diverses	0,86	0,18	0,91	0,37

Dans les cas les plus favorables, le niveau de vie n'expliquerait que 25 % à 30 % de la variance de la consommation d'un produit ou d'un groupe de produits. La conclusion que ces résultats suggèrent est qu'au niveau individuel, la consommation serait fondamentalement aléatoire. Les lois d'ENGEL, bien qu'elles reposent sur une hypothèse relative aux comportements individuels, n'auraient de signification que comme une tendance pour des groupes de la population : elles correspondraient au moins autant à des régularités de caractère sociologique qu'économique.

Cette conclusion est plausible ; les travaux disponibles sur le sujet ne sont cependant pas assez sûrs pour la confirmer sans précautions. Aucun des auteurs n'a tenu compte correctement de la forme des courbes d'ENGEL dans l'analyse et les résultats établis à partir de données individuelles correspondent à des régressions linéaires simples sans transformation des variables ; les erreurs de mesure sont proportionnellement plus fortes sur des données individuelles que sur des groupes ; la courte période sur laquelle sont habituellement relevées les consommations accroît la variance des consommations individuelles, etc...

Il est cependant certain que le revenu courant n'explique pas la majeure partie de la variance des consommations individuelles. Si les paramètres des lois d'ENGEL sont bien déterminés, la portée de ces lois mérite encore un examen. Il est en outre frappant que l'influence du revenu soit d'autant moins nette qu'on considère des données moins globales. Si le coefficient de détermination est plus faible, à partir de la même enquête, dans le cas de données individuelles que de données groupées, on observe de même que la liaison entre le revenu et la consommation à partir des données groupées d'une enquête est, elle-même souvent plus faible que celle qu'on obtient en reliant la consommation moyenne et le revenu moyen d'un pays à partir de séries chronologiques.

La signification des hypothèses

Toute étude économétrique doit traduire une hypothèse explicitement formulée dans un modèle. Il s'agit là d'un point de doctrine récent, mais fondamental, même s'il n'a pas été accepté sans conflit (qu'on se rappelle la controverse entre le National Bureau of Economic Research et la Cowles Commission entre 1945 et 1950). Quelques lignes d'un article de M. MALINVAUD l'explicitent

très clairement : "Aucune induction scientifique ne saurait être rigoureusement justifiée sans le recours à un cadre d'hypothèses préalables. Tout statisticien sait bien que le test d'une hypothèse et l'estimation d'une loi ne se conçoivent qu'à l'intérieur de modèles aléatoires plus ou moins généraux. Cette vérité générale revêt une importance particulière en économie. L'indigence des données disponibles, la complexité des phénomènes étudiés ont pour conséquence qu'une induction fructueuse nécessite le plus souvent un modèle assez étroitement défini, et par suite assez limitatif" (1).

Mais cela ne définit pas la nature et la signification du cadre d'hypothèses qu'on invoque. Or il n'y a pas, d'une part, les méthodes de l'induction statistique, d'autre part un corps de doctrine rigoureusement unifié sur une base axiomatique, qui serait "la théorie économique" et qui déduirait d'un ou de quelques postulats de maximation l'ensemble des hypothèses sous-jacentes au modèles. L'exemple que nous avons traité montre une situation plus complexe, plus empirique et moins logiquement satisfaisante. Le recours à la théorie pure des choix du consommateur a permis, au mieux, de dégager les grandes lignes du raisonnement et de choisir, sans en préciser l'expression, les principales variables qu'il convenait de considérer ; il n'a permis de suggérer aucune restriction empiriquement vérifiable sur la façon dont elles jouent. Pour traiter le problème posé, il a fallu introduire successivement de nombreuses hypothèses complémentaires. Certaines, relatives par exemple à l'influence de la taille du ménage ou de la catégorie de communes, ou bien au choix de la variable explicative, ont surtout traduit des considérations de sens commun sur le comportement des consommateurs, complétant ce qu'avait de trop abstrait la formulation initiale de la théorie. D'autres hypothèses, comme celles relatives à la variation de l'élasticité avec le revenu qu'implique le choix de la forme des courbes d'ENGEL, correspondent à des contraintes dont la signification économique est assez claire, mais peu intéressante ; elles ont surtout été introduites pour des raisons de commodité. Enfin, d'autres hypothèses, qui se réfèrent surtout à la part aléatoire du modèle, ont un caractère plus inquiétant : elles n'ont pas de traduction économique simple, ni même toujours intelligible. Elles ne sont introduites que pour justifier le recours à des techniques d'estimation habituellement empruntées à d'autres domaines d'application de la statistique. C'est là un aspect peu satisfaisant de l'économétrie dans son état actuel : la théorie est formulée de façon à permettre l'application de techniques connues, au lieu que les techniques soient adaptées à l'exploration d'une hypothèse complète. Cet inconvénient n'est cependant pas particulier à l'économie inductive et y remédier exigera des efforts à très long terme.

---

(1) - E. MALINVAUD : L'âge mûr de la Comptabilité Nationale. Journal de la Société de Statistique de Paris, Octobre 1960, p. 237.

La théorie et l'expérience

Enfin, la nature des données empiriques limite aussi les possibilités de vérifier convenablement certains éléments du schéma théorique proposé. Le processus logique serait évidemment de formuler la théorie, de chercher les données quasi-expérimentales permettant de la soumettre à l'épreuve, en utilisant les techniques d'induction statistique les mieux adaptées. Or non seulement les techniques d'induction, mais les observations disponibles se sont avérées être des données dont la nature et la forme étaient fixées avant que les hypothèses à étudier aient été formulées. Cette situation est bien connue et irrémédiable dans le cas où l'on analyse des statistiques de nature administrative, comme des séries chronologiques de consommation et de prix observées sur un marché. Il est important de remarquer que la même situation est apparue ici. Non seulement des erreurs de mesure et d'observation se sont introduites - elles résultent de l'imperfection des techniques d'enquête et elles ne pourront être réduites à l'avenir que moyennant des efforts coûteux et prolongés - mais le plan qu'enquête lui-même a étroitement limité la possibilité de donner des réponses significatives à certaines questions dont l'étude a montré qu'elles pourraient être fondamentales.

Il y a à cela des raisons en partie administratives et financières, telles que les habitudes des organismes d'enquête, l'insuffisance de la documentation permettant de définir un plan de sondage et la nécessité de viser, dans la même opération, plusieurs objectifs qui exigeraient des techniques distinctes. Tout cela est peu intéressant.

Ce qui l'est plus, par contre, c'est que la méthode qui a été choisie n'aurait probablement pas pu être fondamentalement modifiée sans une expérience préalable.

Nous sommes parvenus dans les chapitres précédents à deux conclusions :

- la première est que le plan de sondage aurait gagné à s'inspirer d'un plan expérimental où la disposition et les effectifs des différentes cellules auraient permis de séparer efficacement l'influence de facteurs quantitatifs ou qualitatifs statistiquement liés les uns aux autres.

- La seconde est que les habitudes de consommation des exploitants agricoles ne peuvent probablement pas être dissociées de leur comportement de producteurs.

Ces deux remarques ne pouvaient cependant découler que de l'observation. Ce n'est en particulier qu'à partir de l'étude, que nous avons résumée, d'une enquête générale faite sans idée précise à l'origine que les lignes selon lesquelles un schéma quasi-expérimental d'enquête devrait être dessiné ont pu commencer à se dégager.

Cela conduit à nuancer l'image du raisonnement économétrique que nous avons esquissée plus haut. Il n'y a pas un processus linéaire allant de la théorie à l'estimation en passant par les observations. Le dialogue entre "la théorie", c'est-à-dire un ensemble d'hypothèses souvent partielles et peu cohérentes, et les diverses observations comprend de nombreuses phases. Il s'agit d'un cheminement par approximations successives qui suppose beaucoup de temps et d'inévitables échecs avant qu'une esquisse de conclusion puisse être dégagée. Là encore, il n'y a rien de particulier à l'économétrie. Qu'on y observe ce processus d'une façon de plus en plus générale peut être le signe que l'économie inductive commence à acquérir le caractère d'une science.

B I B L I O G R A P H I E

Anonyme — La consommation de fruits et légumes
Consommation, N° 4 - 1959

Anonyme — Economie et Consommation
Revue et l'Action Populaire, Juillet-Août 1960

AITCHISON, J. et BROWN, J.A.C.
A Synthesis of Engel curve Theory
Review of Economic Studies - 1954

AITCHISON, J. et BROWN, J.A.C.
The lognormal Distribution
Cambridge, 1957

ALBERT, J.F. - ROTTIER, G. - TABARD, N.
La demande de logements en France
Annales du Centre de Recherches et de Documentation sur la Consommation, N° 4 - 1957

ALLEN, R.G.D. Expenditure Patterns of families of different sizes ap. Studies in honor of Henry SCHULTZ, Chicago-1942

ALLEN, R.G.D. et BOWLEY, Al.
Family Expenditure, London-1935

BLACK, G. — Variations in prices paid for food by income level
Journal of Farm Economics - 1952

BOULDING K. A Reconstruction of Economics - 1950

BRADY, D. et FRIEDMAN, M.
Savings and the Income Distribution
ap. Studies in Income and Wealth, Vol. 10 ; 1947.

BROWN, J.A.C. The Consumption of food in relation to household composition an income
Econometrica, 1954

CHOMBART de LAUWE, Ph.
Conditions de vie des familles ouvrières
C.N.R.S., 1958

CLOSON, F.L. Avant-propos
Consommation N° 1 - 1958

CRAMER, J.S. La comparaison internationale des enquêtes de budgets familiaux
Revue Economique, N° 2 - 1961

CRAMER, J.S. The ownership of major consumer durables
Thèse Amsterdam, 1961.

CREDOC-SEEF La consommation de 1950 à 1957
Consommation, N° 2 - 1958

CREDOC-SEEF La consommation des particuliers de 1950 à 1958
Consommation, N° 3 - 1959

DESABIE, J. Theorie et pratique des sondages
Tome I, Ecole d'Application de l'INSEE - 1959

DUESENBERRY Income, Saving and the Theory of Consumer Behavior Harvard, 1949

ENGEL, E. Die Productions und Consumptions verhältnisse des Königreich Sachsen 1857
Rep. Bul. I.S.I., Vol. IX.

ENGEL, E. Die Lebenkosten Belgischer arbeiter familien früher und jetzt, 1895
Rep. Bul. I.S.I., Vol. IX.

EZEKIEL, M. et FOX, K.
Methods of Correlation and Regression Analysis Wiley, 1959

FARRELL The Demand for motor cars in the U.S.
Journal Royal Statistical Society, 1954

FAURE H. Un modèle prospectif du marché de l'automobile
Consommation, N° 4 - 1959

FAURE H. Abaques pour l'utilisation des élasticités
Consommation, N° 1 - 1960

FAURE, H. et LISLE, E.A.
Les dépenses d'automobile des ménages français
Consommation, N° 2 - 1959

FAURE H. et SEIBEL, Cl.
L'étude régional de la Consommation
Consommation, N° 1 - 1959

FRIEDMAN, M. A Theory of the Consumption Function
Chicago, 1957

GALBRAITH, J.K. The Affluent Society
New-York, 1958

GOREUX, L. Perspectives des dépenses d'habillement des Français
Etudes et Conjoncture, Octobre 1956

GOREUX, L.M. Comparaison internationale des courbes d'ENGEL
Communication au Congrès Européen de la Société d'Econométrie -
Luxembourg, 1957

GOREUX, L.M. Income Elasticity of the demand for food,
E.C.E., Agri (W.P.7) 2 - Juin 1959

GUILBAUD, G.Th.
Sur les méthodes d'étude des budgets de ménages
ap. Cahiers de l'I.S.E.A., Série B - annexe 4, 1947

GUILLOT, M. Le crédit à la consommation dans les budgets familiaux
Consommation, N° 4 - 1958

HICKS, J.R. Value and Capital
Oxford, 1939

HOOD, W.C. et KOOPMANS, T.C.
Studies in Econometric Method
Wiley, New-York-1953

HOUTHAKKER, H.S.
The Econometrics of Family Budgets
Journal Royal Statistical Society, 1952

HOUTHAKKER, H.S.
La forme des courbes d'ENGEL,
Cahiers du Séminaire d'Econometrie, C.N.R.S. Vol. II, 1953

HOUTHAKKER, H.S.
An International Comparison of Engel's law,
Econometrica, 1957, N° 1

INSEE-CREDOC.
Budgets des Français en 1956, Dépenses et niveaux de vie
Dunod, 1960

KOOPMANS, T.C.
Three essays on the State of Economic Science
Mc. Graw Hill - 1957

KENDALL, M.G. Regression, structure and the functional relationship
Biometrica, 1951

KLEIN, L. Econometrics, E
Evanston, New(York, 1953

LISLE, E.A. La prévision de la consommation,
Revue Economique, N° 2 - 1961

LISLE, E.A. et VORANGER, J.
Le marché de la chaussure en France
Annales du Centre de Recherches et de Documentation sur la
Consommation, N° 4 - 1957

LISLE, E.A. ; MASSONAUD, M. ; TABARD, N.
L'étude de l'épargne
Consommation, N° 4 - 1960

- LISLE E.A. : L'épargne et l'épargnant : Orientations théoriques récentes,

- MASSONAUD, M. : L'enquête-pilote de 1959 : Méthodologie

- LISLE, E.A. ; MASSONAUD, M. ; TABARD, N. : L'enquête-pilote de 1956 : l'analyse des résultats

- TABARD, N. : L'enquête-pilote de 1959 : les relations entre le revenu, le patrimoine et l'épargne

LIVIATAN, N. Errors in variables and Engel Curve Analysis Communication au Congrès Européen de la Société Internationale d'Econométrie, Naples, 1960, ronéo 30 pp.

MACK (Ruth P.)
Economics of Consumption
ap. A Survey of Contemporary - Economics, Vol. II
BF. HALEY Editor, 1952, p. 32 seq.

MADANSKY, A. The fitting of straight lines when both variables are subject to errors.
Journal American Statist. Associat. 1959

MAINGUY, Y. Détermination statistique de la consommation incompressible.
ap. Cahiers de l'I.S.E.A., Série B, Annexe 4, 1947

MAINGUY, Y. La consommation incompressible
Economie Appliquée, N° 1 - 1948

MAINGUY, Y. Nouvelle contribution à l'étude du comportement du consommateur
Cahiers de l'I.S.E.A., Série B, Annexe 4, 1950

MALINVAUD. E. Initiation à la Comptabilité Nationale - I.N.S.E.E. - P.U.F. 1957

MALINVAUD. E. Les méthodes statistiques en Econométrie - 2 volumes - Ecole d'Application de l'I.N.S.E.E. 1959-1960

MARSCHAK, J. Money Illusion and Demand Analysis
Review of Economic Studies - 1943

MOOD, A.M. Introduction to the Theory of Statistics, New-York, 1950.

MORICE, J. (Melle)
La demande d'automobiles en France,
A. Colin, Paris, 1957.

NERLOVE, M. The implications of FRIEDMAN'S,
Permanent income hypothesis for demand analysis
Agricultural Economics Research

PRAIS, S.J. The estimation of equivalent adult scales from family budgets
Economic Journal - Décembre 1953

PRAIS, S.J. A comment (to summers'paper)
Econometrica, Janvier 1959

PRAIS, S.J. et HOUTHAKKER, H.S.
Les variations de qualité dans les budgets de famille
Economie Appliquée - Janvier-Mars 1952

PRAIS, S.J. et HOUTHAKKER, H.S.
The Analysis of Family budgets - Cambridge 1935

QUENOUILLE, M.H.
An Application of least squares to family budget surveys
Econometrica, 1950

ROCKWELL, G.R. Jr
Income and Household size ; their effects on food consumption
Marketing Research Report N° 340, U.S.D.A. 1959

ROSCH, G. Les dépenses médicales des Français en 1956
Consommation, N° 3, 1958

ROTTIER G. La distribution des revenus non agricoles,
Consommation, N° 1, 1958

ROTTIER G. Quelques problèmes posés par l'analyse des grandes enquêtes de consommation,
Communication, Congrès Européen de la Société Internationale d'Econométrie - Amsterdam, Sept. 1959

ROTTIER, G. Niveau de vie et consommation de la population non-agricole
Consommation, N° 3 - 1959

ROTTIER, G. Niveau de vie et consommation de la population agricole
Consommation, N° 4 - 1959

ROTTIER, G. La prévision de la consommation
Revue de l'Action Populaire, Juillet-Août 1960

ROTTIER, G. Initiation à l'économétrie,
C.R.E.D.O.C., 1961

ROTTIER, G. L'économie de la consommation
Revue Economique, N° 2 - 1961

ROTTIER, G. et SALEMBIEN, E.
Les budgets familiaux en 1956
Consommation N° 1 - 1958

SAMUELSON, P.A.
Foundations of Economic Analysis,
Harvard, 1947

SAVINO Un modello per prevision di circolazione automotociclista in Italia
Industria, 1957

SEIBEL, Cl. Les dépenses de vacances des Français
Consommation, N° 4 - 1958

SEIBEL, Cl. Les jardins et élevages familiaux en France,
Consommation, N° 4 - 1958

STIGLER, G.J. The early history of empirical studies of consumer behavior
Journ. Pol. Economy, 1954

STONE, R et al.
The measurement of consumers' expenditure and behavior in the United Kingdom -
Vol. I - Cambridge, 1954

STUVEL, G. & JAMES, S.F.
Household expenditure on food in Holland
Journal of the Royal Statistical Society - 1950

SUMMERS, R. A note on least-squares bias in household expenditure analysis
Econometrica, Janvier 1959

TABARD N. Les variations saisonnières des dépenses des ménages
Consommation N° 3 - 1958

TOBIN, J. Statistical Demand Function for food in the U.S.A.
Journal Royal Statistical Society - 1950

VORANGER, J. L'élasticité des dépenses des ménages d'après les enquêtes de 1951 et 1952
Annales du Centre de Recherches et de Documentation sur la Consommation, N° 2 - 1956

VORANGER J. Le facteur qualité dans l'analyse de la demande
Consommation N° 4, 1958

WOLD & JUREEN
Demand Analysis -
WILEY - 1953

WORSWICK & CHAMPERNOWNE, D.G.
A note on the adding up criterion
Review of Economic Studies, Vol. 22 - 1954

-:-:-:-:-:-:-:-:-:-

- 7 JUIN 1972

2 ex - n° 1