

# La statistique publique à l'ère du numérique : entre déclin, mission impossible et nouveau départ



## Quatre questions à

Dominique BUREAU

Président de l'Autorité de la statistique publique<sup>1</sup>

**Statistique et société :** Dans l'article du Guardian, William Davies déplore une « perte de crédibilité » des statistiques publiques. « Plutôt que de dissiper la controverse et la polarisation, il semble que la statistique les attise en réalité ». Ce constat semble confirmé, en France, par certaines enquêtes d'opinion<sup>2</sup>. Nombreux sont ceux qui font même le constat d'une crise de la quantification en général. Faites-vous aussi ce constat ? Y apporteriez-vous des nuances ?

**Dominique Bureau :** Notre pays a la chance de disposer d'un système d'information statistique très fourni et qui s'enrichit continuellement. À cet égard, le rapport de l'Autorité pour l'année 2016 signale, par exemple, la publication, pour la première fois, des résultats infra-communaux sur le revenu disponible et la pauvreté monétaire, ainsi que, comme les années précédentes, de nombreuses enquêtes ou publications inédites sur des sujets touchant de près nos concitoyens.

Cependant, malgré la quantité considérable d'informations et les progrès réalisés, le débat public demeure difficile, notamment sur les questions sensibles que constituent l'emploi, le chômage et la précarité de l'emploi, comme l'ont montré les travaux de la Commission d'enquête du Sénat sur les chiffres du chômage. Plus généralement, le public attend des chiffres variés couvrant tous les domaines de l'économie et de la société, reflétant la diversité des situations, permettant les comparaisons internationales... Dans ce contexte, la statistique publique n'échappe pas au constat de défiance et d'affaiblissement général de « l'autorité des chiffres ».

Les prophètes du déclin de la statistique ont cependant tendance à faire « feu de tout bois ». En effet, il ne faut pas confondre les insatisfactions vis-à-vis de certaines politiques ou services publics, ou celles ressenties par rapport aux situations rencontrées, avec les questions sur la production des chiffres. Et il faut distinguer les interrogations sur la qualité de la statistique de celles sur les lacunes en matière d'évaluation des politiques publiques ou dans la mobilisation de l'expertise pour les concevoir, ou encore par rapport à la participation du public à leur élaboration.

Il faut aussi se garder d'embarquer sans discernement la production statistique dans les débats sur la pertinence du modèle de décision rationnel des agents économiques ou du critère utilitariste pour les choix publics. En effet, la statistique est fondamentalement observationnelle et à visée descriptive. De plus, pour éclairer les enjeux macroéconomiques, la statistique doit estimer des agrégats, construits sur des bases transparentes et aussi pertinentes que possible. La résurgence de ces critiques alors même que, par exemple, jamais les travaux empiriques

1. Si les vues exprimées ici s'appuient sur les travaux de l'Autorité ( dont les rapports annuels sont consultables sur le site [www.autorite-statistique-publique.fr](http://www.autorite-statistique-publique.fr) ), elles n'engagent cependant que leur auteur.

2. Voir par exemple, à propos d'une enquête récente du Centre d'études de la vie politique en France ( Cevipof ), l'article de Jean Chiche dans le numéro 2016-4 de Statistique et société.

en économie comportementale et le souci de comprendre les écarts de perception sur les situations économiques et sociales n'ont été aussi actifs, est un autre paradoxe des débats actuels.

Enfin, s'il est vrai que l'irruption du *Big Data* conduit à renouveler la réflexion méthodologique sur les potentialités d'analyses prédictives « sans modèle » par rapport à l'économétrie privilégiant l'énoncé préalable d'hypothèses précises, contrôlables par des tests, force est de noter que le sujet n'est pas totalement nouveau non plus. Il importe donc de distinguer les différents aspects de cette défiance et d'identifier ce qui met plus spécifiquement en cause la statistique. Ceci conduit à observer un tableau plus nuancé.

En effet, les enquêtes de satisfaction réalisées par l'Insee en 2016 montrent beaucoup d'éléments positifs : la notoriété de l'Insee est élevée, associée en général (80 %) à une bonne opinion ; alors que la confiance du public sur les chiffres et données publiés sur la situation économique et sociale ne dépasse pas 43 %, ce chiffre atteint 67 % pour ceux publiés par l'Insee, avec comme première raison pour cette confiance, celle dans l'organisme qui les produit. Toutefois, il est exact que le public se reconnaît plus ou moins dans ces chiffres. Surtout, la première raison de défiance mise en avant concerne plutôt la manière dont les chiffres sont utilisés.

Pour autant, ces constats plus rassurants ne doivent pas conduire à ignorer les menaces et défis pour la statistique publique. En effet, celle-ci se fixant pour objectif de fournir à tout un chacun à des fins de prise de décision, de recherche et de débat public, des informations de qualité, élaborées en toute indépendance, sur l'économie et la société car cela constitue un fondement des processus démocratiques et le progrès de la société, la défiance actuelle interpelle et met nécessairement en première ligne la statistique.

**S&S :** S'interrogeant sur les raisons d'une possible « crise », on peut rapprocher les critiques de William Davies de celles qui avaient été émises au moment du lancement de la commission Stiglitz – Sen – Fitoussi (SSF). Les recommandations de cette commission ont-elles été suivies d'effet ?

**DB :** Le « Ramener les questions sociales et économiques à des agrégats numériques et des moyennes semble à beaucoup violer la décence politique » de Davies fait effectivement écho à la saisine de la Commission « SSF » en 2008, qui faisait état d'une « insatisfaction par rapport à l'état actuel des informations statistiques touchant à l'économie et la société » et lui donnait « mission de déterminer les limites du PIB comme indicateur des performances économiques et du progrès social, en soulignant les problèmes relatifs à sa mesure, d'identifier les informations complémentaires qui pourraient être nécessaires pour aboutir à des indicateurs plus pertinents du progrès social, d'évaluer la faisabilité de nouveaux instruments de mesure et enfin de débattre de la présentation la plus appropriée des informations statistiques ».

Face à ce qu'ils perçoivent comme des critiques excessives ou injustes, les statisticiens ont tendance à se mettre sur la défensive, et à rappeler que l'on n'a pas attendu la crise financière pour développer des indicateurs d'inégalités, cerner au plus près les caractéristiques de la pauvreté et l'exclusion ; ou pour essayer « de rendre compte des situations complexes et diversifiées du marché du travail par une batterie d'indicateurs sur l'emploi, le chômage, le sous-emploi et la précarité de l'emploi »<sup>3</sup> ; et, il y a plus de quarante ans, pour que Carré, Dubois et Malinvaud insistent sur la nécessité de comprendre les ressorts psychologiques et sociologiques de la croissance...

---

3. Cf. le rapport « Emploi, chômage, précarité – Mieux mesurer pour mieux débattre et mieux agir » du groupe de travail présidé par Jean-Baptiste de Foucauld, CNIS, 2008

Le rapport « SSF » reconnaissait d'ailleurs que bon nombre des questions abordées avaient été posées de longue date, par ceux-là même qui avaient contribué à élaborer nos systèmes actuels de comptabilité nationale. Toutefois - et c'est sans doute le point important à prendre en compte par les responsables statistiques - il considérait que, s'il avait été reconnu de longue date que le PIB posait problème en tant qu'outil de mesure des performances économiques, bon nombre des changements intervenus dans la structure de nos sociétés ont rendu ces déficiences plus criantes. Surtout, il mettait en avant que : « si la question de la mesure des performances économiques et du progrès social revêt de nos jours une importance particulière, c'est précisément parce que l'on craint que les mesures usuelles risquent d'encourager nos sociétés à évoluer dans une mauvaise direction, ce qui, la crise actuelle nous le montre, peut être générateur de détresse sociale et de dégradation du bien-être. »

Leur travail débouchait donc sur des recommandations opérationnelles, à partir de l'idée que les progrès de la recherche en de nombreuses disciplines rendent possible la conception de mesures plus larges du bien-être qui en intégreraient davantage d'aspects, notamment par rapport aux questions de soutenabilité de la croissance. En parallèle avec les réflexions sur les nouveaux indicateurs de richesse, notre système statistique s'est attaché à mettre en œuvre le programme de travail esquissé alors<sup>4</sup>, ce qui s'est avéré fructueux et montre que, plus que de nourrir les débats de principe ou idéologiques, ce que nous avons à faire est d'abord de répondre aux attentes du public en matière d'information statistique.

Dans cette perspective, notre système statistique public est maintenant en ordre de marche pour assurer la coordination des travaux statistiques relatifs aux indicateurs de suivi pour les objectifs de développement durable approuvés par l'Assemblée générale de l'ONU en septembre 2015. Le recensement réalisé à cette fin montrait que : parmi les 229 indicateurs (sans les doublons), 198 d'entre eux relèvent effectivement du domaine de la statistique, 31 relevant plutôt de la mise en œuvre des politiques publiques ; les services producteurs sont bien identifiés pour 84 % des indicateurs statistiques, le SOeS, l'Insee et la Drees<sup>5</sup> étant les plus grands pourvoyeurs ; par ailleurs, 65 % des indicateurs existent déjà dans une version exacte ou approchée.

**S&S :** Êtes-vous d'accord avec l'idée selon laquelle le *Big Data* serait un nouveau « mode » de quantification qui remplacerait les anciennes méthodes légitimes comme les enquêtes ou les comptes nationaux ? Pensez-vous qu'il existe un risque réel que les analyses tirées de *Big Data* par des organismes privés concurrencent les résultats de la statistique publique ?

**DB :** La transformation numérique bouleverse le fonctionnement des entreprises et des marchés, avec l'émergence, autour d'internet, de nouveaux canaux d'information pour mettre en relation les différents acteurs, et le développement par ceux-ci de stratégies mobilisant l'abondance de nouvelles données et leur traitement par les « data-sciences »<sup>6</sup>.

L'économie « numérique » se caractérise ainsi par la production de flux importants de données reflétant l'activité économique, issues de l'internet ou de différents capteurs, stockées sous des formes variées. L'exploitation de ces données associées au phénomène *Big Data* pour la production statistique suscite un intérêt croissant, ces données étant susceptibles de fournir de nombreuses opportunités : pour réduire les délais de publication, compte-tenu de

4. Cf. l'article. « Les préconisations du rapport Stiglitz-Sen-Fitoussi : quelques illustrations », Clerc M., Gaini M. et Blanchet D. in « L'économie française, édition 2010 » Insee

5. L'auteur cite trois services statistiques publics français : le Service de l'observation et des statistiques ( SOeS ) du ministère de l'Environnement, de l'Énergie et de la Mer, l'Institut national de la statistique et des études économiques ( Insee ), et la Direction de la recherche, des études, de l'évaluation et des statistiques ( DREES ) du Ministère des Affaires sociales et de la Santé.

6. Pour une description générale de la rupture que représentent ces données en termes de volumes, d'instruments pour les traiter et d'applications potentielles, cf. « Analyse des big data. Quels usages, quels défis ? », Hamel et Marguerit, France stratégie, 2013

la disponibilité immédiate de l'information ; pour disposer d'observations à des échelles plus fines ; pour compléter les indicateurs existants... Le projet « données de caisse » a ainsi été lancé en 2015, après une phase expérimentale en 2011, avec objectif de l'intégrer en production à l'horizon 2019.

Plus généralement, le groupe Cnis-Insee (2015) sur les potentialités de développements statistiques à partir de systèmes d'information (SI) de gestion privés a enclenché une dynamique. Ce groupe avait identifié notamment trois secteurs particulièrement prometteurs : outre l'utilisation des « données de caisse » pour la production d'indices de prix, les données de téléphonie pour mesurer la population présente et celles des cartes bancaires pour la consommation.

D'autres applications cherchent à utiliser d'autres sources, telles que : les données « satellites » (accessibles grâce au projet Copernicus) ; les requêtes des internautes, pour enrichir ou développer des méthodes alternatives pour la prévision conjoncturelle de la consommation (cf. *Google Trends*<sup>7</sup>), dont l'Insee a analysé les potentialités, avec, en l'état, des résultats peu concluants ; d'autres encore, à utiliser les données de systèmes de réservation ou celles d'usage de sites internet pour enrichir, par exemple, les statistiques culturelles, ou, plus généralement, combler les lacunes dans la mesure du volume d'activité du secteur tertiaire... Ce n'est qu'en les testant que la valeur de ces opportunités nouvelles pourra être établie. A ce titre, les projets en cours seront précieux.

L'adoption de la loi sur la République numérique fournit un cadre pour leur réalisation, puisque la statistique publique pourra désormais, pour les besoins d'enquêtes statistiques obligatoires, se voir transmettre sous forme électronique sécurisée des informations issues de certaines bases de données des personnes de droit privé concernées. Les conditions de confidentialité des informations communiquées par ces fournisseurs de données, socle de la confiance entre ceux-ci et le système statistique, sont ainsi établies, mais elles devront se décliner dans des conventions signées ensuite avec ceux-ci pour construire des cadres de coopération.

Ces nouvelles données offrent donc de nouvelles opportunités pour la statistique publique, qui doit par ailleurs développer les méthodes appropriées pour mesurer la valeur créée par la numérisation de l'économie. Elles obligent aussi à anticiper l'évolution de son rôle, notamment par rapport à l'émergence de nouveaux producteurs d'indicateurs statistiques, même si ceux-ci ne répondent pas, ou seulement partiellement, aux besoins d'information fiable et de qualité pour mesurer l'économie et les transformations sociales.

Le système statistique public (SSP) doit donc s'y préparer activement, en se gardant de relativiser les enjeux, même lorsque leur maturité semble encore incertaine. Certes, l'offre de nouveaux producteurs de données se situe encore « à côté » de la statistique, un peu comme, par exemple dans le domaine de la santé, avec les applications « data-mobiles » qui ont d'abord - mais cela évolue maintenant - concerné le bien-être plutôt que les traitements médicaux. Plus spécifiquement, il y a encore beaucoup de grain à moudre dans le développement de la statistique administrative, grâce notamment à l'appariement de ses fichiers, qui permet actuellement de renouveler l'éclairage de nombreux sujets.

Mais ceci ne doit pas conduire à négliger les perspectives de développement de la statistique à partir de nouvelles sources et, potentiellement l'émergence de nouveaux acteurs. En effet, la rigidité de l'offre publique la rend toujours fragile lorsqu'émergent des producteurs potentiellement « concurrents », même si dans un premier temps, le recouvrement des champs

---

7. Cf. dossier de la note de conjoncture publiée par l'Insee en mars 2015 : « Apports de Google Trends pour prévoir la conjoncture française : des pistes limitées » Clément Bortoli et Stéphanie Combes

d'activités peut sembler marginal. De plus, dans un monde où l'information est disponible de façon quasi-instantanée, il faut s'attendre à un renforcement des exigences du public en termes de réactivité, de capacité à qualifier les phénomènes et à publier des données fiables, avec aussi des difficultés pour en faire reconnaître la qualité ou l'objectivité dans un contexte de prolifération de l'information.

**S&S** : Faisant preuve d'un optimisme volontariste, le « statisticien national » anglais, John Pullinger, a répondu au pessimisme de l'article de William Davies en se réjouissant de la demande croissante adressée aux statistiques publiques, mais sans citer de piste pour satisfaire cette demande. Pour votre part, quels remèdes envisagez-vous pour faire face à la situation actuelle ?

**DB** : Effectivement, si l'on adhère à l'idée exposée par John Pullinger selon laquelle c'est le moment pour la statistique d'apporter sa meilleure contribution à la société, encore faut-il établir les conditions pour cela. Une première réponse se situe au niveau de la définition des programmes de travail appropriés pour mieux satisfaire les attentes des utilisateurs, de la résolution des problèmes de méthodes rencontrés, de la qualité des publications... Mais il faut placer cela dans une stratégie plus globale car la gouvernance est cruciale pour cela, avec, évidemment, comme premier principe celui d'indépendance.

Fondamentalement, l'existence même de l'Autorité de la statistique publique s'inscrit dans cette perspective, sa mission étant de veiller « au respect du principe d'indépendance professionnelle dans la conception, la production et la diffusion de statistiques publiques ainsi que des principes d'objectivité, d'impartialité, de pertinence et de qualité des données produites ». L'objectif ultime est ainsi que le public puisse avoir confiance en son système statistique. Et c'est la perspective de contribuer à celle-ci qui guide son action.

Pour cela, le contrôle du respect exemplaire des principes du Code de bonnes pratiques de la statistique constitue le cœur de l'activité de l'Autorité, avec leurs trois dimensions : des facteurs institutionnels et organisationnels déterminants de la crédibilité, notamment l'indépendance professionnelle, l'engagement sur la qualité, le secret statistique, l'impartialité et l'objectivité ; des procédures pour organiser, collecter, traiter et diffuser les statistiques ; et des résultats statistiques en termes de pertinence, exactitude et fiabilité, mais aussi d'actualité, de cohérence et d'accès pour les utilisateurs.

Mais c'est évidemment l'ensemble du système statistique qui est concerné. À ce titre, le plan d'action de gestion de la qualité au sein des processus de production statistique qui fait suite à la création du Comité stratégique de la qualité à l'Insee fait partie des avancées importantes à porter au crédit de 2016. Il est aussi essentiel que les programmes statistiques répondent aux attentes du public et que les parties prenantes se les approprient, avec une vision partagée. Le CNIS est donc aussi une pièce maîtresse.

Enfin, on ne peut laisser dans l'angle mort la question de l'utilisation qui est faite des statistiques, qui souvent nuit à l'image d'ensemble de la « marque » lorsqu'elle est inappropriée. Le rapport de Foucauld observait, par exemple, que, dans le débat public, le taux d'emploi est moins considéré que le taux de chômage, dont les variations pour ce qui concerne les chiffres mensuels de Pôle emploi, peu significatives, prennent une importance exagérée, par rapport aux chiffres sur l'ancienneté au chômage, par exemple. Plus généralement, l'enquête sur l'image de l'Insee auprès des « Inseenautes » pose une question sur les raisons de la défiance dans les données sur la situation économique et sociale de la France. La réponse principale met en cause non pas la valeur des chiffres mais ce qu'on leur fait dire. Ceci ne peut être négligé.

---

8. Op.cit. note 3.