

# Le service public de la statistique en occident : le point actuel<sup>1</sup>



**Edmond MALINVAUD**

Professeur au Collège de France  
Ancien Président des Sociétés de statistique

Alors que la production statistique a beaucoup crû en quantité et en qualité, sa précision ne répond pas toujours aux attentes et des lacunes subsistent. Face aux difficultés budgétaires, à un accueil souvent moins favorable du public et à des freins entravant une utilisation efficace des données administratives, les statisticiens réussissent à s'adapter.

De par sa vocation et sa tradition, la Société de Statistique de Paris s'intéresse au service public de la statistique. Or, il se trouve qu'elle en a peu parlé récemment. Mon départ de l'INSEE fournit un prétexte pour combler ce qui pourrait apparaître comme une lacune ; prétexte légitime, car j'aurais pu tout aussi bien faire cet exposé l'an dernier mais l'éloignement risquerait de rendre plus difficile que je le retarde à l'an prochain.

Notre intérêt doit se porter sur les tendances de fond et les problèmes principaux, non sur les détails anecdotiques auxquels nos contemporains, fussent-ils intellectuels, sont trop enclins à prêter attention. Je voudrais donc m'interroger aujourd'hui sur les grandes évolutions qu'ont connu les services statistiques des pays occidentaux et sur les défis auxquels ils sont actuellement confrontés. Cet objectif, de même que la longueur raisonnable de mon exposé, imposent que je sois sélectif. Mes collègues statisticiens officiels trouveront sans doute que je passe trop sous silence des aspects qui leur tiennent à cœur ; effectivement mon choix peut être contesté.

Je parlerai d'abord des produits fournis par le service public de la statistique, puis du processus de production.<sup>2</sup>

## I - LES PRODUITS

Bénéficiant d'investissements souvent anciens et des progrès techniques concernant le traitement des informations, la production statistique a beaucoup crû depuis quarante ans en quantité et en qualité. Cependant cette croissance ne suffit pas à répondre à la demande qui se développe à un rythme comparable. En particulier deux causes d'insatisfaction demeurent : la précision de beaucoup de statistiques ne répond pas à ce que l'on attend ; le service public de la statistique est incapable de fournir des informations fort demandées. Reprenons ces divers points.

1. Communication faite devant les Sociétés de statistique le 22 juin 1988

2. J'ai traité récemment d'un thème voisin, mais en insistant sur d'autres aspects du sujet dans E.Malinvaud, « Production statistique et progrès de la connaissance », ISTAT, Annali di Statistica, Anno 116, Serie IX, vol.7, 1987

## I. Des statistiques plus nombreuses et de meilleure qualité

La comptabilité nationale ne nous fournit pas d'évaluation valable du volume des services rendus par les producteurs publics de statistiques. Je ne me hasarderai donc pas à avancer des chiffres pour le taux de croissance de ce volume. Je me contenterai d'évoquer une image, celle d'une courbe logistique. Plus ou moins tôt, les divers pays occidentaux ont décidé durant ce siècle de se doter d'une information statistique de bonne qualité, couvrant les domaines principaux de l'activité économique et de la vie sociale; les difficultés de mise en place des organismes et des instruments statistiques ont souvent fait que le démarrage a été lent ; mais après un temps la croissance de la quantité et de la qualité des statistiques fournies s'est accélérée ; puis les besoins les plus urgents paraissant satisfaits, les coûts de progrès supplémentaires devenant de plus en plus élevés, le rythme de développement s'est ralenti. Aujourd'hui la plupart des pays occidentaux en sont arrivés à cette troisième phase ; mais la croissance y est encore positive et appréciable.

Ce progrès se constate de quatre façons à propos desquelles je peux donner l'exemple de quelques réalisations françaises récentes, à titre illustratif.

Premièrement, des statistiques entièrement nouvelles apparaissent dans certains domaines non précédemment bien étudiés. C'est le cas pour les statistiques françaises sur les entreprises de service qui font depuis quelques années l'objet d'une enquête annuelle. C'est le cas aussi pour la statistique mensuelle des créations d'entreprises. Ce sera bientôt le cas pour la connaissance précise des retraites perçues par toute catégorie de retraités que l'on imaginera.

Deuxièmement des statistiques anciennes sont enrichies, tantôt ayant une couverture plus large, tantôt fournissant un plus grand nombre d'informations. C'est le cas pour les statistiques françaises de prix à la production, ou pour les évaluations régionales et locales de l'emploi.

Troisièmement la qualité de statistiques anciennes est améliorée, et ceci de diverses façons. Il peut s'agir d'une plus grande rigueur conceptuelle dans la définition des grandeurs, par exemple pour les prélèvements obligatoires ou pour le sous-emploi. Il peut s'agir d'une plus grande précision de la mesure, comme pour le suivi mensuel des demandeurs d'emploi. Il peut s'agir de l'élaboration de données plus homogènes couvrant des unités géographiques diverses ou une longue durée, comme ceci se fait de mieux en mieux à l'occasion de chaque changement de base des comptes nationaux ou d'un indice important. Il peut s'agir enfin d'une meilleure lucidité et d'une meilleure information sur la précision effective des données élaborées, comme ceci ressort d'études méthodologiques ou d'un calcul plus fréquent des erreurs d'échantillonnage.

Quatrièmement la diffusion des statistiques est améliorée, en sorte que les utilisateurs puissent en disposer plus aisément. Ce peut être l'établissement de bureaux ayant cette vocation particulière, comme les observatoires économiques régionaux et leurs réseaux de correspondants. Ce peut être l'ouverture d'un service nouveau d'information comme le 3615 - INSEE sur Minitel. Ce peut être la constitution de banques de données comme la Banque de Données Locales regroupant sur chacune des communes françaises tout un ensemble de données les concernant, avec la possibilité de regroupements géographiques quelconques. Ce peut être enfin la mise au point de fichiers statistiques propres utilisables par des chercheurs en vue d'exploitations particulières, comme ceci se fait pour les grandes enquêtes par sondage.

## 2. Cependant une précision souvent encore médiocre

Ce n'est pas ici qu'il faut rappeler que le grand avantage de l'évaluation chiffrée est d'apporter la précision, mais que c'est aussi son danger quand la précision est illusoire. La lecture ou

l'audition d'un résultat chiffré donne toujours, même chez les plus avertis, une première impression d'exactitude. Or les statistiques sont rarement précises jusqu'à ce qu'il est convenu d'appeler leur dernier chiffre significatif. La situation commune est même celle dans laquelle ce dernier des chiffres dits significatifs n'a rien de significatif.

Sans doute dois-je faire ici une incidente afin de rappeler pourquoi une telle pratique est normale. C'est une vérité, et il est malheureux que l'éducation française évite de la mettre en valeur, que l'activité humaine opère dans un monde incertain, mais qu'il faut bien distinguer incertitude et ignorance. Dès lors qu'elle est objective, c'est-à-dire non intentionnellement biaisée, une information imprécise est meilleure que pas d'information du tout. L'utilisateur d'une statistique a ainsi intérêt à connaître le résultat obtenu tel quel, et ceci d'autant plus qu'il s'agit d'un utilisateur plus sérieux. De plus, une statistique est rarement isolée ; elle figure le plus souvent à côté d'autres avec lesquelles elle entretient des relations de cohérence, souvent apparentes dans les tableaux statistiques. Or les arrondis perturbent ces cohérences, ce qui interdit de les pousser jusqu'à la limite du significatif.

Donc une statistique imprécise est meilleure que pas de statistique du tout. Cependant, j'ajoute immédiatement que l'imprécision devrait être reconnue et même annoncée. Certes le statisticien ne doit pas pêcher par masochisme et dévaluer son travail ; mais renseigner sur le degré de précision d'une statistique est souvent aussi important que de fournir le résultat obtenu. Le renseignement peut rarement être lui-même précis puisque, à côté des erreurs d'échantillonnage, figurent d'autres erreurs de mesure non quantifiables ; il est néanmoins valable pour beaucoup d'utilisations. Je dois reconnaître avoir souvent été mal à l'aise à me dire que les statisticiens français faisaient des efforts peut-être insuffisants pour renseigner sur la précision de leurs résultats.

Quoi qu'il en soit, cette précision est souvent médiocre par rapport aux besoins d'information, et ceci surprend périodiquement ceux qui faisaient une confiance aveugle aux statistiques. En voici deux exemples révélés l'un et l'autre par l'incohérence entre des statistiques d'origines différentes.

Selon les définitions adoptées, l'effectif de la population active est rigoureusement égal à la somme du nombre des personnes ayant un emploi et du nombre des chômeurs. Cependant les évaluations, provenant pour ces trois nombres de sources différentes, ne respectent pas dans leurs variations temporelles l'identité en cause ; le calcul du chômage par différence entre population active et emploi donne, certaines années, une évolution nettement différente de celle fournie par la statistique des chômeurs. Le problème avait été discuté publiquement aux États-Unis il y a vingt ans. Il l'a été à nouveau en France il y deux ans.<sup>3</sup>

Autre exemple : « le trou noir » des balances des paiements, plus exactement le fait que, depuis un peu moins de dix ans, le total des soldes des balances des paiements courants de l'ensemble des pays du monde est fortement négatif, alors qu'il devrait être rigoureusement nul. Le trou a dépassé 100 milliards de dollars en 1982 et semble être actuellement de l'ordre de 60 milliards de dollars. L'étude d'un groupe d'experts réunis par le Fonds monétaire international a localisé la cause essentielle de cette « discordance statistique », ou tout au moins de sa croissance à partir de 1980 ; elle réside dans les revenus de certains investissements financiers qui, recensés dans les pays débiteurs, ne le sont pas dans les pays créditeurs.<sup>4</sup>

Dans un cas comme dans l'autre, l'étude de ces incohérences statistiques suggère des propositions constructives qui devraient, espère-t-on, améliorer les choses. Mais elle révèle

3. E. Malinvaud, Sur les statistiques de l'emploi et du chômage. Documentation française, Paris 1986.

4. Voir P. Esteva: « Paiements internationaux: le mystère de la discordance statistique est-il levé? », Courrier des statistiques, Octobre 1987.

aussi que le problème subsistera, car il provient des limitations inévitables des sources primaires d'information. Or ces exemples, où l'imprécision se manifeste au grand jour, ont évidemment bien des analogues moins visibles dans les domaines les plus divers.

### 3. De nombreuses demandes encore insatisfaites

Plus abondantes sont les informations diffusées, plus la soif d'informations s'exprime. A vrai dire, les demandes adressées au statisticien ont parfois un caractère abusif ou contradictoire ; le statisticien doit alors ne pas chercher à répondre à tout. On l'a bien vu il y a quelques années aux États-Unis quand l'information la plus attendue et commentée était celle relative à la masse monétaire hebdomadaire, comme si la situation économique et financière d'un pays pouvait changer d'une semaine à l'autre. On le voit de façon continuelle alors que quasiment les mêmes commentateurs demandent un jour aux statisticiens des évaluations précoces et détaillées, et se gaussent le lendemain des révisions apportées aux évaluations statistiques.

Mais il y a aussi des demandes apparemment raisonnables auxquelles le statisticien voudrait pouvoir répondre et auxquelles il ne le peut pas bien, compte-tenu de ce qu'est l'organisation de la société dans laquelle il vit et des moyens dont il peut raisonnablement espérer disposer. J'en donnerai trois exemples.

Là où existe une statistique nationale couramment utilisée, on imagine trop aisément que puisse aussi exister une statistique pour chaque région et même pour chaque localité ; mais le statisticien ne dispose pas de moyens réels pour fournir cette information localisée. La demande peut émaner d'un simple souci de connaissance désintéressé. Elle peut aussi avoir un objectif décisionnel direct. Ainsi, il arrive souvent que l'on ait à répartir des crédits ou des prélèvements obligatoires entre unités géographiques auprès desquelles la gestion est décentralisée. Pour asseoir cette répartition sur des bases objectives et pertinentes, on demande au statisticien des données souvent assez spécifiques sur les populations-cibles. Celui-ci se trouve alors confronté à l'alternative de soit refuser sa collaboration, soit fournir des données très imprécises ou très inadéquates aux objectifs poursuivis. Le problème s'est posé tel quel aux États-Unis pour la répartition des fonds fédéraux, et dans la Communauté Économique Européenne pour la politique régionale. Bien que les objectifs aient été plus diffus, les statisticiens français ont été aussi mis sur la sellette de façon fort désagréable en 1982 au moment de la décentralisation des pouvoirs.

De même existe de façon récurrente une demande pour une information précise sur des activités cachées, plus ou moins frauduleuses quoique souvent tolérées. Tantôt il s'agit des revenus réels de certaines professions non salariées, par opposition à leurs revenus fiscaux. Tantôt il s'agit de productions totalement irrégulières ou se situant en dehors de l'économie normalement organisée. Tantôt il s'agit de l'immigration clandestine ou des avortements clandestins, etc. Parfois le statisticien arrive à répondre à de telles demandes par des voies détournées, comme par exemple pour les revenus des agriculteurs ; mais il s'agit alors de cas concernant des activités non totalement cachées. Il subsiste des cas dans lesquels les bases manquent pour confirmer ou infirmer les informations fantaisistes qui ne manquent pas de circuler.

Mon troisième exemple concerne la demande d'un suivi statistique des techniques de production. On voudrait avoir une information qui décrive la transformation des techniques dans ses diverses composantes : procédés, degré d'automatisation, consommations de matières premières et énergie, etc. Il s'avère d'une part que les méthodes de production sont à notre époque si complexes, nombreuses et variées que la description statistique, nécessairement simple, se conçoit difficilement, d'autre part qu'aucun enregistrement systématique des techniques employées n'existe à l'intérieur des entreprises, de sorte que l'information disponible

à cet égard dans les ateliers ne remonte guère.

Que conclure vis-à-vis des cas où des statistiques ont une précision insuffisante et de ceux où des demandes légitimes d'information ne sont pas satisfaites, ou le sont mal?

Le statisticien doit évidemment éviter l'attitude qui consisterait à nier l'existence du problème, simplement parce qu'il se sent en paix avec sa conscience. D'abord il ne peut pas être certain d'avoir vraiment épuisé toutes les possibilités, toutes les voies d'approche ; il peut manquer d'imagination ou de pugnacité. Ensuite son attitude serait incomprise par les utilisateurs dont les besoins sont bien réels. On comprend dès lors que des groupes de travail, associant producteurs et utilisateurs de statistiques, se réunissent pour étudier les problèmes coriaces identifiés par les uns ou les autres ; on le comprend quoique la création et la vie de ces groupes fassent souvent penser à la célèbre phrase de Guillaume d'Orange: « Point n'est besoin d'espérer pour entreprendre ni de réussir pour persévérer »

Le rôle du statisticien dans de tels cas est en effet de faire comprendre la nature des obstacles qui s'opposent aux progrès de l'information. Certains de ces obstacles au moins pourraient être levés, mais à des coûts financiers ou autres que la société refuse. Le choix du statu quo est admissible mais le problème ne doit pas être dissimulé.

## II. LES PROCESSUS DE PRODUCTION

Passant à la seconde partie de mon exposé, je ne vais pas m'étendre sur les gains considérables que la statistique publique a retiré de l'informatique au cours des dernières décennies. L'utilisation des modes modernes de traitement de l'information a fait disparaître de nos ateliers certaines des tâches parmi les plus fastidieuses qui soient. Elle a surtout permis de conserver à l'information au moment du dépouillement statistique toutes les dimensions de sa complexité. Alors qu'il fallait autrefois limiter assez étroitement le nombre des caractères que l'on relevait et arrêter, souvent une fois pour toutes, la liste des tableaux croisés que l'on produisait par le dépouillement, on peut aujourd'hui conserver et obtenir à coût assez faible des résultats sous des formes adaptées à telles demandes qui se manifesteront.

De plus la généralisation de l'informatique dans les administrations a ouvert la voie à des exploitations statistiques assez aisées des données rassemblées par les services dans le cadre de leur gestion courante. S'étant engagée tôt dans cette voie, la France est souvent apparue comme précurseur d'une évolution qui se manifeste un peu partout.

Je vais négliger aussi les progrès des techniques proprement statistiques. Ces progrès furent cependant considérables. Il y a quarante ans, les enquêtes par sondage aléatoire n'en étaient qu'à leurs débuts ; les concepts de la comptabilité nationale se cherchaient ; beaucoup de nos actuelles nomenclatures étaient inexistantes ou inadéquates ; la méthodologie pour le suivi de certains aspects de la vie économique et sociale était fruste, etc.

Soucieux d'insister ici plus sur les problèmes que sur les motifs de satisfaction, je vais concentrer mon attention sur trois questions: les difficultés budgétaires, l'accueil du public, enfin les freins à une utilisation efficace des données administratives.

### 1- Les budgets

Les déficits budgétaires, créés ou aggravés par le ralentissement de la croissance économique, ont partout amené les gouvernements à chercher à alléger le coût des charges collectives. Le service public de la statistique a souvent été visé dans la recherche des économies. Avec

l'expérience d'un certain nombre d'années et de nombreux pays, nous pouvons nous interroger sur les sacrifices qui en ont résulté et en résulteront.

Dans cette phase de l'histoire, les services statistiques ont souffert de deux handicaps et de trois avantages pour faire face.

Le premier handicap réside en ce que le coût budgétaire des statistiques est beaucoup plus aisément identifié que leur utilité et que l'efficacité des services chargés de les produire. Partout on a pu entendre des censeurs dire sans preuve que telles statistiques ne servaient à rien, ou qu'elles seraient obtenues à moindre coût par des organismes privés, ou qu'il y avait des doubles emplois, ce dernier langage étant d'autant plus facile que le système statistique était plus décentralisé et faisait donc intervenir un plus grand nombre d'unités. Parfois même le fait que la fourniture d'une information statistique ait la nature d'un service public a été perdu de vue.

Le second handicap tient au très long délai qui s'écoule entre la décision de construire une statistique nouvelle et le moment où elle a effectivement sa pleine utilité. Il serait facile de montrer que le délai moyen est de l'ordre de dix ans. Ce qui est à craindre est donc moins l'abandon de tout un pan de l'information statistique, que le tarissement des investissements dans un secteur important, pour lequel même les préoccupations de maintien de la qualité seraient abandonnées. La dégradation progressive de l'information statistique est ainsi probable sur de longues années dans les pays où les services statistiques ont fait l'objet des coupes les plus sombres.

Pour se protéger les statisticiens publics ont pu d'abord plaider. Au total l'information statistique ne coûte pas très cher par rapport à bien d'autres missions des administrations. A un âge où la collectivité dépense des sommes considérables en faveur des médias, il semblerait paradoxal d'abandonner le souci de la collecte régulière d'une information objective dont l'intérêt est reconnu, quoique non évaluable.

Puis les statisticiens étaient souvent prêts à réaliser des économies de personnel au fur et à mesure que l'informatique devenait plus performante, plus opérationnelle, plus généralement implantée dans les bureaux et les ateliers. De fait, une dégradation du service n'est probable que là où les économies ont imposé une réduction de l'effectif du personnel très qualifié, ce que dans d'autres pays on appelle « les professionnels ». Le cas s'est produit. Je ne pense pas qu'il ait été fréquent.

Enfin les statisticiens ont appris à réaliser quelques recettes en vendant des services qu'ils étaient capables de produire et pour lesquels une gestion efficace exigeait qu'il en soit ainsi. Tantôt il s'agit d'éviter le gâchis, comme lors de la vente des publications, de la tarification des services videotex ou de la cession de fichiers informatiques. Tantôt il s'agit de fournir à des entreprises ou organismes un service plus ou moins spécialisé dont il est normal que les bénéficiaires supportent le coût. Ces recettes ne constituent évidemment qu'un appoint, mais un appoint qui maintenant peut parfois atteindre 10 % des crédits budgétaires.

Au total j'estime quant à moi que, dans la plupart des pays, la rigueur budgétaire n'a pas gravement affecté le service public de la statistique. Il y a toutefois quelques exceptions.

La plus notable concerne les organismes internationaux, plus précisément même ceux appartenant à la famille des Nations Unies. Je n'ai pas besoin d'insister sur le double rôle que doivent jouer ces organismes, dans l'harmonisation des statistiques nationales, grâce à l'adoption de normes communes, et dans la conduite d'opérations statistiques internationales,

telle que celle concernant la comparaison du pouvoir d'achat des monnaies et des revenus réels<sup>5</sup>. Or la situation des services statistiques des Nations Unies est assez déplorable pour que l'on puisse avoir de sérieuses craintes sur la pérennité de ce rôle.

## 2. Accueil du public

Les statisticiens de pays occidentaux ont eu à faire face au cours des vingt dernières années à un autre groupe de difficultés concernant leurs rapports avec les enquêtés, auprès desquels ils collectent les informations de base. La multiplication des enquêtes et formulaires en tout genre provoque la lassitude et parfois le rejet. Les modes d'existence modernes rendent l'accès aux individus plus problématique qu'autrefois. Les craintes vis-à-vis du pouvoir d'inquisition de l'informatique compliquent nombre d'investigations. Considérons successivement ces trois sources de difficultés.

L'intérêt porté aux données statistiques a amené beaucoup d'administrations et d'organismes à lancer des enquêtes pour leur propre compte. Comme simultanément les programmes statistiques publics se développaient, la durée totale du temps consacré à répondre a beaucoup augmenté. Le phénomène n'a pas de gravité pour les ménages qui ne sont pas souvent individuellement concernés, les taux de sondage restant faibles pour eux. Il en va différemment avec les entreprises, qui ont aussi à répondre à beaucoup d'investigations émanant d'administrations ayant des pouvoirs réglementaires ou des fonctions de contrôle. Aussi les statisticiens se heurtent-ils parfois à la mauvaise humeur des entreprises enquêtées, ou même à l'exigence, imposée par les pouvoirs publics, d'avoir à alléger la charge des enquêtes. Comme précédemment à propos de l'observation des activités cachées, la véritable question est celle du bon arbitrage entre coût de collecte de l'information, y compris auprès des enquêtés, et utilité de cette information. Pour orienter cet arbitrage, il existe dans tous les pays avancés des organes consultatifs ayant reçu mission d'émettre des avis sur le programme des enquêtes statistiques publiques et sur ses modalités d'exécution. En France cet organe est le CNIS, Conseil National de l'Information Statistique, où siègent notamment les représentants des entreprises. L'expérience montre que de tels conseils peuvent bien fonctionner et éclairer utilement les décisions, surtout après quelques années, quand les contraintes et techniques de l'activité statistique ont été comprises par tous les membres. Mais ceci ne suffit pas toujours à aplanir les résistances à la base.

Bien que la charge effective de réponse aux enquêtes statistiques soit faible pour les ménages, on déplore aussi une certaine tendance à l'augmentation du taux de ceux que l'on n'arrive pas à enquêter, soit qu'ils refusent soit qu'on ne puisse pas les joindre. Le phénomène s'explique par l'importance croissante de la population urbaine, par la généralisation du modèle dans lequel tous les adultes du ménage travaillent en dehors du logement, par une certaine aggravation du sentiment d'insécurité etc. Pour y faire face, les statisticiens publics essaient d'adapter leurs méthodes d'enquête et de familiariser à l'avance le public avec leurs investigations.

Enfin, on a assisté à la diffusion d'une inquiétude, voire d'une psychose, vis-à-vis de la généralisation des fichiers informatiques de données individuelles. Or à certains stades au moins de leurs travaux les statisticiens utilisent de tels fichiers, qu'ils les aient constitués eux-mêmes ou qu'ils les aient obtenus en vue d'extraire l'information statistique latente dans des données collectées par les administrations pour leur gestion. Dans divers pays des lois ont été prises afin de régir la tenue des fichiers ; en France ce fut la loi de janvier 1978 sur l'informatique, les fichiers et les libertés.

---

5. M. Aufrant et H. Picard, « Les parités du pouvoir d'achat entre pays », *Courrier des statistiques*. Octobre 1987.

La législation a habituellement créé un organe chargé du suivi de son exécution, en France la CNIL, Commission Nationale de l'Informatique et des Libertés. Les statisticiens ont eu à exposer précisément l'objet et les modalités de leurs investigations et à tenir compte d'un certain nombre de contraintes qui leur ont été imposées. Celles-ci ont concerné tantôt la constitution et la conservation des fichiers statistiques, tantôt l'accès des statisticiens aux fichiers administratifs, tantôt les fusions de fichiers qui peuvent s'avérer utiles pour la richesse de l'information statistique<sup>6</sup>.

Suivant les pays l'adaptation fut plus ou moins facile ; mais opérer en accord avec un organe indépendant chargé de la protection des libertés a permis d'éviter le rejet massif de certaines investigations statistiques par la population, comme on l'a vu aussi parfois là où aucune législation spécifique n'existait, l'exemple le plus flagrant ayant été l'annulation en dernière minute du recensement de la population de l'Allemagne Fédérale en 1983.

### 3. Statistiques d'origine administrative

Le dernier point sur lequel je voudrais insister n'est la source de déception que dans la mesure où l'on avait pu entretenir à son sujet des illusions. Conscient de l'abondance des informations collectées par les administrations modernes dans le cadre de leur gestion, conscient aussi de la puissance de l'informatique, on a souvent imaginé que des statistiques toutes prêtes sortiraient aisément des données administratives. Comme il arrive souvent, l'imagination n'était pas foncièrement fautive, mais trop simpliste : entre son rêve et toute réalisation pratique il y avait bien des obstacles à franchir. Aujourd'hui, avec l'expérience acquise, notamment en France, nous pouvons mieux apprécier les contraintes inhérentes à cette voie pour l'obtention de statistiques<sup>7</sup>.

L'avantage en termes de coûts est évident. Pour les personnes qui ont déjà eu à répondre aux administrations comme pour les services statistiques, l'économie d'une nouvelle collecte est très importante, bien que la mobilisation statistique des données administratives comporte aussi des coûts qu'on avait toujours sous-estimés dans les débuts. Dans l'ensemble, la qualité des données élémentaires, si on les utilise avec discernement, est meilleure que dans les réponses fournies aux statisticiens. Il ne s'agit pas par exemple d'évaluer les revenus agricoles à partir des forfaits retenus par les administrations fiscales. Mais la réponse donnée à une administration est fournie avec plus de soin que celle donnée à une enquête purement statistique parce qu'une réponse erronée peut entraîner plus d'ennuis ; le plus souvent c'est donc aussi une réponse plus fiable.

Il reste que la production de statistiques a la nature d'un sous-produit de l'activité administrative. Ceci entraîne une certaine fragilité et une certaine inadéquation.

Une certaine fragilité, car les pratiques administratives peuvent varier. Après qu'ait été organisée une exploitation de fichiers administratifs d'un certain type, il se peut que les procédures de constitution de ces fichiers soient modifiées dans un souci d'efficacité propre à l'administration gestionnaire, que la définition des données change plus ou moins à la suite d'une législation ou réglementation nouvelle, que même les fichiers disparaissent totalement avec la suppression légale des contrôles administratifs qu'ils permettaient. Comme l'information statistique tire souvent sa valeur de sa continuité, tous ces avatars sont fort dommageables aux statisticiens qui doivent chercher à en corriger les effets mais ne le peuvent pas toujours.

6. Pour les règles élaborées en France, voir le chapitre 9, « Recherche et statistiques » dans CNIL, *Dix ans: d'informatique et libertés*, Economica, Paris 1988.

7. Pour un traitement plus complet, voir E. Malinvaud, « From statistics to data management - the French difficulties », *Statistical Journal of the United Nations*, vol. I, n° 3, p. 285-290, 1983.

Les données d'origine administrative concernent rarement exactement ce que le statisticien a pour mission d'observer. Elles sont conçues essentiellement pour les besoins des administrations. Elles sont donc plus ou moins adéquates. Ceci est vrai même en France où, à leur demande, les statisticiens interviennent souvent de façon précoce en vue de manifester leurs besoins quand une opération administrative est conçue ou réformée. Le champ couvert concerne celui de la compétence de l'administration et celui visé par la loi ou le règlement, non celui sur lequel on souhaite avoir une information économique ou sociale ; la liste des grandeurs observées comporte des lacunes par rapport à ce que le statisticien doit faire connaître ; la définition des grandeurs est conforme à l'objectif de l'administration, non à celui de la statistique, etc.

Vis-à-vis de ces manques d'adéquation, qu'il ne faudrait d'ailleurs pas dramatiser, la bonne solution est évidemment de concevoir les statistiques d'origine administrative comme complémentaires de celles obtenues directement par enquête statistique. Ainsi l'enquête annuelle sur l'emploi permet de cadrer et de mieux interpréter les résultats de la statistique mensuelle des demandeurs d'emploi. Ainsi encore l'enquête annuelle d'entreprise est conçue comme devant être complémentaire de l'exploitation systématique des déclarations que les entreprises font chaque année au fisc sur leurs résultats comptables.

\*  
\* \*

Ce tour d'horizon, trop rapide, a visé à décrire la situation présente et ses problèmes. Il témoigne de ce que le service public de la statistique n'a eu ni le goût ni le loisir de se complaire dans la routine. C'est une activité vivante en continuelle évolution, dont les progrès ne sont pas toujours faciles mais sont réels.