

MICHEL ARMATTE

**Pierre Thionet et l'introduction en France des méthodes de sondage aléatoire**

*Journal de la société française de statistique*, tome 144, n° 1-2 (2003), p. 227-255

[http://www.numdam.org/item?id=JSFS\\_2003\\_\\_144\\_1-2\\_227\\_0](http://www.numdam.org/item?id=JSFS_2003__144_1-2_227_0)

© Société française de statistique, 2003, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

# PIERRE THIONET ET L'INTRODUCTION EN FRANCE DES MÉTHODES DE SONDAGE ALÉATOIRE

Michel ARMATTE \*

## RÉSUMÉ

Après avoir rappelé les étapes d'une histoire des sondages de Laplace à Neyman (1934), notre apport se concentre d'abord sur l'introduction par Stoetzel des sondages d'opinion (pooling) et de la méthode des sondages (sampling) « à choix raisonné » dans les Instituts de Statistique français (IFOP, SNS de Carmille, Fondation A. Carrel) avant et pendant la Seconde Guerre Mondiale. Mais ce texte est principalement dédié à Pierre Thionet (1916-2002), récemment disparu, et à qui l'on doit, en liaison avec Raymond Lévy-Bruhl, l'introduction des sondages aléatoires à l'INSEE. Le premier publie à son retour d'Allemagne une synthèse critiques des méthodes mathématiques découvertes dans la littérature américaine, et qu'il commence à enseigner à l'École d'Application. Le second rapporte d'un stage aux États-Unis l'expérience du Census et des Administrations. Dans un climat de controverse avec l'INED et l'IFOP, tous deux mettent en place les premières enquêtes par sondage sur les budgets de famille (1946 et 1948), puis les plans de sondage de l'enquête emploi (1950). Le recensement de 1954 est l'occasion de réajuster ce plan de sondage, mais cette date est déjà celle du départ du tandem Thionet-Lévy Bruhl bientôt remplacé par Jacques Desabie à la tête du service.

## ABSTRACT

At first, we give an historical survey of sampling from Laplace to Neyman (1934), then we describe how Stoetzel introduced sampling and pooling, more exactly "purposive selection" and "opinion", in several French Institutes of Statistics (IFOP, SNS, Carrel Foundation) just before and during World War II. But this paper is principally dedicated to the late Pierre Thionet (1916-2002) who was responsible with Raymond Lévy-Bruhl of the introduction of random sampling at the new created INSEE. He published in 1946 a critical review of sampling methods in American Journals while Lévy-Bruhl was visiting the *Bureau of Census* and the *Bureau of Labor Statistics*. Both of them were then working at INSEE on the design of sample surveys on family budgets (1946-48) and on employment (1950-54) and were involved into some controversies with INED or IFOP statisticians who defended purposive sampling.

---

\* Université Paris-Dauphine – UFR Économie appliquée – Place du Maréchal de Lattre de Tassigny, F-75775 Paris Cedex 16. E-mail : michel.armatte@dauphine.fr

## 1. Introduction

Le texte qui suit <sup>1</sup> se veut une synthèse de différentes sources d'informations sur la question des premières enquêtes par sondage mises en place en France par l'INSEE, l'INED, l'IFOP et quelques autres institutions plus éphémères, entre 1938 date à laquelle Jean Stoetzel crée l'IFOP, et 1954, date à laquelle Pierre Thionet quitte le service des sondages de l'INSEE. Le décès récent de Pierre Thionet (le 27 avril 2002) nous a incité à en faire le personnage principal de notre narration, et à rendre ici un hommage tout particulier à ses travaux sur les sondages et à une œuvre mal connue. Ce texte s'appuie d'ailleurs sur diverses publications et documents d'archives, mais aussi sur des interviews très instructifs des principaux acteurs de la période, dont un entretien avec Pierre Thionet réalisé en 1996, mais aussi d'autres réalisés avec son prédécesseur et premier employeur Henri Bunle, avec son compère à l'INSEE Raymond Lévy-Bruhl, et avec son successeur au service des sondages Jacques Desabie <sup>2</sup>.

En situant le point de départ de cette étude dans la France occupée de la seconde guerre mondiale, nous ne prenons pas l'histoire des sondages à ses débuts et dans sa dimension internationale. Une première phase de cette histoire a déjà fait l'objet de nombreux travaux dus à : Seng (1951), Kruskal et Mosteller (1980), O'Muircheartaigh et Wong (1981), Desrosières (1988), Armatte (1988), Tassi (1988), Desrosières (1993 et 2000), Fienberg et Tanur (1996). Nous commençons par en rappeler les principales étapes.

## 2. Les enquêtes représentatives avant 1936

Les travaux de Laplace sur la théorie des erreurs et leur application à l'estimation de la population par le ratio population/naissance assorti d'un intervalle de confiance offrent les premiers fondements probabilistes d'un jugement sur échantillon à la fin du XVIII<sup>ème</sup> siècle. Mais l'intense activité de production de chiffres par les bureaux de Statistique dès la fin des années 1820 conduit les Administrations et la Statistique Générale de la France à privilégier les enquêtes exhaustives qui semblent fournir les cadres comptables principaux des activités sociales et économiques, tandis que les monographies de famille ou d'atelier telles que pratiquées par Le Play, Villermé, et Cheysson après 1850, appuyées sur l'établissement de budgets détaillés, offrent une information complémentaire d'une incomparable richesse mais qu'il s'avère

---

1. Une première version de ce texte a été rédigée et présentée au Séminaire de Recherche de l'INSEE le 9 mai 1996, puis diffusée en interne à l'occasion des cérémonies du cinquantenaire de l'INSEE. Cette première version participait à une réflexion sur *l'évolution des enquêtes « conditions de vie » à l'INSEE de la fin du XIX<sup>ème</sup> siècle à nos jours*, publiée en interne, et contenant également les contributions de Françoise Dussert, Alain Desrosières, Michel Glaude, et François Héran, qui doivent ici être remerciés.

2. Les entretiens utilisés sont les suivants : *Henri Bunle* par Alain Desrosières, *Raymond Lévy-Bruhl* par Alain Desrosières et Michel Armatte, *Pierre Thionet* par Jean-Claude Deville et Michel Armatte, *Jacques Desabie* par Alain Desrosières, Michel Glaude, François Héran, Françoise Dussert et Michel Armatte.

impossible d'agrèger. À l'enquête statistique exhaustive de constituer « le squelette » des phénomènes sociaux alors que la monographie en fournit « la chair », telle est la méthodologie duale prônée par Cheysson à la fin des années 1880.

La pratique des sondages représentatifs semble s'être développée assez tôt dans le contexte des enquêtes agricoles, et combiner ces deux traditions de la monographie et de l'enquête exhaustive. Emmanuel Didier (2000) a montré comment le Département d'Agriculture des États-Unis emploie dès 1863 des « crop reporters », fermiers ou non, volontaires et parfois élus, qui « représentent » la communauté des fermiers d'une circonscription, et informent l'administration sur les pratiques et les récoltes de celles-ci. Après 1921, les enquêteurs sont tous fermiers et ne représentent plus que leur propre ferme ; on est donc passé de l'opinion d'un témoin à la déclaration individuelle d'un acteur choisi comme représentant<sup>3</sup>. Après 1924, on fait appel aux postiers pour distribuer les questionnaires, et ce n'est plus le répondant qui est représentant d'une communauté, mais tous les fermiers, gros et petits, d'une portion de route constituée comme échantillon dont la moyenne sera représentative. En Norvège (Lie, 2001), les enquêtes agricoles faites au milieu des années 1870 par Kiaer et Mohn dans le cadre du Central Bureau of Statistics sont faites sur la base de 5 fermes représentatives de leur *township* et sont complétées par des études approfondies des registres de population pour 56 townships considérées chacune comme représentative de tout un district. En Russie, comme l'a montré Martine Mespoulet (2000), Kaufman, et Chuprov (Tchuprov) pratiquent dès les années 1880 une forme de tirages systématiques (« mécaniques ») parfois combinée au choix d'aires-types, tandis que la méthode des choix raisonnés est mise en œuvre par Gur'ev en 1900 dans l'enquête de Vjatka, avec un succès que confirme la concordance établie avec le recensement de 1880. Dans ces différents exemples, le statisticien semble hésiter encore pour construire son échantillon entre la représentativité de l'individu moyen typique au sens de Quetelet et de Le Play, et celle d'un échantillon doté d'une variabilité similaire à celle de la population par simple effet d'un tirage au hasard ou systématique.

Au tournant du siècle, des enquêtes partielles sur le paupérisme, puis sur les salaires, sur la condition ouvrière, sur les budgets et les consommations ou encore sur les assurances vont constituer le terreau d'une réflexion sur la représentativité des résultats établis sur des sous populations qui n'apparaissent pas encore comme des échantillons. Les enquêtes de Booth et Rowntree sont très artisanales et d'échelle modestes. Mais l'enquête norvégienne de Strom en 1888 sur la consommation de sucre, café et pétrole utilise un échantillon de type aléatoire de 18 000 ménages. L'enquête de Kiaer sur les revenus et les actifs des années 1890 porte sur un échantillon de 11 500 personnes, et la sélection combine encore un choix raisonné de villes et de communes, et le tirage systématique de personnes dans ces villes (selon la première lettre de leur nom). L'enquête du Board of Trade anglais sur les coûts comparés de

---

3. E. Didier rappelle que le premier sens de « sample » au *Department of Agriculture* fut celui d'un cas individuel exemplaire avant que d'être un échantillon de plusieurs personnes.

la main d'œuvre (BOT, 1909) lancée à une échelle européenne dès 1905, les enquêtes de Bowley à Reading en 1913, ou les enquêtes allemandes utilisées par Halbwichs dans sa thèse de 1913 s'adressent à de grands échantillons, mais ceux-ci entièrement constitués d'ouvriers des grandes villes sont choisis et consultés par le biais des organisations ouvrières syndicales ou coopératives. L'enquête SGF de 1913-14 menée par Dugé de Bernonville élargit son champ aux familles d'ouvriers, d'employés, d'artisans et de salariés agricoles mais l'échantillon important (12 500 questionnaires distribués, 3 700 récupérés, 2 000 exploités) n'est pas aléatoire, bien qu'on ait déjà recommandé un tirage systématique dans les listes nominatives du dénombrement de 1911 : la consigne ne fut pas imposée, donc peu suivie, et le choix des familles enquêtées a été laissé aux bons soins de maires, instituteurs, médecins et industriels, dont les rapports privilégiés aux enquêtés ne pouvaient pas garantir l'objectivité et l'homogénéité de l'opération.

Le débat sur la légitimité des enquêtes partielles est porté par Kiaer (1895) devant les instances de l'IIS à propos de la méthode représentative et reçoit les critiques conjointes de deux groupes : il y a ceux qui, comme von Mayr, rejettent *a priori* le principe de substitution d'un échantillon à l'enquête exhaustive et rejettent aussi toute approche mathématique de la question, et ceux qui, comme Bortkiewicz ou Hjorth, acceptent le principe de l'enquête partielle, mais prouvent par le calcul de probabilité que la méthode utilisée n'est pas représentative, les écarts de l'estimation à la vraie valeur ne pouvant être attribués au hasard mais bien à un biais systématique. La seconde des enquêtes de Kiaer est mobilisée pour mettre en place un système d'assurances, et la controverse technique s'inscrit alors dans les enjeux politiques et une rivalité professionnelle entre actuaires et statisticiens (Lie, 2000). Kiaer est mort depuis 6 ans quand le débat reprend à l'IIS et aboutit à la résolution de 1925, reconnaissant la validité des sondages sous les deux formes de la «purposive selection» et du sondage aléatoire (rapport Jensen et Bowley, 1926).

La controverse sur la possibilité même «*de prendre la partie pour le tout*» (Desrosières, 1988) fait place à une seconde controverse entre méthodes représentatives et méthodes aléatoires. Les premières méthodes, de type quotas, proposent simplement des garde fous «raisonnés» visant à limiter certains biais systématiques et à assurer *a posteriori* la représentativité. Par contre, les approches aléatoires se fondent sur les travaux probabilistes de Gauss et Laplace dont il faut rappeler qu'ils sont vers 1900 largement méconnus des enquêteurs sociaux, des sociologues, des économistes, voire des statisticiens eux-mêmes.

L'histoire des sondages est alors une histoire parallèle des pratiques administratives et des résultats en statistique mathématique. On doit d'abord à Bowley d'avoir reconstruit dans les éditions successives de son traité (Bowley, 1901) et dans sa *Presidential Adress* (Bowley, 1906) le minimum d'un savoir sur la théorie des sondages simples : loi de la fluctuation d'échantillonnage, mesure de la précision, intervalles de variation. Entre 1920 et 1925, la théorie des sondages se voit considérablement enrichie par les apports de

Ronald Fisher sur la théorie de l'estimation, qui introduisent la distinction entre paramètres de la population et statistique d'échantillon, les propriétés de ces dernières quand elles sont prises pour estimateurs, leur loi exacte pour des petits échantillons, la méthode du maximum de vraisemblance et son lien avec l'efficacité des estimateurs (Armatte, 1988). Dans la même période, l'analyse de variance et la randomisation dans les plans d'expérience sont introduites et vulgarisées à large échelle par la publication de *Statistical Methods for Research Workers* (1925). L'exposé de Bowley (1926) rédigé à l'occasion de la résolution Jensen à l'IIS, témoigne des progrès accomplis pendant les vingt ans qui séparent ce texte de sa *Presidential Adress*.

L'étape suivante commence avec la première visite de Jerzy Neyman à Londres en 1925 et la reconnaissance simultanée de ses écrits polonais (sous l'influence de Bernstein et Markov) et de ceux de Chuprov. La méthode des choix raisonnés, pratiquée par Gini et Galvani – et présentée à l'IIS (Gini, 1928) – pour sélectionner 29 districts sur 214 en utilisant le recensement italien de 1921, est discutée pas à pas et finalement disqualifiée par Neyman (1934) qui, dans le même article, introduit les intervalles de confiance et énonce les principes de la stratification optimale (Fienberg et Tanur, 1996). Sur cette question, Neyman reconnaîtra que ses résultats avaient été anticipés par l'École russe, et en particulier par Chuprov et Kovalevskij, une dizaine d'années plus tôt (Mespoulet, 2000). Sur la question des intervalles de confiance, la dispute avec Fisher qui défend sa méthode des intervalles fiduciaires sera d'autant plus vive qu'elle entre en résonance avec deux controverses sur les carrés latins et sur les tests statistiques, dans un contexte rendu sensible par la crise succession de Karl Pearson à *University College*. Avec la publication de *The Design of Experiments* (1935), Fisher inscrit cette question dans le cadre général de la logique inductive expérimentale. La séparation devient inévitable entre les problématiques liées aux plans d'expérience que Fisher met au point à Rothamsted, celles qui sont relatives à des décisions sur échantillon comme dans le contrôle de fabrication, ou encore celles qui accompagnent le traitement d'enquêtes par sondage dans les administrations, malgré les similitudes et les multiples échanges qui se poursuivent entre les mathématiques développées pour ces trois problèmes. L'épisode hexagonal que nous nous proposons de relater fait suite à ces avancées et controverses des années trente et en possède les multiples facettes bien qu'il se rattache essentiellement à la tradition des enquêtes.

### 3. Opinion, sampling et pooling en 1936-45

Si l'on revient en France, les enquêtes plus confidentielles de l'entre deux guerres – Antonelli à Lyon en 1921 sur le coût de la vie, Delpech à Toulouse en 1937 sur 100 ménages toulousains, Rist et Bunle à Paris, Lyon, et Mulhouse en 1936-37 auprès de familles de chômeurs restent plus proches dans leurs tailles et leurs objectifs de la monographie que de l'enquête statistique, et semblent ignorer ces avancées théoriques sur les sondages aléatoires. Comme le montre bien Blondiaux (1998), la technique des sondages (appelés d'abord *straw*

*polls*) ne naît pas dans l'univers académique mais dans celui du marketing et du journalisme. C'est le célèbre coup d'éclat de Gallup aux élections présidentielles américaines de 1936, qui fait passer la question des sondages du stade de la polémique à celui de l'expérimentation et de la prédiction en vraie grandeur, donc au statut de méthode scientifique aux yeux du plus grand nombre. Prévoyant à l'aide d'un échantillon judicieusement choisi de quelques milliers de personnes<sup>4</sup> la victoire de Roosevelt, alors que les « votes de paille » de plusieurs millions de lecteurs de grands quotidiens comme le *Literary Digest* donnaient pour vainqueur son adversaire Landon, Gallup eut la science (et la chance ?) de faire une juste prévision à faible coût, ce qui contribua grandement à la légitimité des sondages<sup>5</sup> mais plus malheureusement aussi à l'assimilation entre sondage et opinion, entre *sampling* et *pooling*, brouillant pour longtemps ce qui dans l'innovation était dû à la méthode (le jugement sur échantillon représentatif) et ce qui relevait de l'objet (l'opinion), et de sa possible ou impossible mesure.

Mais la méthode des sondages a aussi profité de cette confusion puisqu'elle est importée en France sous prétexte d'études d'opinions. Le premier qui ait introduit les sondages en France est incontestablement Jean Stoetzel, un agrégé de philosophie qui a fondé l'Institut Français d'Opinion Publique dès 1938, à son retour d'un voyage aux États-Unis. L'IFOP emprunte ses méthodes des instituts privés américains comme Gallup, et publie ses premiers résultats d'enquête sur la dénatalité sous la plume de Henri Paoletti dans les trois premiers numéros de sa revue *Sondages*, en 1939. Mais la guerre interrompt les activités de l'IFOP qui ne seront reprises qu'à la Libération. Jean Stoetzel transpose alors ses activités à la *Fondation française pour l'étude des problèmes humains*, une institution de recherche eugéniste fondée en novembre 1941 par Alexis Carrel avec l'aide de la Fondation Rockefeller, et rattachée au Secrétariat d'État à la famille (Drouard, 1983 et Girard, 1986). Dirigée par François Perroux puis J.J. Gillon, cette Fondation comptait 150 à 160 personnes à la veille de la suspension de Carrel en août 1944. Dès février 1942, la *Fondation* s'est dotée d'un *Service de Sondages et Statistiques* que dirige Stoetzel, avec l'aide de d'Alain Girard, ex-professeur de lettres au lycée d'Évreux et bibliothécaire à la Bibliothèque Nationale. Stoetzel et Girard mobilisent des données sanitaires ou scolaires, et conduisent à la *Fondation* des enquêtes principalement sur les questions démographiques et sanitaires : enquête sur la natalité auprès des jeunes femmes, grande enquête de 1944 sur les 100 000 enfants, qui sera exploitée plus tard à l'INED.

La Fondation Carrel n'est pas seule à pratiquer les sondages sous l'occupation. Dans les années 1941-44, l'activité de l'IFOP en matière d'opinion publique a été relayée par d'autres organismes privés qui se réclament comme lui de l'influence de Gallup. Un énigmatique *Service des Sondages et de Statistique*

---

4. 1 500 selon certaines légendes, près de 100 000 tout de même selon Blondiaux.

5. Cette « expérience décisive » fut d'autant plus probante que les écarts de prévision du *Literary Digest* et de l'Institut Gallup étaient énormes : 42,5 % contre 54 % pour Roosevelt, mais elle doit être relativisée par le fait que d'autres Instituts sondeurs confirmèrent la prévision de Gallup (comme Roper avec 3 000 interviews seulement) et que Roosevelt obtint en fait 61 % des suffrages.

(SSS) fondé en 1944 par Max Barioux dans la clandestinité, se dit issu de la Résistance et envoie ses enquêtes d'opinion à Alger et à Londres. La CEGOS, fondée par le polytechnicien Milhaud, s'est également dotée en 1943, sous l'impulsion de son directeur M. Poudroux, d'un bureau de statistiques et de sondages qui développe conjointement les études de marchés et les sondages d'opinion. L'assureur Jacques Dourdin a monté sa propre société depuis 1938, pour développer la première invention de Gallup à savoir des études de marché qu'il vend aux industriels.

Comme le fait remarquer Loïc Blondiaux, la période de Vichy n'est pas spécialement favorable au développement des méthodes de sondage d'opinion, fortement associées aux principes démocratiques et aux libertés individuelles depuis la rhétorique de Gallup. Le régime de Vichy préfère se faire une idée secrète de l'opinion publique à travers un dispositif d'écoutes à large échelle<sup>6</sup>. Peu d'établissements publics se sont intéressés aux sondages à cette période, à l'exception d'une institution comme l'Institut National d'Hygiène, créé par une loi du 30 novembre 1941, qui a procédé régulièrement à des enquêtes alimentaires et sanitaires en marge de ses travaux de laboratoire.

Du côté de la statistique officielle, l'introduction des premières enquêtes par sondage est à replacer dans la dynamique de la constitution du SNS. Le contrôleur général des armées René Carmille qui s'est distingué tout autant par ses faits d'arme en 1918 que par ses réformes comptables et ses conférences à Sciences-Po, est appelé dès août 1940 par le Ministre de la Guerre à reconstruire notre armée à l'aide d'un système « complet et cohérent » d'information démographique et économique dont les éléments clés sont le renseignement continu, l'identification des personnes, la tenue de fichiers mécanographiques, et la coordination de toutes les statistiques (Touchelay, 1993). Carmille crée d'abord en 1940 un service de Démographie, qui fusionne officiellement en 1941 avec la SGF dans un nouveau *Service National des Statistiques*. Celui-ci s'installe à Lyon mais s'adjoint bientôt de nombreux centres régionaux richement dotés en hommes et en machines. Le rattachement administratif de la petite et vieille SGF au SNS n'est pas immédiat dans les faits : si Michel Huber ancien directeur de la SGF et directeur de l'ISUP, est pris comme conseiller par Carmille dès les débuts, puis est rejoint en septembre 1941 par Alfred Sauvy, dont l'*Institut de Conjoncture* créé en 1938 au Ministère de l'Économie Nationale est également rattaché au SNS, il semble que le nouveau directeur Henri Bunle rechigne un peu à ce mariage forcé [entretien].

C'est dans ce climat de bouleversements incessants et de dualité entre administrateurs civils et militaires, entre zone occupée et zone libérée, que les techniques de sondage, *a priori* concurrentes du système de fichiers de Carmille, vont être timidement évaluées. On en trouve la trace dès 1941 dans des études de biométrie humaine impulsées par Pierre Lombardy et menées par les médecins militaires qui dirigent les ateliers régionaux (Touchelay, 1993),

---

6. Le service des « contrôles techniques » rattaché directement à Laval aurait intercepté 2,5 millions de lettres, 1,8 millions de télégrammes, et 21 000 conversations téléphoniques pour le seul mois de décembre 1943 (Blondiaux, 1998, p. 396).



et dans une note du 26 juin 1941 de Paoletti (délégué à l'équipement) à Dulac, administrateur du Service de la Démographie, qui discute des mérites des sondages de l'IFOP, de leurs usages privés (par les professionnels de la publicité<sup>7</sup>) et peut-être, pourquoi pas, publics. Cette question des sondages a du embarrasser René Carmille qui sollicite une commission d'experts composée de Stoetzel, de l'administrateur Sassi, et des professeurs François Divisia, Paul Lévy, Jacques Chapelon. Alors que Stoetzel développe les points techniques et émet la conclusion « *qu'il n'y a pas de différence essentielle entre un sondage dans une population et n'importe quelle recherche expérimentale* », Divisia conteste les hypothèses de normalité des distributions et d'indépendance des tirages qui sont au fondement des calculs et qui ne sont pas vérifiées dans un groupe social, *a fortiori* en matière d'opinions, « *à cause de cela toutes méthodes statistiques générales peuvent être d'une application d'autant plus dangereuse qu'elles se présentent, en la forme, comme plus perfectionnées* » conclut-il. Paul Lévy retient la première objection de Divisia et conclut que « *chaque application pratique du théorème de de Moivre relève du bon sens, et non pas d'une règle que l'on puisse appliquer automatiquement* ». Chapelon insiste quant à lui sur la dernière objection de Divisia à savoir qu'en matière de variable subjective, « *les réponses peuvent être orientées systématiquement par toutes sortes d'influences psychologiques* » tout particulièrement rajouta-t-il « *dans les circonstances actuelles* ». Il prolonge enfin cette inquiétude par une question bien iconoclaste : « *est-on sûr que le pourcentage p à mesurer existe bien ?* » car

*« le fait qu'il existe des opinions est dû à l'ignorance, mais si l'ignorance est trop grande, la notion d'opinion disparaît. Or les circonstances actuelles ont créé cette ignorance et ont entraîné une sorte de liquéfaction de l'opinion. Les gens pensent différemment selon les interlocuteurs, selon les heures de la journée. Ils n'ont plus d'opinion mais oscillent entre le scepticisme et la crédulité. Dans les temps normaux, quand l'opinion est cristallisée, elle est déjà évasive et difficile à saisir. Aujourd'hui autant vaudrait saisir un feu follet. Certes un sondage fournira un nombre, mais ce nombre ne représentera rien. Autant vaudrait s'efforcer de déterminer la température de la France ou la longueur d'une barre de caoutchouc étirée irrégulièrement. »*

Le ralliement de l'université à la méthode Gallup, que Blondiaux repère dès 1940 aux États-Unis, avec les publications favorables de Cantril et Katz et l'enrôlement de Lazarsfeld dans des recherches sur l'opinion, est visiblement plus difficile en France. Selon Pierre Thionet [entretien], les doutes et réserves émis par cette commission, dont l'actualité est d'ailleurs remarquable, n'ont fait que renforcer la défiance de Carmille ; il conclura dans son rapport au Ministre des Finances que « *les conditions théoriques nécessaires pour qu'un sondage devienne légitime sont réunies dans des situations extrêmement rares pratiquement* », et que dans les autres cas, on ne saurait attribuer aux sondages

7. Le décès de M. Bleustein-Blanchet en 1996 a été l'occasion de rappeler dans la presse le rôle qu'il a joué dans l'introduction des sondages en France.

par la méthode représentative la même valeur mathématique que celle que l'on peut attribuer aux sondages aléatoires. Enfin, rappelle-t-il, « *un sondage ne saurait être substitué systématiquement à une Statistique bien faite.* »

Cette défiance n'a pas empêché que Carmille accepte l'offre de services de Jean Stoetzel auprès d'une section des sondages créée début 1942 au sein du SNS et placée sous la direction de Pierre Delagrangé, centralien et administrateur. Mais « *les préventions de Carmille contre les sondages étaient nettes : la section des sondages n'a jamais été composée que de M. Delagrangé et sa secrétaire* » dit Thionet, critiquant fortement une brochure sur « les enquêtes par sondage » à destination des Administrations éditée par ce service qui reprenait les thèses de Stoetzel (voir Stoetzel, 1943). Le SNS a néanmoins mené sous l'autorité de Delagrangé et avec l'aide de Stoetzel et la logistique de ses trente-cinq directions régionales toute une série d'enquêtes par sondages, fussent-ils non aléatoires. La toute première, commandée par le Secrétariat d'État à la Production Industrielle et au Commerce, porte sur l'opinion des français sur les échanges de vieux textiles contre des tickets. Il y eut aussi l'enquête sur le logement (2 000 enquêtés) et sur le cinéma (18 960 enquêtés). Des fichiers les plus divers servent de base aux enquêtes sanitaires, comme les dossiers médicaux des chantiers de jeunesse (1941) ou les fichiers des administrations hospitalières de certaines grandes villes (1942), ou encore les fichiers des services agricoles des communes (1941), que le SNS exploite pour la première des enquêtes annuelles agricoles commandée par le ministère de l'Agriculture et poursuivie jusqu'en 1944 malgré la faible qualité des résultats. Une expertise s'est donc constituée au SNS, et c'est elle qu'entérine le décret du 30 mars 1944 sur « les recensements approximatifs par voie de sondage » : toute enquête par sondage faite par une administration publique ou un organisme privé d'intérêt général doit être soumise à son approbation. Comme le souligne Béatrice Touchelay dans sa thèse, « *le SNS bénéficie d'une législation très favorable. L'encadrement des personnes et des activités imposées par les Autorités de Vichy produit une abondance de données administratives et professionnelles qui est sans précédent dans l'histoire de la statistique publique.* » (p. 120).

#### 4. Explosion des sondages à la Libération

Dès l'automne 1944, l'IFOP, hébergée un peu plus tard par la CEGOS, reprend ses activités sous forme d'enquêtes publiées dans sa revue *Sondages* et dans différents journaux qui se piquent d'une mesure de l'opinion. L'IFOP subit néanmoins la concurrence du SSS sorti de la clandestinité. Animé par P. Copeau, P. Hervé, Y. Morandat et M. Barioux, celui-ci se dote d'un comité directeur composé de personnalités universitaires (G. Gurvitch, D. Lagache, H. Wallon, E. Labrousse, G. Lutfalla), dispose de 160 agents et correspondants et publie dans son mensuel, jusqu'aux années 1960, les résultats bruts à des questions d'opinion politique parfois instructives, parfois naïves et mal-

adroites<sup>8</sup>, choisies par son comité directeur ou par le gouvernement, et posées à 2 000 personnes environ.

Dans une France libérée qui doit reconstruire entièrement un consensus politique, l'opinion publique est à nouveau un levier important. Elle l'est aussi du point de vue de la géostratégie mondiale : l'IFOP reçoit le soutien des Fondations américaines qui agissent dans le cadre d'un volet recherche du Plan Marshall consacré au développement des sciences sociales et à la lutte contre le communisme. Stoetzel, devenu professeur de psychosociologie à Bordeaux et chargé de cours à l'IEP de Paris, diffuse sa méthode à travers de nombreuses conférences sur le sol national et même international puisqu'il intègre son institut dans un cadre mondial industriel – l'AIPO qui regroupe une quinzaine d'instituts – et académique, devenant en 1947 le premier président de la *World Association of Public Opinion Research* (WAPOR).

La *Fondation* d'Alexis Carrel est mise en extinction après août 1944, mais ses personnels sont incorporés dans le nouvel *Institut National d'Études Démographiques* créé par ordonnance du 24 octobre 1944. L'INED, nouvellement créé, grâce aux efforts de Robert Debré, Adolphe Landry, Henri Laugier et de son directeur Alfred Sauvy<sup>9</sup>, a l'initiative des premières enquêtes par sondage sur les dépenses des ménages. En effet Stoetzel et Girard s'y retrouvent à la tête d'une *section de psychosociologie* qui succède au service des sondages de l'ex-Fondation et dont la fonction principale est d'organiser des enquêtes sur des thèmes aussi divers que la mobilité sociale, l'opinion sur l'alcoolisme, les attitudes vis-à-vis de l'avortement et la contraception ou de l'immigration<sup>10</sup>, et enfin les budgets de famille. La plupart de ces enquêtes sont faites par sondage, mais l'approche, rappelée par Sauvy dans une préface de 1950 à la présentation des résultats, est celle

« de la méthode si contestée des quotas et des enquêtes à domicile. La question n'est pas de savoir si l'échantillon était ou non représentatif, ni si les réponses ont été sincères... La question est de savoir si les résultats de l'enquête permettront de prendre une décision plus opportune ».

---

8. 1945 : *Êtes-vous contents de la conduite des militaires américains en France ?*

1947 : Question 126 : *Si l'on vous demandait de choisir parmi les personnalités suivantes qui désigneriez vous comme Président de la République : Vincent Auriol ? Marcel Cachin ? Yvon Delbos ? Le général de Gaulle ? Francisque Gay ? Félix Gouin ? Edouard Hériot ?*

Question N°128 : *Quelles seraient à votre avis les principales causes responsables de l'ensemble des difficultés que traverse le pays à l'heure actuelle ?*

Question N°130 : *Aimeriez-vous que les villes de France possèdent des gratte-ciel comme en Amérique ?*

Question N°131 : *Quelle est à votre avis la plus belle région de France ?*

9. Sauvy qui était revenu à la Statistique Générale après un passage aux Cabinets de Charles Spinasse (1936) puis Paul Reynaud (octobre 1938), et au Comité des programmes et des achats alliés avec Jean Monnet, avait été quelques mois secrétaire général à la Famille et à la Population.

10. « Enquête auprès des immigrés italiens en France ...sur la possibilité de faire venir un nombre suffisant de travailleurs italiens par appel provenant d'immigrants déjà installés » ; les temps ont bien changé. Enquête « pour déterminer l'accueil qui pourrait être réservé à des mesures tendant à reculer l'âge de la retraite » ; celle-ci est redevenue d'actualité.

Il s'agit donc avant tout pour Sauvy de « guider l'action publique », et dans l'exemple qu'il choisit des budgets de famille, d'asseoir un indice du coût de la vie sur des bases moins contestables que celles des budgets-types d'avant-guerre :

« Le procédé d'échantillonnage même imparfait, même non représentatif, n'est pas seulement une commodité, mais parfois une nécessité. Aucun autre moyen ne permet par exemple d'approcher la connaissance des budgets familiaux d'une catégorie sociale déterminée. Les enquêtes effectuées depuis la guerre dans ce domaine ont rendu d'appréciables services, bien que les échantillons utilisés ne fussent pas représentatifs et que les réponses fussent affectées d'erreurs systématiques ».

Raison de plus pour trouver assez étrange que le texte suivant de la même publication soit un exposé technique (par Simone Lefebvre) de la méthode statistique et du calcul des erreurs d'échantillonnage dans les sondages aléatoires, sondages aléatoires que ne pratique pas l'INED. Alain Girard se montre plus clair dans un texte antérieur de *Population* (Girard, 1947a) en rappelant que « la loi des grands nombres ne s'applique que dans les limites du hasard ». Mais il lui semble suffisant « d'égayer les enquêteurs » et de les laisser « enquêter dans leur entourage ». « La confiance que l'enquêteur inspire autour de lui est le meilleur gage de la sincérité des réponses », d'où le recrutement de « correspondants » faiblement rémunérés mais motivés parmi « les instituteurs, ecclésiastiques, et militants syndicalistes ». L'entretien personnalisé avec une bonne dose d'empathie à la Rogers, et de récompense soigneusement calculée, est le modèle psychosociologique qui s'articule normalement avec un choix raisonné et qui s'oppose totalement à la neutralité froide du sondage aléatoire. Et c'est pourtant un calcul probabiliste qui permet à Girard d'évaluer à 2000 la taille d'un échantillon donnant une précision suffisante.

Si l'on regarde plus précisément les indications méthodologiques de l'enquête menée sur les dépenses de 1080 familles de la région parisienne en mars et juin 1946 (Girard, 1947b), il ressort du tableau des nombres de carnets de compte distribués par catégorie de famille (taille et profession) que les quotas ont été très souples. Dans l'enquête de novembre auprès de 936 familles de militaires (Girard, 1947c) les carnets de comptes ont été remplis par des volontaires. Et dans la synthèse des enquêtes sur les budgets familiaux de 1946, Girard admet que les 2910 budgets « n'ont pas été prélevés selon le principe de la méthode représentative ». Il reste que ces premières enquêtes, faites dans des conditions économiques difficiles, ont permis de préciser la variation des budgets selon les catégories de familles (lois d'Engel). L'INED tout en continuant à s'intéresser aux conditions de vie abandonnera assez vite l'observation des dépenses effectives. « Il avait ouvert une voie. Il appartient ensuite à un organisme spécifique, le CREDOC (créé en 1953) et à l'INSEE, avec des moyens d'une toute autre ampleur et dans une autre optique, de procéder à l'étude des consommations des ménages. » écrit Alain Girard beaucoup plus tard (1986).

## 5. Construction d'une expertise au SNS

Carmille est arrêté par la Gestapo en février 1944. Il mourra à Dachau en janvier 1945. Bunle assure tant bien que mal l'interim du SNS. Les Directions Régionales victimes directes ou indirectes des bombardements fonctionnent au ralenti. *La période troublée traversée suspendit à peu près toutes les enquêtes en 1944-45 et elles ne reprirent jamais sur le même rythme*<sup>11</sup>. Pour la statistique officielle, les années 1944 et 1945 ont été marquées par une très faible productivité et un recul de la pratique des sondages. Cependant, un cortège de demandes institutionnelles venant des ministères du Travail, de l'Intérieur, de la Production Industrielle provoque, à la Libération, une relance du service des sondages bientôt suivie d'une réorganisation complète du SNS. La loi de finance du 27 avril 1946 officialise la création de l'INSEE, placée sous la direction de Francis Closon, et sous la tutelle du Ministère de l'Économie et des Finances. Les décrets de juin en précisent les attributions. La section des sondages, après avoir dépendu du Cabinet, est rattachée à la même Direction que la Statistique Générale. Le non renouvellement de la mission de Stoetzel en janvier 1946 s'inscrit dans une cure générale d'amaigrissement des effectifs, mais peut être aussi dans la volonté de laisser le champ libre à l'équipe qui se forme autour de Pierre Thionet et Raymond Lévy-Bruhl fin 1946.

L'itinéraire intellectuel de Pierre Thionet est fortement marqué par les circonstances exceptionnelles de la guerre. À peine a-t-il terminé sa troisième année à l'École Normale Supérieure de la rue d'Ulm, qu'il est mobilisé en septembre 1939, puis réformé temporaire à la suite de problèmes de santé, des problèmes d'audition qui l'handicaperont toute sa vie. Affecté en classe de Navale de Saint Louis (repliée en province) jusqu'en août 1940 à l'arrivée des Allemands, puis à Henri IV l'année suivante, il n'a pas gardé un bon souvenir de ces deux années d'enseignement, qui se sont terminées par une demande de congé puis une quatrième année (1941-42) à l'École Normale, consacrée à la recherche. En l'absence de Darmois (en semi liberté à Londres), de Dugué à Alger, que Thionet remplace pour certains travaux pratiques, notre normalien passe son temps à dévorer la littérature statistique anglo-saxonne d'avant-guerre. C'est alors que le sous directeur de l'École le dirige vers la Statistique, et que Bunle le recrute à la SGF comme Administrateur de troisième classe. Il suit à l'ISUP, et peut-être à l'École d'Application qui a ouvert ses portes en octobre 1942 sous la direction d'Eugène Morice<sup>12</sup>, les cours d'économétrie de Roy, de probabilité de Fortet, de démographie de Bunle, de mécanographie de Viennot. Pierre Thionet se souvient être passé par toutes les sections du SNS, y compris à l'Institut de Conjoncture chez Alfred Sauvy, mais d'avoir eu peu de contacts directs avec Carmille hormis deux voyages à Lyon. Il a fait comme beaucoup de statisticiens, et tout particulièrement Bunle lui-même, un stage de quelques mois en Allemagne, au service statistique de la zone française d'occupation de Baden-Baden. Et c'est à son retour en France en novembre 1946 que commence vraiment sa carrière de spécialiste des sondages, par la

11. Thionet, 1947-48, p. 168.

12. Mais il n'est pas un des quatre élèves administrateurs de la première promotion, tous polytechniciens.

publication d'un ouvrage novateur sur les méthodes statistiques des administrations américaines.

Ce livre – (Thionet, 1946) – est une synthèse critique des expériences américaines en matière de recherche et de pratique sur les sondages, qu'elles soient empruntées au *Bureau of the Census* ou au *State College of Iowa* qui sont les principaux centres de recherche appliquée américains, ou qu'elles soient glanées au fil des écrits d'Hansen, Cochran, Hurwitz, Jessen, Marlow que Thionet a lus dans les collections du *JASA* ou dans la documentation très conséquente que lui a confiée un certain Mayer. Ce biologiste professeur au Collège de France se pique de statistique et a ramené de ses voyages aux États-Unis une documentation récente qui vient heureusement compléter les collections d'avant guerre interrompues. Mayer est un proche de Raymond Lévy-Bruhl qui a convaincu celui-ci, après des études de Droit et de Mathématiques, de s'orienter vers la statistique, et lui a obtenu une bourse Rockefeller d'étude aux États-Unis. Muni de cette bourse R. Lévy-Bruhl a été voir Closon, lequel l'a nommé administrateur au 1er septembre 1946 avant de l'envoyer avec différentes recommandations se former auprès des administrations américaines. Deux hommes incités par un même amateur éclairé à s'informer des méthodes américaines, allaient donc se construire parallèlement une expertise unique en France sur les techniques de sondage. Tandis que l'un, Thionet, peaufinait en Allemagne le manuscrit d'un ouvrage entièrement fondé sur une étude livresque des techniques d'échantillonnage, l'autre entreprenait un voyage de neuf mois, six mois de cours intensifs de statistique mathématique à Columbia et trois mois à Washington, partagés entre le *Bureau of Labor Statistics* où il s'initie à la statistique économique et sociale et l'équipe de Hansen et Hurwitz qui travaillent aux méthodes d'enquête du *Census* (et où il rencontre Deming).

Il est intéressant de noter que l'Amérique est alors non seulement le pays des libérateurs de la patrie occupée et des financeurs du Plan Marshall, mais aussi l'Eldorado de tous les chercheurs français de l'immédiat après guerre. Les économistes (dont le futur Directeur Général de l'INSEE Edmond Malinvaud) y vont découvrir l'économétrie, la théorie des jeux ou encore la recherche opérationnelle, développée par ceux-là même qui ont fui le nazisme et l'Europe dans les années trente. Alain Drouard (1982) fait de la référence aux États-Unis une des bases de l'essor des sciences sociales en France et donne une liste des sociologues et psychologues français – parmi lesquels Bourricaud, Crozier, Mendras, Touraine – qui vont aux États-Unis avec une bourse Ford ou Rockefeller et la conscience d'un important retard à combler. Pour la Statistique, « *l'exemple des États-Unis est instructif* » comme le dit Thionet, parce que, malgré une organisation éclatée entre le Bureau of the Budget et les Federal Agencies des Ministères, entre le Census chargé du recensement, de la production et du commerce, et le BLS chargé des prix et du travail, on s'y préoccupe vraiment de centralisation ou de coordination des statistiques depuis au moins 1922; parce que l'esprit d'équipe et l'organisation des tâches frappent tous les visiteurs (Mossé, 1945); parce que la recherche théorique en statistique mathématique fait preuve d'une avance considérable sur ce que l'on peut lire et entendre en France dans les années quarante – il faut se souvenir

que Ronald Fisher commence à peine à être connu de quelques spécialistes comme Fréchet et Darmois à la fin des années 30 -, et qu'elle est de plus intimement associée aux pratiques administratives. Ce n'est pas le cas en France avec la méfiance des milieux savants vis-à-vis des probabilités et de la statistique et la faible formation mathématique des statisticiens de la SGF et des militaires du SNS plus orientés vers la gestion de grands fichiers :

« Dans cette voie on peut prévoir la disparition prochaine de tout emploi du Calcul des Probabilités et même de toute science statistique, à laquelle se substituerait une vaste comptabilité nationale, et avouer la vanité de chercher des nombres approchés puisqu'on disposerait ainsi de nombres exacts. Toutefois l'expérience montre qu'actuellement au moins, les principaux éléments d'une telle Comptabilité Nationale seraient, en France, profondément entachés d'inexactitudes... ».

On voit comment le pessimisme exagéré de notre auteur sur le développement futur des sondages se tempère d'un autre pessimisme sur la qualité de la Comptabilité Nationale, fort salutaire à son art !

Que trouve-t-on au juste dans ce premier ouvrage de Thionet ? Essentiellement une théorie de l'échantillonnage assez élémentaire donnant les précisions des échantillons aléatoires simples de Bernoulli et Poisson, des échantillons stratifiés (Neyman et Markoff), des échantillons à deux degrés (Friedmann et Wilcox), des échantillons systématiques et en grappes (ce qui mobilise la notion fishérienne de corrélation intraclasse), et donnant aussi les calculs et arbitrages en terme de coûts introduits par Neyman et Jessen. Ces thèmes forment la base de la théorie mathématique des sondages que Thionet développera à fond pendant une dizaine d'années, dans près d'une centaine d'articles, dans ses cours à l'École d'Application (Thionet et Lévy- Bruhl, 1947-48), dans les fameux numéros 5 et 6 des « Études théoriques » (Thionet, 1953), ou encore dans son dernier ouvrage chez un éditeur allemand (Thionet, 1978). La seconde partie de l'ouvrage de 1946 est plus originale à nos yeux d'historien car elle relate trois types d'expériences américaines postérieures à 1938 qui ont servi d'études de cas aux pionniers français : 1°) le sondage au vingtième fait, selon une technique de choix systématique, sur certaines questions particulières, en complément du recensement général de la population de 1940. 2°) le sondage aréolaire stratifié de Hansen sur le chômage, celui de Eckler sur les migrations internes. 3°) diverses techniques d'échantillonnages agricoles : échantillonnage par quadrillage de Jessen dans l'Iowa, estimation des récoltes, utilisation d'un échantillon maître.

Lorsque Lévy-Bruhl rentre des États-Unis en septembre 1947, il est aussitôt affecté aux sondages, et même s'il consacre une première année à suivre les cours de l'École d'Application que lui indique son Directeur, il forme avec Thionet un tandem qui va se trouver maître de la conception des enquêtes par sondage au sein de la Direction de Rivet, tandis que la partie opérationnelle (liaison avec les Directions Régionales et les enquêteurs) allait très vite leur échapper au profit de la Direction de l'exploitation (Chevry, assisté de Brichler et Desabie après 1949). Tous deux vont pouvoir passer du transfert de

technologie à un programme d'enquêtes par sondage. Celles-ci sont devenues entre-temps indispensables dans le cadre des nouvelles missions de l'INSEE : lorsqu'à la fin de la guerre Michel Huber (1944) dresse dans le JSSP, à la demande du Conseil Supérieur de la Statistique reconstitué, la liste fort longue des « principales lacunes et insuffisances des statistiques françaises », il note en effet la faiblesse du système des enquêtes, et réclame

« qu'une grande enquête sur les budgets familiaux soit entreprise en France dès que les circonstances le permettront. Elle seule peut fournir des bases pour un calcul convenable d'indices du coût de la vie... À côté d'enquêtes portant sur les familles ouvrières, les moins favorisées, il est nécessaire d'avoir des données sur les familles des classes moyennes, fonctionnaires, employés, sur des familles de cultivateurs... ».

Thionet explique (1947-48, p.168) que les activités de son service explosent, du fait qu'il est chargé, dans le cadre de la mission de coordination de l'INSEE, de la vérification des questionnaires et des plans de sondage des différents Ministères de la Démographie, de l'Agriculture, de l'Industrie et du Commerce, du Travail, des Finances et des Colonies. Et il se plaint de

« l'extrême faiblesse de nos moyens de calcul numérique (qui) nous a amené à rechercher souvent des méthodes mécanographiques pour effectuer les calculs statistiques qui nous sont posés. L'INSEE ne possède même plus (en décembre 1947) de calculatrice à cartes perforées ».

## 6. Premières enquêtes par sondage aléatoire à l'INSEE

La première enquête par sondage sur les budgets de famille, a été réalisée à l'INSEE en avril et novembre 1946 par R. Rémary, à la demande de la Commission supérieure du coût de la vie, dans le double objectif de fournir « les coefficients de pondération des indices » et « une appréciation de la valeur quantitative et qualitative du régime alimentaire de diverses catégories de population ». Les questionnaires en forme de carnets de compte ont été confiés aux familles par des enquêteurs qui sont encore en grande partie extérieurs à l'Institut. Les (1052 + 1077) familles interrogées, limitées au département de la Seine et à un niveau social « inférieur ou égal à celui d'ouvrier qualifié et d'employé subalterne », sont choisies « d'une part à l'aide d'un tirage au hasard dans le fichier de la Caisse interdépartementale de la Sécurité Sociale, d'autre part parmi les connaissances des enquêteurs ». On peut donc parler comme le fait Jacques Desabie de sondage « partiellement aléatoire », et d'un échantillon qui n'est guère représentatif du champ très limité de l'enquête, encore moins de l'ensemble des ménages français.

Les premières enquêtes par sondage de Thionet<sup>13</sup> et Lévy-Bruhl réalisées en 1948 sont des enquêtes de marché rendues nécessaires par la suppression partielle du rationnement après 7 années de restrictions, et portant l'une sur

13. Pierre Thionet (1954) date de la fin 1947 le moment à partir duquel il a eu la responsabilité des enquêtes par sondage à l'INSEE.



l'achat de textiles (rationnés et non rationnés) (Thionet et Lévy-Bruhl, 1948) à la demande du Ministère de l'Industrie, l'autre sur la consommation de produits alimentaires non contingentés (Thionet et Lévy-Bruhl, 1949) à la demande du Haut Commissariat au Ravitaillement. Le questionnaire sur les textiles (février 1948) concernait les achats du ménage mais était rempli par la mère de famille. Les ménages étaient pour la première fois choisis au hasard avec probabilités égales, à partir des feuilles du dénombrement de 1946, dans toute la France et toutes les couches de la population. Plus précisément, afin de « réduire les frais d'enquête », c'est un tirage à deux degrés que l'on a inauguré : d'abord des communes tirées au hasard dans 7 strates définies par la taille et la proportion de population agricole, puis des ménages tirés par grappes de trois et de façon systématique (à intervalle régulier) dans les feuilles de ménage de ces communes. La taille de l'échantillon quoique limitée par des raisons financières était tout de même de 4 086 ménages. Une campagne de presse et de radio exposant les objectifs de l'enquête avant le sondage avait permis d'obtenir un très faible taux de refus, inférieur à 5 %. L'enquête textile permit de constater « *qu'un rationnement par les prix avait succédé à un rationnement par les tickets* ». L'enquête sur le ravitaillement de décembre 1948 portant sur 4237 ménages a suivi un plan de sondage analogue légèrement amélioré et complexifié : une première stratification en 18 régions précédait la stratification par taille et type de commune. Dans chaque strate on a choisi un nombre de ménages à interroger proportionnel au nombre de ménages de la strate et non pas au nombre d'habitants comme en février. Des adresses supplémentaires ont été prévues pour remplacer les ménages qui seraient absents ou refuseraient l'enquête. Toutefois, la nécessité de prendre en compte les nouveaux logements construits, bien qu'entrevue, a été repoussée aux prochaines enquêtes. Dans les deux enquêtes, les enquêteurs sont des agents des Directions Régionales de l'INSEE, sauf à Paris où ils ont été complétés par 15 étudiants. Ils sont l'objet d'une formation orale initiale « sur les principes de la théorie des sondages », et munis d'instructions écrites spécifiques à l'enquête, ce qui n'empêche pas qu'ils subissent des contrôles *a posteriori* par visites ou par téléphone. Si l'on sait que le questionnaire lui-même a été minutieusement testé et même complètement remanié après une première tentative de questionnement par correspondance jugée désastreuse (sur des enquêtés volontaires choisis par la méthode des quotas), on est en droit de considérer que ces deux enquêtes ont été les prototypes soigneusement calibrés d'un modèle de sondage aléatoire que Thionet et Lévy-Bruhl ne feront par la suite qu'améliorer, défendre, et généraliser.

## 7. Le modèle Thionet et Lévy-Bruhl (1948-54)

Les enquêtes INSEE-INH de 1948-51 sur les budgets de familles constituent un nouveau terrain d'expérimentation de la méthode des sondages. La première de ces enquêtes par exemple (Thionet et Lévy-Bruhl, 1949b) date du printemps 1948 ; elle est motivée par quatre objectifs clairement énoncés : étudier les effets de la suppression progressive du rationnement sur les dépenses alimentaires, suivre l'évolution de la répartition des dépenses par poste, évaluer

la répercussion des hausses de prix et/ou de salaires sur les budgets familiaux, réajuster la pondération des indices du coût de la vie. Les dépenses alimentaires ou diverses hebdomadaires (désignation, quantité, coût, menus), et les dépenses sur plus longue durée d'habillement et de logement sont recueillies par un questionnaire rempli par l'enquêteur lors d'un entretien minutieux au cours duquel il note également l'état du logement. L'échantillon pour la première fois dans ce domaine est totalement aléatoire, et tiré dans la seule base des logements du recensement de 1946, mais encore restreint en 1948 à la région parisienne, et aux familles de 0 à 3 enfants dont le chef était salarié ou sans emploi, avec un salaire inférieur à un certain seuil et une situation inférieure à celle de contremaître ou de comptable, ces restrictions étant imposées par une taille économique de 1000 ménages ne permettant pas de reproduire la diversité des genres de vie. Malgré le recrutement massif de 60 enquêteurs « parmi les étudiants en sociologie, les assistantes sociales, et le personnel de l'Institut », nécessité davantage par la lourdeur du procédé (3 visites) que par la taille de l'échantillon, malgré la récompense en nature offerte aux familles, le taux de refus a atteint la valeur inquiétante de 40 %. Une étude assez systématique des facteurs de ces refus, c'est-à-dire de sa variation dans les différents types de familles n'a pas suffi pour se consoler de cet échec. Pour Thionet (1954) cette « mauvaise grâce » des enquêtés jette un doute sur la sincérité des réponses obtenues, et c'est une cause essentielle de l'abandon pour quelques temps de la méthode aléatoire dans les enquêtes de budget au début des années 1950 au profit de la méthode des quotas.

Thionet perd d'ailleurs la responsabilité de ces enquêtes à partir de l'arrivée de Jacques Desabie, au profit de l'enquête emploi (Affichard, 1987) dont la première version semestrielle, date de 1950. Elle est présentée par Lévy-Bruhl (1951a et b) qui en a importé le concept de son voyage américain, et plus précisément de l'enquête *mensuelle* par sondage auprès de 25 000 ménages du *Census*. L'enquête emploi devient le nouveau cadre dans lequel peuvent s'appliquer les sondages aléatoires. Comme pour les autres enquêtes de cette période, les rôles sont assez bien partagés entre Pierre Thionet, qui construit le plan de sondage, rédige des consignes aux enquêteurs et fait tous les calculs de précision se rattachant aux estimations, et Raymond Lévy-Bruhl qui prend en charge la difficile mise au point du questionnaire en liaison avec les choix de contenu. Ce dernier rappelle les insuffisances des méthodes d'investigations précédentes : les enquêtes trimestrielles du Ministère du Travail ne s'adressent qu'aux employeurs des entreprises de plus de 10 salariés et ne couvrent pas toutes les activités ; les fichiers administratifs de la Sécurité Sociale ne sont pas disponibles et ceux des fonds de chômage sont incomplets compte tenu des conditions restrictives qui s'attachent aux indemnités (par exemple avoir travaillé pendant au moins six mois) ; le recensement quinquennal enfin est « lourd, onéreux et long à dépouiller ». La nouvelle enquête par sondage qui est proposée doit, sur le modèle américain, répondre au contraire à des exigences d'exhaustivité, de faible coût et de disponibilité. Le plan de sondage à deux degrés est en gros celui qui a été testé dans les enquêtes de marché et de budget en 1948 et 1949 : stratification en 8 grandes régions et en 10 catégories de taille et de ruralité pour le tirage au hasard d'une commune

avec probabilité proportionnelle au nombre de ses ménages (à l'aide de la table de Tippett), puis tirage systématique des ménages dans les feuilles de la DR correspondant à la commune, en sur-représentant les communes urbaines par un taux de sondage plus élevé pour tenir compte de l'hétérogénéité plus grande des comportements, et en complétant pour la première fois les logements du recensement par un échantillon de logement neufs. Tous les individus de plus de 14 ans du ménage sont interrogés. Comme précédemment, les 149 enquêteurs INSEE épaulés par 11 étudiants ont été solidement préparés et contrôlés, et les ménages enquêtés prévenus individuellement par courrier, tandis que la tentative de publicité par presse et radio était abandonnée dès la seconde enquête. Le ménage occupant au moment de l'enquête le logement sélectionné est impérativement celui qui doit être interrogé, fut-ce au prix de 2, 3, 4 visites ou plus, ce qui a permis, en 1950, de baisser le taux de non réponses (pour absence ou refus) à 13 %. Le redressement ne se fait pas par pondération, mais par simple duplication de cartes mécanographiques tirées au hasard dans la strate déficitaire. Enfin, si l'échantillon (taille 5 700) est totalement renouvelé entre les deux enquêtes de 1950, il est ensuite décidé de n'opérer que son renouvellement partiel. Certaines questions sont légèrement modifiées pour pouvoir étudier expérimentalement les impacts de la formulation. Ces mesures, ainsi que la généralisation du plan de sondage à de nombreuses autres enquêtes au début des années 1950, marquent la fin d'un période de mise au point, et inaugurent une sorte de standardisation de la procédure.

Malgré les difficultés rencontrées dans l'enquête-budget, Thionet a en effet réussi à construire un plan original d'échantillonnage aléatoire à deux degrés (communes tirées au hasard, ménages tirés systématiquement), avec stratification des communes par régions et catégories, qui est commun à plusieurs enquêtes : enquête sur l'emploi, enquête sur l'écoute téléphonique (1952), mais aussi enquête sur le marché de la chaussure (1953) et sur celui des assurances (1953). Laissons Thionet (1954) expliciter lui-même ce que ce modèle a d'original :

« Cette méthode est mathématique, c'est-à-dire se prête à l'application correcte du calcul des probabilités. Mais alors que la méthode du Bureau of Census consiste à tirer au sort des petites aires, la méthode de l'INSEE, comme d'ailleurs celle du Social Survey britannique, repose sur des tirages au sort d'adresses dans les fichiers ou sur des listes préétablies. La différence essentielle entre le procédé de l'INSEE et celui du Social Survey est que les administrations britanniques sont en mesure de fournir au Social Survey des listes d'adresses adaptées au but visé ; tandis qu'en France, après avoir essayé (Enquête emploi 2<sup>ème</sup> semestre 1951 et Enquête Loyers 1952) le tirage au sort dans les calepins de l'administration des Contributions Directes, l'INSEE a préféré s'en tenir à l'emploi de sa documentation personnelle : les feuilles de ménage du recensement 1946 complétées par un échantillon de nouveaux logements (grâce à l'institution du permis de construire). »

La standardisation de la méthode trouve son point d'orgue dans la notion d'*échantillon-maître* introduite au moment du recensement de 1954. Un tel échantillon-maître constitue une sorte d'investissement de forme (au sens de Thévenot) qui décuple l'efficacité de la méthode en définissant une fois pour toutes la liste des cantons qui serviront de première étape à toutes les enquêtes de 1954 à 1960; on ne renouvelle que les tirages des communes et des ménages. La première enquête INSEE-CREDOC de 1956 sur les budgets familiaux (CREDOC, 1960) profitera de cette percée de l'enquête emploi et de la nouvelle base que constitue le dernier recensement, pour réintroduire en force le sondage aléatoire dans ce domaine où il avait été abandonné au profit de la méthode des quotas. L'enquête saisit pour la première fois toutes les consommations des ménages de toutes les catégories sur tout le territoire. Elle vise en effet à mieux connaître les agrégats de la Comptabilité Nationale, à permettre l'estimation économétrique des élasticités prix et revenus, et à fournir un cadre à certaines études de marchés. 22 800 ménages seront choisis par un tirage à trois degrés – cantons, communes, logements –, et seront invités par les 300 enquêteurs à remplir un carnet de compte assez complexe à plusieurs périodes. Enquête lourde se traduisant par trois visites et 45 types de cartes différentes possibles pour un même ménage, l'enquête de 1956 n'a pas rempli toutes les attentes en matière de précision des estimations, souvent pour des raisons indépendantes du plan de sondage, et le fait que le total des dépenses ait dépassé le total des ressources déclarées a été vécu comme un échec relatif. Mais le modèle de sondage était déposé et ne devait guère être modifié jusqu'au recensement suivant.

Toutefois, la conception de cette enquête INSEE-CREDOC, bien qu'inspirée du travail de Thionet et Lévy-Bruhl, n'était déjà plus de leur responsabilité. En effet Pierre Thionet quitte le service des sondages de L'INSEE à sa demande et pour des raisons personnelles dès 1954. Il est recueilli dans un des bureaux du Service des Études Économiques et Financières monté par Gruson, où il semble qu'en dehors de très nombreuses notes de lectures, il ait essentiellement travaillé, avec Nataf et Darmois, à sa thèse sur les notions de précision et d'information en liaison avec les sondages, thèse qu'il soutiendra en 1958. La suite de sa carrière sera dès lors universitaire, à Poitiers d'abord depuis novembre 1960, à Paris-Dauphine après 1968 et jusqu'à sa retraite, une carrière consacrée à l'enseignement de la statistique et à une recherche sur la théorie de sondages poursuivie avec ferveur, malgré la coupure totale et définitive d'avec les travaux appliqués de l'INSEE. Quant à Raymond Lévy-Bruhl, il devait lui aussi quitter le service des sondages pour aller au Ministère du Travail prendre la responsabilité de l'enquête sur les salaires et les conditions d'emploi (Lévy-Bruhl, 1977). Dès lors c'était sur Jacques Desabie que devait reposer la théorie et la pratique des sondages (Desabie, 1965), et plus particulièrement des enquêtes sur les conditions de vie des ménages (Desabie, 1987).

Mais l'on ne peut pas en terminer ainsi avec l'œuvre du tandem Thionet-Lévy-Bruhl sans s'élever un peu au dessus de leur innovation et évaluer l'extension et la réception de leurs travaux au delà des réalisations concrètes que sont les enquêtes INSEE de la période 1948-54. La première de ces extensions,

c'est l'enseignement. Thionet enseigne les sondages à l'école d'application de l'INSEE, dès 1943, aux administrateurs et aux attachés, mais ces premières années sont interrompues plusieurs fois par les problèmes de santé ou le voyage en Allemagne. Le cours publié de 1947-48 de *Théorie et pratique des sondages* est cosigné avec Lévy-Bruhl qui a rédigé les deux dernières leçons plus descriptives sur «l'enquête verbale» et sur «les sondages aux États-Unis». Truffée d'exemples et de contre-exemples empruntés à la France et aux États-Unis, c'est la première présentation didactique de la théorie des sondages que l'on peut trouver en France<sup>14</sup>. Thionet assurera même un moment la direction par intérim de l'École. Les deux compères se retrouvent encore associés avec Hansen et Yates dans une publication d'un éphémère *Centre européen d'application statistique agricole et démographique* créé par Darmois sous les auspices de l'ONU, de l'UNESCO de la FAO et du gouvernement français, et chargés d'une formation de leurs personnels aux sondages en agriculture. Il faut noter à ce sujet que ce domaine de l'agriculture est plutôt celui des américains que celui des français, et Lévy-Bruhl, qui se souvient avoir enseigné la pesée des betteraves à l'École d'Application, a toujours souligné le peu du moyen du ministère de l'Agriculture en cette période<sup>15</sup>. Darmois est également à l'origine de l'entreprise de traduction de l'ouvrage de Yates (1949) à laquelle participeront Thionet et Lévy-Bruhl.

Comment évaluer ces travaux pionniers de Thionet et Lévy-Bruhl ? D'un point de vue d'historien, et non d'expert technique, la question n'est pas de le faire à l'aune des améliorations ou même des révolutions qui seront opérées à la suite de leurs travaux, mais plutôt d'évaluer grossièrement les choix dans lesquels ils ont engagé la statistique officielle, en s'appuyant sur les matériaux de l'époque qui témoignent des conditions, des risques pris, des succès et des controverses engagés.

## 8. Évaluation : les profits et pertes du hasard

Le tout premier de ces choix semblerait être celui du sondage contre celui des grands fichiers du SNS puis de la comptabilité nationale. Mais on a vu que ce choix est dans l'air du temps pour un grand nombre d'investigations privées ou publiques, et qu'il n'est pas incompatible avec les deux autres pour la statistique publique, puisque le fait même de choisir l'échantillonnage au hasard oblige à s'appuyer sur une base de sondage exhaustive et fiable.

Le choix majeur est celui du hasard contre la méthode des choix raisonnés. Thionet en fait remonter la responsabilité au SNS de 1942 et à Carmille, mais il semble bien que les choses n'ont pas été tranchées avant le départ de Sauvy et Stoetzel pour l'INED. La méthode du sondage aléatoire est alors

14. Son cours sera repris par M. Chartier après 1954.

15. On trouve néanmoins une participation de Thionet à l'enquête annuelle sur les salariés de l'agriculture, réalisée par sondage, à partir du recensement et des fichiers des Caisses d'Assurances sociales, et sur le modèle des enquêtes faites au ministère du Travail. Le document de synthèse (Thionet 1954b) contient, sous forme d'extraits de l'instruction aux enquêteurs, un luxe de détails sur l'élaboration des taux de sondage.

fermement et régulièrement affirmée par Thionet et Lévy-Bruhl contre la « méthode Gallup » du sondage par quotas défendue par Sauvy, Stoetzel et Girard. Si ce dernier dit (1986) que « *la querelle est aujourd'hui dépassée* », il reconnaît qu'elle fut vive aussi bien à l'intérieur même de l'INED, où ne manquaient pas les mathématiciens ouverts à la statistique mathématique, qu'entre l'INED et l'INSEE, dont la création même et les attributions ont constitués une grosse épine dans le projet scientifique d'une équipe formée dans la matrice de la Fondation Carrel et marquée par l'idée d'une science sociale unifiée. Malgré des textes statutaires qui invitaient à la coopération et au partage des moyens, la situation de quasi monopole de l'INSEE a plusieurs fois été amèrement critiquée par l'INED, que ce soit en terme de moyens financiers et mécaniques, ou en termes d'attributions légales, comme l'obligation faite aux interrogés de répondre et le secret statistique qui lui font contrepartie, ou encore le monopole du visa nécessaire à toute enquête sous couvert du comité de coordination (Girard 1986, p.76). Le choix du sondage aléatoire ou du choix raisonné est bien en effet une question économique avant que d'être une question scientifique : c'est par l'argument de l'économie à précision donnée (ou vice-versa de la précision pour un coût donné) que les statisticiens ont défendu le premier de ces choix, mais c'est aussi parce qu'il nécessite une base de sondage et des enquêteurs en plus grand nombre, plus dispersés sur le territoire et mieux formés, que les psychosociologues de l'INED ont dû faire le second choix. La question de coût est centrale comme le révèle les attaques continuelles de Stoetzel ou Sauvy sur ce thème<sup>16</sup>. De leur côté, Thionet et Lévy-Bruhl ne manquent pas une occasion de critiquer la faible valeur scientifique de « la méthode Gallup ». Dès 1945, Thionet s'est taillé une réputation de spécialiste du fleuret, dans une belle controverse avec Galvani, au sujet des mérites de l'école italienne et de son leader Corrado Gini, dépeint comme « intransigeant », « éternel batailleur contre les théories anglo-saxonnes », ayant « un sens très vif de la réclame », et témoignant « un profond mépris pour ses grands rivaux ». Il n'est pas sûr qu'il ait pensé plus de bien de Stoetzel. Tous deux sont en tout cas les représentants d'une statistique sans mathématique et sans probabilité qu'il condamne<sup>17</sup>. « Seule la méthode aléatoire est *mathématique* et se prête à l'application correcte du calcul des

16. Par exemple dans la discussion de l'exposé de Lévy-Bruhl à la Société de Statistique de Paris [Lévy-Bruhl 1951] : « - M. Rivet rend hommage à ses collaborateurs...

- M. Stoetzel s'associe à M. Rivet pour féliciter le conférencier et lui demander approximativement le coût d'une telle enquête

- M. Lévy-Bruhl les évalue approximativement à 500 000 francs pour les frais de déplacement et à 2 000 000 pour les heures de travail des enquêteurs (...) M. Thionet précise qu'à son avis le travail des administrateurs représente en plus une dépense de 500 000 francs environ

- M. Stoetzel émet le voeu que lorsqu'une étude statistique est présentée, il soit précisé quel en a été le coût." (il ne dira plus rien d'autre dans la discussion)

17. « *Sa devise, au lieu d'être « Statistique avec mathématiques » devrait être « Statistique avec le minimum de mathématiques »* (J.R.S.S. 1926, p. 706-7). « *Nous ne saurions partager pareille conception car, si elle signifie, comme c'est bien en effet le cas, que la moindre formule de Calcul des Probabilités est une faute, alors que Gini se permettra des formules d'algèbre très élémentaire effroyablement compliquées et inutiles, on peut véritablement parler alors d'esprit de système ; mais cette extrême méfiance de notre auteur*

probabilités» dit-il à plusieurs reprises (1947, 1954). Lévy-Bruhl écrit tout un article dans le JSSP (Lévy-Bruhl, 1951c) pour brocarder le dernier échec de Gallup : si on connaît son coup d'éclat de 1936, on se souvient en effet moins bien aujourd'hui de sa prévision erronée du succès de Dewey aux élections présidentielles de 1948, et du démenti que lui a infligé le succès, il est vrai très mince, de Truman.

« Gallup et ses collègues devinrent l'objet de plaisanterie non seulement de la presse mais de l'ensemble du public. C'est ainsi, par exemple, que le jour où Truman prenait officiellement ses fonctions de Président à Washington, au Sénat de l'État d'Indiana, tous les sénateurs unanimes, debout, observaient une minute de silence à la mémoire du Dr Gallup ».

Au-delà de cette arme de la dérision qui est de bonne guerre, l'auteur analyse les caractéristiques de la méthode, qui, sauf à envisager une modification toujours possible de l'opinion elle-même, indiquent un défaut de l'appareil de mesure :

« Il apparaît que la technique utilisée par les Instituts d'Opinion publique peut introduire des erreurs systématiques non mesurables. Il semble donc que la méthode utilisée puisse être tenue pour responsable, pour une part, de l'erreur totale ».

La possibilité d'un *calcul* de l'erreur est bien sûr le grand avantage de la méthode aléatoire sur la méthode des quotas, laquelle ne peut que *réduire* l'erreur par des choix raisonnés sans connaître le degré final de l'approximation. Certes il s'agit seulement de l'erreur d'échantillonnage et Thionet ne manque pas de rappeler qu'il faut y ajouter les erreurs propres à toute investigation empirique, que l'on pourrait ventiler en erreurs sur l'objet, erreurs liées à l'appareil de mesure, à l'observateur...

Levy-Bruhl (1949) liste une douzaine de ces causes d'erreurs<sup>18</sup> dans le cas des enquêtes par questionnaire. L'analyse de l'erreur totale associée à un sondage est très délicate encore aujourd'hui. Pierre Thionet a développé le calcul d'erreur d'un point de vue théorique dans son cours de 1947-48. En pratique les choses sont plus complexes car tout n'est pas calculable. Mais l'effort dont il fait preuve pour adapter ces calculs aux cas de sondages les plus complexes est remarquable. Même le passage répété des enquêteurs est exploité pour calculer le biais dû aux non-réponses. Ce qui l'est plus encore, c'est sa volonté de publier de tels calculs d'erreur pour chacune des enquêtes, ce qu'il fait explicitement dans l'enquête emploi dès 1951 : on trouve dans le BMS une annexe à la présentation de l'enquête qui est consacrée à l'estimation et la discussion des précisions des résultats, exprimés en terme de coefficients

---

à l'égard de toute hypothèse supplémentaire gratuite, de tout schéma (surtout probabiliste), caractérise véritablement ses travaux et son école. » (JSSP, 1945, p.250)

18. Indétermination de l'univers à enquêter, indétermination de l'information désirée, fluctuation physique, erreurs de rédaction du questionnaire, erreurs dans les réponses, biais enquêteur, interaction enquêteur-enquêté, champ de l'enquête mal couvert, influence de l'organisme enquêteur, mauvaise rédaction des instructions, mauvais report dans les tableaux, biais d'interprétation des résultats.

de variation du nombre total des heures travaillées (variable selon les régions de 2 % à 15 %), ou du nombre des chômeurs (10,6 %). Et la synthèse présentée en 1954 à la SSP traite pour l'essentiel de l'évaluation des erreurs d'échantillonnage dans les enquêtes à deux degrés de sondage.

« L'idéal serait de pouvoir publier côte-à-côte les résultats de l'enquête et des calculs d'erreur. Malheureusement ceci retarderait notablement la publication des résultats; et il a fallu y renoncer une fois pour toutes à l'I.N.S.E.E. Les calculs d'erreur sont donc effectués dans les temps morts, tant pour la partie mécanographique que pour la partie manuelle du travail.(...) Pour terminer nous formulons le vœu que des recherches analogues puissent être effectuées à propos des nouvelles enquêtes de l'I.N.S.E.E.; nous souhaitons qu'on se rende compte de l'absolue nécessité de ce genre de spéculation (pour une administration qui entend employer des méthodes scientifiques) et que par conséquent, on sache y consacrer les moyens nécessaires » (Thionet 1954, p. 184 et 195).

Or il semble bien que ces vœux pieux mais raisonnables n'aient guère été entendus; que l'on ait considéré que toute cette activité calculatoire de Thionet était une sorte de manie de mathématicien, une sorte de luxe qui ne portait pas à conséquence, et que le renoncement indiqué par lui ait perduré jusqu'à inclure toute publication, même différée, d'un quelconque calcul d'erreur. Ce n'est pas le moindre des paradoxes que de voir un choix justifié par la principale raison qu'il est le seul à autoriser un calcul d'erreur, se prolonger par toutes sortes de raisons sauf celle-là. La récente crise de confiance envers les sondages politiques a remis à l'ordre du jour cette nécessité d'assortir les estimations d'une mesure d'erreur, et quelques timides initiatives font écho au souci exprimé 50 ans plus tôt par Thionet. Mais la lourdeur des calculs n'est plus la raison principale des renoncements qui perdurent.

## 9. Évaluation : du choix de l'unité statistique

Cette première facette de la controverse en cache une autre cependant qui est peut-être plus importante qu'il n'y paraît. C'est celle qui porte sur le choix de l'unité statistique privilégiée. Il est clair en effet que les chercheurs de l'INED se penchent sur toute une série de problèmes très variés, de faits, d'opinions, de goûts, de vœux, d'orientation professionnelle, de santé et d'hygiène, de natalité, d'immigration, pour lesquels l'unité signifiante est l'individu, que ce soit au niveau des dispositions, des opinions, ou des décisions et des actions, même si leurs sentiments natalistes et leurs professions de psychosociologues les conduisent à situer ces événements dans le cadre de la famille ou du groupe social. Les préoccupations et attributions de l'INSEE lui font au contraire privilégier *a priori* les acteurs économiques, au premier rang desquels les ménages. Que cette logique institutionnelle soit redoublée par des raisons pratiques et historiques ressort avec évidence des travaux de Thionet. C'est même un point sur lequel il s'éloigne du modèle américain, choisissant,



plutôt que le sondage aréolaire du Census, un tirage au hasard d'adresses de logements, pour la seule raison que c'est la forme pratique de l'unité du recensement et qu'ayant essayé en vain, dit-il, de travailler avec des fichiers d'autres administrations, c'est pour finir envers la sienne qu'il s'est tourné avec le plus de confiance. Il ne reste alors plus qu'à définir le ménage comme l'ensemble des personnes partageant le même logement (à la différence de la famille qui intéresse le démographe, ou du foyer qui intéresserait le comptable national), et de prévoir un dernier degré de sondage (avec en général un tirage exhaustif, ou plus tard le « Kish », que Lévy-Bruhl a d'ailleurs introduit théoriquement en France) pour atteindre les individus dans le cas où des questions leurs sont destinées. Mais ce choix est évidemment un compromis, c'est-à-dire un point de faiblesse du sondage aléatoire que les partisans de Gallup ne manquent pas d'exploiter (par exemple dans la discussion de (Lévy-Bruhl, 1951), en disant d'une part que l'effet de grappe de ce sondage à deux degrés (ménages puis individus) ne manquera pas d'affaiblir la précision des résultats en cas de forte corrélation intra-grappe, selon la théorie même de ses partisans, ce dont Thionet se défend par l'argument que dans la plupart des cas cette corrélation est faible; en disant d'autre part que les statisticiens ne construisent pas des échantillons représentatifs des individus et ne peuvent pas répondre correctement à des questions qui concernent les choix individuels. Il est patent en effet que ce choix de méthode a déterminé en retour une sélection des thématiques et des champs des enquêtes, au moins pour la période étudiée. L'exclusion des célibataires au début des enquêtes budgets, et celle des individus appartenant à des ménages collectifs (militaires, religieux...) ne s'est pas toujours justifiée dans le fond mais par la forme. La forme ménage des unités a privilégié dans les budgets de famille l'étude des dépenses alimentaires et de logement aux dépens des dépenses d'habillement, des sorties et des vacances, et des consommations prises à l'extérieur du logement. Cette même forme a imposé un déclarant intermédiaire (la « mère de famille », le « chef de ménage ») entre le consommateur final des biens et services et le statisticien (CREDOC, 1960), ce qui ne pouvait que dégrader la qualité de l'information. On voit que le choix du ménage comme unité statistique a été particulièrement inadapté aux enquêtes de budgets et l'on devine qu'il n'a pas toujours été optimal dans les enquêtes emploi.

Si l'on considère que la standardisation de ces plans d'enquêtes, observée après 1950 a pu induire le caractère uniforme de la méthodologie et son autonomisation comme « forme », on peut se demander si l'industrialisation du procédé qui en a résulté n'a pas amoindri sérieusement la souplesse originelle d'une enquête par sondage et réduit à néant toute possibilité d'adapter la technique de sondage aux nouveaux objets d'enquête qui sont apparus dans l'agenda des décideurs politiques : mieux évaluer les consommations, le chômage, la précarité nécessitait parfois de quitter la logique « ménage ». Par ailleurs, la représentativité de l'échantillon au hasard n'est pas toujours une nécessité absolue, et quelques chercheurs (au CREDOC par exemple) ont montré qu'il fallait au contraire surseoir à l'exigence de représentativité et choisir la sur-représentation de certaines catégories lorsque certains projets de recherches comparatives ou heuristiques ne s'intéressaient pas à des estimations de pro-

portions ou de moyennes mais à des effets qualitatifs, des oppositions, ou des enchaînements comportementaux. Le sondage aléatoire n'était pas forcément la panacée.

## 10. Conclusion

Nous laisserons à d'autres études, passées (Desabie, 1987), (Glaude, 1996), (Desrosières, 1996) ou à venir, le soin de répondre à cette question d'une sclérose possible de la méthode ou de sa capacité à évoluer dans la période suivante. Étant rappelé tout de même que c'est le devenir normal et souhaitable de toute innovation que de se standardiser pour s'industrialiser, mais aussi d'être remise en cause lorsque cette standardisation empêche toute évolution et toute adaptation aux nouveaux besoins.

Le bref épisode que nous avons évoqué a joué un rôle important dans la méthodologie d'enquête qui s'installe pour un demi siècle. Sans doute est-il marqué par les conditions de l'après-guerre : le contexte d'une reconstruction appuyée sur la relance de la consommation et l'aide américaine, l'émergence des disciplines économiques (Armatte, 2003) et l'essor des sciences sociales (Drouard, 1982), la mise en place de nouvelles institutions étatiques, le rôle qu'y joue le corps des ingénieurs-statisticiens (Desrosières, 1994), les nouveaux enjeux posés aux enquêtes par l'introduction de la comptabilité nationale et la modélisation macroéconométrique (Armatte et Desrosières, 2000). L'enquête devient dans les années cinquante un dispositif important pour la reconstruction de l'État-Providence et pour sa mutation vers l'État-Entrepreneur : couplée aux méthodes de la Comptabilité Nationale, l'enquête permet d'asseoir les prévisions à courts terme et la planification à moyen terme sur des bases empiriques, et constitue le tableau de bord indispensable d'une action raisonnée de politique économique. Cette innovation qu'est l'enquête par sondage participe d'un système complexe de régulation économique et social qui domine les Trente Glorieuses et ne se réduit pas à une invention technique.

Rendons par conséquent un dernier hommage à la personne de Pierre Thionet qui sut transformer cette invention en innovation, qui sut batailler pour ne jamais dissocier exigences scientifiques et investissement dans des formes administratives souvent contraignantes, mais nécessaires à l'effort de reconstruction et à l'esprit de service public. Même s'il perdit quelques batailles. Même si celles qu'il a gagnées voient parfois leurs conquêtes menacées.

## RÉFÉRENCES

- AFFICHARD J. (1987), L'enquête sur l'emploi, *Pour une Histoire de la Statistique*, tome 2. Paris, INSEE.
- ANTONELLI A. (1921), Une enquête régionale sur le coût de la vie, *Revue d'Économie Politique*, **19**.
- ARMATTE M. (1988), «La construction des notions d'estimation et de vraisemblance chez Ronald A.Fisher», *Journal de la Société Statistique de Paris*, tome

- 129, N° 1-2, 1988, repris dans *Estimation et sondages*, J. Mairesse ed., Paris, Economica.
- ARMATTE M. (1995), *Histoire du Modèle Linéaire. Formes et usages en Statistique et en Économétrie jusqu'en 1945*, Thèse de Doctorat, dir. Jacques Mairesse, EHESS, 1995, 970 p. Chap.12.
- ARMATTE M. (2003), « Les sciences économiques et la pax americana », in Pestre & Dahan (eds), *les sciences dans et pour la guerre, 1940-1960*, Paris, Presses de l'EHESS, à paraître.
- ARMATTE M. et DESROSIÈRES A. (2000), Méthodes mathématiques et statistiques en économie : nouvelles questions sur d'anciennes querelles, dans J.P. Beaud et J.G. Prevost (eds.), *l'ère du chiffre, systèmes statistiques et traditions nationales*. Montréal, Presses Universitaires du Québec, 431-481.
- BLONDIAUX L. (1998), *La fabrique de l'opinion. Une histoire des sondages*. Paris, Le Seuil.
- BOARD OF TRADE (1909), *Cost of living in French towns*. Londres, Darling & Son.
- BOWLEY A. L. (1901), *Elements of Statistics*, King and Son, London. 2<sup>ème</sup> éd. 1902, 335 p.; 4<sup>ème</sup> éd. 1920, 454 p.; Traduction française sur la 5<sup>ème</sup> édition par L. Suret et G. Lutfalla, 1929.
- BOWLEY A. L. (1906), Presidential Adress to the economic section of the British Association, *JRSS*, 540-558.
- BOWLEY A.L. (1926), Measurement of the precision attained in sampling, *Bulletin I.I.S.*, XXII-I.
- BRU B. (1988), Estimations Laplaciennes, *JSSP*, 129, 1-2, 1988, et Mairesse (ed), 1988, 7- 46.
- CREDOC (1960), La consommation des ménages français en 1956, *Consommation*, VII-2, (avril-juin). Paris, Dunod.
- DARMOIS (1948), *Statistiques mathématiques*, Leçons données à l'École d'Application en 1937-38, 1944-45, 1947-48, et rédigées par P. Thionet, INSEE.
- DELPECH H. (1938), *Recherches sur le niveau de vie et les habitudes de consommation (Toulouse 1936-38)*. Paris, Sirey.
- DEMING W.E. (1950), *Some Theory of Sampling*. New York, Wiley.
- DESABIE J. (1965), *Théorie et pratique des sondages*. Paris, Dunod.
- DESABIE J. (1987), Les enquêtes sur les conditions de vie des ménages, dans *Pour une histoire de la Statistique*, tome 2, éd. J. Affichard, Paris, Economica/INSEE, 253-286.
- DESROSIÈRES A. (1988), La partie pour le tout : comment généraliser ? la préhistoire de la contrainte de représentativité, dans *Estimation et sondages*, J. Mairesse éd. Paris, Economica.
- DESROSIÈRES A. (1994), De l'École de l'INSEE à l'ENSAE et l'ENSAI : 1942-1994. Éléments d'histoire d'une école de statistique et d'économie, Conférence ICOTS4. Marrakech, Publ. I.I.S.
- DESROSIÈRES A. (1996), Essai sur l'évolution des usages des enquêtes sur les budgets de familles 1850-1970, dans *L'évolution des enquêtes « condition de vie » de l'INSEE du XIXème siècle à nos jours*, INSEE (cinquantenaire).
- DESROSIÈRES A. (2000), *La politique des grands nombres. Histoire de la raison statistique*, 1<sup>ère</sup> éd. 1993, 2<sup>ème</sup> éd. 2000. Paris, La Découverte.

- DIDIER E. (2001), Les premiers sondages américains, naissance de la représentativité et démocratie, dans *Enquêtes, modèles et applications*, J.J. Drosesbeke et L. Lebart. Paris, Dunod, 29-43.
- DROUARD A. (1982), Réflexion sur une chronologie : le développement des sciences sociales en France de 1945 à la fin des années soixante, *Revue Française de Sociologie*, 55-85.
- DROESBEKE J.-J. et LEBART L. (2001), *Enquêtes, modèles et applications*. Paris, Dunod.
- DROUARD A. (1983), Les trois âges de la Fondation française pour l'étude des problèmes humains, *Population*, XXXVIII, 6, 1017-1047.
- DUGE DE BERNONVILLE L. (1916-1917), Enquête sur les conditions de la vie ouvrière et rurale en France en 1913-1914, *Bulletin de la S.G.F.*, tome IV, 1916, 85-108; 1917, 185-221.
- FIENBERG S.E., TANUR J.M. (1996), Reconsidering the fundamental contributions of Fisher and Neyman on experimentation and sampling, *International Statistical Review*, 64, 237-253.
- GINI C. (1928), Une application de la méthode représentative aux matériaux du dernier recensement de la population italienne (1<sup>er</sup> décembre 1921), *Bulletin I.I.S.*, XXIII-2, 198- 215.
- GIRARD A. (1947a), Sondages et Démographie, *Population*, 273-282.
- GIRARD A. (1947b), Les Conditions d'existence des familles. Les dépenses de 1 080 familles en mars et juin 1946, *Population*, 245-266.
- GIRARD A. (1947c), Les Conditions d'existence des familles. Les dépenses de 939 familles de militaires en novembre 1946, *Population*, 421-449.
- GIRARD A. (1948), Les Conditions d'existence des familles. Les budgets familiaux en 1946, *Population*, 245-266.
- GIRARD A. (1986), *L'Institut National d'Etudes Démographiques. Histoire et Développement*. Paris, INED, 253 p.
- GLAUDE M. (1996), Les enquêtes budget de famille des années 1960 à nos jours, dans *L'évolution des enquêtes « condition de vie » de l'INSEE du XIX<sup>ème</sup> siècle à nos jours*, INSEE (cinquantenaire).
- HALBWACHS M. (1913), *La classe ouvrière et les niveaux de vie*, recherches sur la hiérarchie des besoins dans les sociétés industrielles contemporaines. Paris, Alcan.
- HANSEN M.H., HURWITZ W.N., MADOW W.G. (1953), *Sample Survey Methods and Theory*. New York, John Wiley and Sons.
- HUBER M. (1944), Principales lacunes et insuffisances des statistiques françaises, *JSSP*, 219- 231.
- INSEE et CREDOC (1960), La consommation des ménages français en 1956, *Consommation*, VII, avril-juin 1960, N°2. Paris, Dunod.
- JENSEN A. (1926), Report on the representative method in statistics, *Bulletin I.I.S.*, XXII, 359-377.
- KIAER A.N. (1895), Observations et expériences concernant des dénombrements représentatives, *Bulletin de l'IIS*, IX-1, 176-183.
- KRUSKAL W.H. et MOSTELLER F. (1980), Representative sampling, IV : the history of the concept in statistics, 1895-1939, *International Statistical Review*, 48, 169-195.
- LETELLIER G. (1949), *Dépenses des chômeurs et valeur énergétique de leur alimentation*. Paris, Sirey.

- LEVY-BRUHL R., THIONET P., HANSEN H., YATES F. (1949), *Sondages en Agriculture*, Centre Européen d'Application Statistique Agricole et Démographique. Paris, INSEE.
- LEVY-BRUHL R. (1951a), Enquêtes récentes par sondage sur l'emploi; avec discussion, *JSSP*, 108-125.
- LEVY-BRUHL R. (1951b), Une enquête par sondage sur l'emploi, *Supplément au BMS*, janv.-mars, p.1. Seconde partie « Erreurs d'échantillonnage » (par P. Thionet), avril-juin.
- LEVY-BRUHL R. (1951c), Les sondages d'opinion et l'échec des prévisions électorales de Gallup en 1948, *JSSP*, 92, 63-70.
- LEVY-BRUHL R. (1977), Souvenirs de la période 1945-1960, dans *Pour une Histoire de la Statistique*, tome 1. Paris, INSEE.
- LIE E. (2001), The Rise and Fall of the Sampling Surveys in Norway, 1875-1906, in *Sampling humans*, Preprint 166, Max-Planck-Institut.
- MAIRESSE J. (ed), (1988), *Estimation et sondages, cinq contributions à l'histoire de la statistique*. Paris, Dunod.
- MARIE G. (1945), Nécrologie de René Carmille, *JSSP*, juillet-août.
- MESPOULET M. (2000), Du tout à la partie; l'ge d'or du sondage en Russie (1885-1924), *Revue d'études comparatives Est-Ouest*, 31-2, 5-49.
- MORICE E. (1945), *Les méthodes statistiques modernes et leur expérimentation à l'expérimentation agricole*. Paris, INSEE, *Études théoriques*, N°1.
- MOSSER R. (1945), La Statistique aux États-Unis, *JSSP*, 213-218.
- NEYMAN J. (1934), On the two different aspects of the representative method : the method of stratified sampling and the method of purposive selection, *J.R.S.S.*, 97, 558-606.
- O'MUIRCHARTAIGH C. et WONG S.T. (1981), The impact of sampling theory on survey sampling practice : a review, *Bulletin I.I.S.*, tome XLIX-1.
- PORTE J. (1954), Une enquête par sondage sur l'auditoire radiophonique, *Supplément BMS*, janv.-mars, p. 31.
- REMERY R. (1948), Budgets de famille de salariés de la région parisienne en 1946, *Bulletin de la SGF*, Avril-juin, p. 155.
- REMPP J.M. Historique des enquêtes françaises sur les dépenses des ménages, Document INSEE.
- S. G. F. (1919), *Condition de la vie ouvrière et rurale en France en 1913-14*. Paris, Alcan, 100 p.
- SAUVY A. (1949), *Statistique et conjoncture économiques*, Cours à l'Institut d'Études Politiques 1948-49. Paris, Centre de documentation universitaire.
- SAUVY A. (1950), Cinq enquêtes sociales. Préface, *Travaux et documents*, N°9, INED, 7-16.
- SENG Y.P. (1951), Historical survey of the development of sampling theories and practice, *JRSS*, 214-231.
- STOETZEL J. (1943), *Théorie des opinions*. Paris, Presses Universitaires de France; compte-rendu par J. Bourgeois, *JSSP*, 1944, 41-46.
- TASSI P. (1988), De l'exhaustif au partiel : un peu d'histoire sur le développement des sondages, dans *Estimation et sondages*, J. Mairesse (ed), Paris, Economica, 117-133.

- THIONET P. et HEISCH M. (1954), L'enquête annuelle par sondage sur les salaires et la main-d'œuvre salariée permanente en agriculture, *Supplément BMS*, avril-juin, p.1.
- THIONET P. et LEVY-BRUHL R. (1947-48), *Théorie et pratique des sondages*. INSEE, École d'Application, 193 p.
- THIONET P. et LEVY-BRUHL R. (1948), Une enquête par sondage sur le marché des textiles, *Bulletin de la SGF*, juillet-sept, p. 250.
- THIONET P. et LEVY-BRUHL R. (1949a), Une enquête par sondage sur le ravitaillement, *BSGF*, p. 162.
- THIONET P. et LEVY-BRUHL R. (1949b), Premiers résultats concernant les dépenses des familles de salariés de la région parisienne au printemps 1948, *BSGF*, p. 37.
- THIONET P. (1945), L'école moderne des statisticiens italiens, *JSSP*, 245-255. Réponse de L. Galvani dans le *JSSP*, 1947, 196-208.
- THIONET P. (1946), *Méthodes statistiques modernes des administrations fédérales aux États-Unis*. Paris, Hermann.
- THIONET P. (1947), Rapport sur les sondages du 7 mai, Archives INSEE.
- THIONET P. (1953), *La théorie des sondages*. INSEE, *Études théoriques*, N°5.
- THIONET P. (1954a), Quelques données sur les sondages de l'INSEE en 1952-53, *JSSP*, 165-196.
- THIONET P. (1954b), L'enquête annuelle par sondage sur les salaires et la main-d'œuvre salariée permanente en agriculture, *Suppl. BMS*, Avril-juin, p.1.
- THIONET P. (1958), *La perte d'information par sondage*, Thèse de doctorat, Paris.
- THIONET P. (1978), « *Quelques problèmes concernant les sondages* ». Göttingen, Vandenhoeck & Ruprecht, 137 p.
- TOUCHELAY B. (1993), « *L'INSEE des origines à 1961 : évolution et relation avec la réalité économique, politique et sociale* », Thèse Paris XII.
- YATES F. (1949), *Sampling Methods for Censuses and Surveys*. Londres, Griffin. Trad fr. sous la dir. de G. Darmais, Paris, Masson, 1950.