

Autorité de la statistique publique

Séance du 22 juin 2016

Remarque

Les comptes rendus détaillés des débats de l'Autorité de la statistique publique ne sont pas publics.

Le texte ci-dessous expose toutefois les sujets abordés et les décisions ou avis qui en sont éventuellement issus.

**La séance est ouverte à 14h30
sous la présidence de M. Dominique Bureau**

I) Questions d'actualité

1) Avis du Président

En tant que Président de l'ASP, Dominique Bureau peut être amené à rendre des avis ou être saisi sur différents sujets. Le 14 avril 2016, il a ainsi rendu un avis favorable au renouvellement de Monsieur Stefan Lollivier, Inspecteur général de l'Insee dans ses fonctions de président du conseil d'orientation de l'ONDRP.

Il a également été saisi par le Défenseur des droits suite à la plainte d'une entreprise contre la charge occasionnée par les enquêtes statistiques.

2) Rapport 2015 de l'ASP

Le Président, accompagné du rapporteur de l'ASP, a présenté le rapport d'activité 2015 à plusieurs conseillers, notamment le conseiller économique du cabinet du ministre de l'Économie, de l'Industrie et du Numérique, le conseiller pour les questions économiques, financières et budgétaires du ministre de la Défense, le conseiller construction, urbanisme et habitat durable de la ministre du Logement et de l'Habitat durable.

3) Révision du règlement intérieur de l'ASP

La principale modification du règlement intérieur de l'ASP porte sur l'introduction d'un nouvel article visant à permettre à l'ASP d'émettre un avis à l'occasion des nominations du directeur général de l'Insee, chef de l'institut national de la statistique, et des responsables des services statistiques ministériels qui sont directeurs d'administration centrale. Les avis correspondants portent à la fois sur le contexte de la nomination (le cas échéant de la fin de fonction) et sur les compétences professionnelles dans le domaine de la statistique de la ou des personnes envisagées. Ils sont émis exclusivement au regard des critères concourant à l'indépendance professionnelle, au sens du 1° du principe du code de bonnes pratiques de la statistique européenne, mentionné dans le règlement européen (CE) n° 223/2009. Ils sont transmis au comité d'audition pour la nomination du directeur d'administration centrale concerné.

Il serait souhaitable que ces nouvelles dispositions soient consolidées dans le décret n° 2009-250 du 3 mars 2009 relatif à l'Autorité de la statistique publique.

4) Rupture d'embargo sur la note de conjoncture

Le Président de l'ASP signale une rupture d'embargo sur la note de conjoncture de juin 2016. M. Eckert, secrétaire d'État chargé du Budget au Ministère des Finances et des Comptes publics, a annoncé le 15 juin à la commission des finances du Sénat que « *l'Insee allait revoir sa prévision de croissance à 1,6 % en 2016* », chiffre de la note de conjoncture devant paraître le lendemain à 18 heures et dont il avait été destinataire sous embargo.

Au terme d'un débat, les membres de l'Autorité adoptent à l'unanimité la délibération suivante :

La note de conjoncture de l'Insee est un des produits-phares de l'Insee. Elle est toujours très attendue et relayée par les médias. La note exploite la production statistique de l'Insee (comptes nationaux, enquêtes de conjoncture, indice des prix à la consommation, emploi, chômage, etc.) pour fournir une analyse de la situation et des perspectives à court terme de l'économie française. Cette publication se situe donc dans le prolongement direct de la production statistique de l'Insee.

La note de conjoncture fait l'objet d'un cadre de diffusion précis¹ (calendrier pré-annoncé, règles d'embargo), porté à la connaissance du public, qui vise à respecter les principes de neutralité et d'équité de traitement des utilisateurs, que préconise le Code de bonnes pratiques de la statistique européenne, notamment ses principes 1 et 6 qui font de l'indépendance professionnelle, de l'impartialité et de l'objectivité des principes fondamentaux de l'organisation du système statistique.

L'Autorité de la Statistique publique constate que certaines de ces règles ont été méconnues dans le cas de la Note de conjoncture de juin 2016, qui a fait l'objet d'une rupture d'embargo². De plus, la levée de cet embargo pour rétablir l'égalité d'accès à l'information a été faite avec délai.

¹ Cf « Règles de diffusion des indicateurs conjoncturels et des publications » et note sur « les relations de l'Insee avec la presse », sur www.insee.fr

² Référence à ses prévisions de croissance par le Secrétaire d'Etat auprès du ministre des Finances et des Comptes publics, chargé du Budget, devant la Commission des Finances du Sénat, le 15 juin 2016.

L'Autorité juge regrettable ce non-respect des règles de diffusion s'appliquant à la note de conjoncture. En effet, celle-ci doit scrupuleusement obéir aux principes et normes édictés, pour assurer la crédibilité de l'information produite par l'Insee aux yeux des utilisateurs de la statistique publique, l'indépendance et la transparence de notre institut national de statistique constituant le socle de confiance sur lequel est fondée la statistique publique. Il convient donc de rétablir les conditions pour que ce type de situation ne puisse se reproduire.

II) Suivi du SSM de la DGFIP dans le cadre de la clause de revoyure émise par l'ASP le 18 juin 2014 : Audition de Monsieur Audran Le BARON, chef du service de la gestion fiscale à la DGFIP accompagné de Monsieur Brice Le PETIT, chef du SSM DGFIP

Le Président de l'ASP rappelle que l'ASP avait reconnu en septembre 2011 que le bureau de la DGFIP possédait la qualité de service statistique ministériel, assortissant son avis à un certain nombre de conditions. L'Autorité avait alors mis l'accent sur la question d'accès aux données fiscales et la bonne compréhension de la fiscalité par le public. Elle a pu constater des progrès en 2014, à l'occasion d'une première réunion de revoyure, notamment dans la mise à disposition de données fiscales aux chercheurs. Néanmoins, l'autorité avait demandé à auditionner de nouveau le SSM afin de pouvoir constater les autres progrès annoncés, notamment en matière de mise en ligne de données plus nombreuses et plus récentes au grand public. C'est la raison de cette nouvelle revoyure en 2016.

1) Présentation du SSM

M. LE PETIT, chef du bureau GF3C, précise que la DGFIP est issue de la fusion en 2008 de la Direction générale des impôts et de la Direction générale de la Comptabilité publique. Le bureau compte 40 agents, dont 5 agents de l'Insee. Il comporte trois sections métier organisées par grandes familles d'impôts (particuliers, professionnels, fiscalité locale) et deux sections plus transverses pour la diffusion et la prévision ainsi qu'une cellule informatique. Le bureau possède en effet un serveur dédié à ses propres besoins dont il assure la maintenance.

Le bureau a pour mission d'opérer pour le compte de la DGFIP et du service statistique public l'acquisition de données administratives en lien direct avec la gestion de l'impôt. Il assure également la diffusion de ces données brutes ou retravaillées légèrement à différents organismes (service statistique public, chercheurs, Parlement, corps d'inspection, etc.) et contribue à l'aide à la diffusion en matière fiscale par le biais de chiffrages, prévisions, suivis d'exécution de recettes budgétaires et de dépenses.

M. LE BARON souligne que l'acquisition de données représente une partie très substantielle du travail du service, même si elle est moins visible que ses autres missions. Elle constitue le noyau de l'activité des sections métier. Ses travaux sont fondés sur sa capacité à mobiliser des données administratives en matière fiscale qui présentent une grande qualité et un niveau de détail suffisant pour être travaillées, et qui suivent aussi l'actualité législative et réglementaire de la matière fiscale, une matière particulièrement vivante. C'est la raison pour laquelle il importe que le bureau soit rattaché au service de la gestion fiscale, puisqu'il est ainsi

associé de façon très directe à l'évolution des applicatifs, des formulaires déclaratifs et de tous les moyens mobilisables pour l'acquisition de données utiles.

Le service mobilise des outils statistiques pour faire en sorte que les bases de données soient complètes et mobilisables facilement. Il est amené à utiliser des données de volumétrie importante, notamment les données exhaustives d'impôts sur le revenu et doit donc mobiliser des méthodes statistiques pour produire des échantillons légers (50 000 foyers) ou plus lourds (500 000) pour son compte propre ou pour d'autres organismes d'administration. Enfin, il recourt à ces méthodes statistiques pour des travaux assez classiques de modélisation, d'étalonnage, etc.

Le service contribue assez fortement au service statistique public par le biais d'une activité conventionnelle très riche. Des conventions ont été conclues avec de nombreux services statistiques pour organiser l'échange de données, que ce soit l'Insee, la Drees, le SOeS, le ministère de l'Enseignement supérieur et de la Recherche, etc. Cette activité se révèle aussi très dynamique, puisque de nouvelles conventions sont signées chaque année, au gré des demandes que le service reçoit et de sa capacité à acquérir de nouvelles données susceptibles d'être mobilisées pour les travaux de ses partenaires.

S'agissant des perspectives, le service souhaite poursuivre sa contribution déjà significative au service statistique public par la production d'indicateurs divers. De nouveaux fichiers seront mis à la disposition de l'Insee (cessions immobilières, droits de mutation à titre onéreux, comptes bancaires) pour enrichir cette collaboration. Des chantiers sont également en cours avec l'Insee et d'autres SSM. Le ministère de la Défense souhaite décliner pour ses sous-traitants l'indicateur de chiffre d'affaires. Des travaux sont développés sur le crédit d'impôt recherche pour dupliquer l'expérience menée récemment sur le CICE. Enfin, le service et l'Insee étudient la façon dont les données de déclaration préremplie d'impôt sur le revenu peuvent être utilisées pour construire un indicateur avancé de pauvreté.

En termes de diffusion, le service s'efforce d'enrichir l'offre des données mises à disposition des chercheurs au niveau du CASD. La précédente audition s'était tenue quelques semaines avant la parution des textes réglementaires qui encadreraient l'ouverture des données fiscales aux chercheurs. Deux ans plus tard, l'appropriation de la procédure par les chercheurs comme par l'administration fiscale est en bonne voie. Le service a reçu 250 demandes de projets de recherche sollicitant des données fiscales et démontre sa capacité à mettre à disposition des données de qualité.

L'axe de diffusion privilégié vis-à-vis du grand public reste le site www.impots.gouv.fr. Le service s'efforce de publier des données fraîches. Il a mis en place, pour la première fois cette année, un calendrier de diffusion et s'efforcera d'annoncer la disponibilité des données au fil de l'eau. Enfin, le service enrichira le portail de diffusion des données de fiscalité locale, avec le REI, le fichier exhaustif des éléments de base d'assiette et de taux de cette fiscalité locale.

2) Délibération

Au terme d'un débat entre les membres du collège, la délibération suivante est adoptée à l'unanimité :

Suite à l'audition du chef de service de la gestion fiscale et du chef du service statistique ministériel de la DGFIP, et en référence aux recommandations faites par l'Autorité lors de sa

séance du 18 juin 2014, l'Autorité de la statistique publique constate que le bureau GFC3 de la DGFIP a désormais mis en place les instruments et cadres d'action (programme de travail, calendrier de publication, réduction des délais de production, etc.) associés au statut de service statistique ministériel. Elle observe que ce SSM se trouve en première ligne du chantier visant à élargir l'accès et garantir la qualité des données administratives de nature fiscale pour les rendre adaptées à l'utilisation statistique.

L'Autorité a constaté les progrès réalisés dans la mise à disposition des bases de données individuelles que gère la DGFIP aux chercheurs. Elle invite le SSM à poursuivre dans cette voie.

Compte tenu de l'ampleur du débat fiscal et des attentes du public en ce domaine, elle recommande à ce service de bâtir un projet d'ensemble cohérent et efficace en matière de diffusion et d'accès des données fiscales au profit du public, afin de rendre les données plus transparentes et en faciliter la compréhension.

Suite aux auditions de nombreux SSM à la demande de l'ASP, le **Président de l'ASP** propose de formuler une recommandation plus générale afin d'appuyer le processus d'homogénéisation des SSM. :

Cette recommandation est la suivante :

L'Autorité de la Statistique publique observe que la qualification de service statistique ministériel répond à un certain nombre de prérequis désormais largement partagés (programme de travail, calendrier de publication, etc.).

Elle insiste cependant sur la nécessité de former les agents des SSM au code de bonnes pratiques ainsi qu'à la charte des chefs de SSM.

III) Présentation de la loi Santé par Monsieur Franck VON LENNEP, directeur de la DREES, Monsieur Jean-Louis LHERITIER, chef de service, adjoint au directeur de la DREES, et Madame Chantal CASES, présidente de l'Institut des données de santé

Le **Président de l'ASP** rappelle que l'audition de l'Insee a démontré l'intérêt d'aborder le sujet du big-data de façon plus prospective, pour appréhender la façon dont l'institut peut modifier l'offre du système statistique. À cet égard, la santé constitue l'un des secteurs les plus sensibles.

M. VON LENNEP indique qu'il existe des bases médico-administratives très riches, encore peu utilisées bien qu'elles offrent des opportunités pour la recherche, mais aussi pour la statistique. La plus grande partie des publications sur le big data cite la santé comme l'un des domaines d'application principaux, avec des potentialités importantes, mais aussi des risques. Les données de santé constituent effectivement des données sensibles et forment donc un environnement très délicat.

3) La loi de modernisation de notre système de santé – Article 193

L'article 193 de la loi de modernisation de notre système de santé a été élaboré par le ministère de la Santé à la demande de nombreux acteurs publics et privés déplorant un trop faible accès aux données médico-administratives. La ministre de la Santé a instauré une commission de concertation qui a élaboré ce texte sur lequel un certain consensus s'est installé au Parlement, dans le sens d'une ouverture des données dans un cadre raisonné, en assurant la protection de la vie privée.

Cet article vient modifier le code de la santé publique pour créer un système national des données de santé à partir des données de l'assurance maladie (SNIIRAM), des données hospitalières du PMSI (séjours hospitaliers avec durée de séjour et codes diagnostiques), des causes de décès (INSERM), voire à l'avenir, lorsque cela sera possible, des données issues des Maisons départementales des personnes handicapées (MDPH) ou des organismes complémentaires. Ces bases de données n'étaient pas considérées dans le code de la santé publique comme des outils de recherche. Le texte marque donc une avancée sur ce point.

Le texte définit ensuite les règles d'accès en promouvant l'open data et une gouvernance générale : l'État pilote le système d'un point de vue stratégique, l'assurance maladie le met en œuvre sur le plan opérationnel et un institut national des données de santé, dont Chantal Cases est la préfiguratrice, représente les utilisateurs. La CNIL conserve son rôle d'autorisation des traitements.

Le texte vient par ailleurs modifier la loi Informatique et libertés sur les chapitres IX et X en unifiant autant que possible les règles d'accès aux données de santé à des fins de recherche et d'étude. Il supprime le décret en Conseil d'État nécessaire aux chercheurs pour les appariements nécessitant le NIR en vue d'une recherche ou étude en santé. Enfin, l'article élargit la possibilité d'établir des procédures CNIL allégées lorsque les demandes s'avèrent récurrentes, identiques, ou portent sur des données faiblement identifiantes. Cette modification devrait permettre de réduire les délais d'instruction de la CNIL qui dépassent aujourd'hui dix mois, voire un an. Elle répond à une demande à la fois des utilisateurs et de la CNIL.

La richesse de ces bases tient au fait que les données sont exhaustives et présentent, grâce au NIR, un chaînage de qualité entre les différentes bases et dans le temps. Elle tient aussi aux possibilités d'utilisation de ces données dans le cadre d'appariements avec différents registres, enquêtes, cohortes, etc. Ces données sont riches en elles-mêmes, mais elles le sont encore plus lorsqu'elles sont appariées avec d'autres.

Dérogeant à la loi Informatique et libertés, l'article 193 redéfinit les conditions d'accès en introduisant un droit d'accès permanent au profit d'organismes exerçant des missions de service public (ministère de la Santé, agences de santé, assurance maladie, INSERM et quelques autres acteurs). Un décret en Conseil d'État, en cours d'élaboration, définira ces accès.

En dehors de ces accès permanents, la procédure CNIL est revue. Le sérieux de l'étude et de ses finalités sera étudié par un comité d'experts, version élargie du comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS). L'intérêt public sera examiné quant à lui par l'Institut national des données de santé, dont le comité d'éthique est en cours de composition. La loi élargit les possibilités d'accès aux données pour des acteurs privés. Il a donc semblé nécessaire d'encadrer cet accès pour éviter que des données soient communiquées à des acteurs dont les intérêts sont discutables. La loi prévoit également des limitations dans l'accès des industriels de santé et des assureurs, qui devront passer par des intermédiaires pour obtenir les données. Enfin, un droit d'opposition des personnes est introduit

pour toutes les études qui ne font pas l'objet d'un droit d'accès permanent. Ce droit n'est pas aisé à mettre en œuvre, puisque la base comporte des données non directement identifiantes.

Enfin, les accès aux données à caractère personnel vont aussi donner lieu à des procédures simplifiées auprès de la CNIL et à des procédures spécifiques, prévues par la loi, pour les données à faible risque de réidentification. La procédure conduisant l'Institut national des données de santé à accorder cet accès en lieu et place de la CNIL est en cours d'élaboration. Il s'agit là encore de décharger la CNIL et de réduire les délais d'instruction.

En synthèse, cet article renforce les contraintes et la sécurité de l'accès aux données. Il met fin à la diffusion du PMSI sur cédérom qui interdisait toute traçabilité des usages. Même si les utilisateurs considèrent que les contraintes restent fortes, la sécurité permet de construire de la confiance sur ces données qui concernent l'ensemble des Français. Autre point de progrès, le texte met fin au décret en Conseil d'État et clarifie la procédure suivie par la CNIL. Les procédures simplifiées et l'articulation avec les comités de protection des personnes (procédures de la loi Jardé) seront appréhendées à l'usage. Des problèmes de frontière pourraient se poser avec des enquêtes qui peuvent aborder des questions de comportement, de santé mentale, etc. Une jurisprudence se construira progressivement sur le sujet.

Des travaux en cours concernent deux décrets d'application et la convention constitutive de l'Institut national des données de santé qui sera créé sous la forme d'un GIP (groupement d'intérêt public). La Caisse d'assurance maladie travaille à la mise en œuvre du système national des données de santé (SNDS) et du nouveau référentiel de sécurité, plus contraignant que l'existant. Collectivement, les services s'attachent à améliorer la documentation des données pour rendre l'utilisation de ces données plus facile pour les chercheurs français et étrangers. Le dispositif devrait entrer en vigueur au premier trimestre 2017.

En parallèle sont menés des travaux sur les données à faible risque de réidentification et sur des appariements entre le SNDS et des enquêtes ou l'échantillon démographique permanent. Ces données médico-administratives permettent d'étudier certains phénomènes, mais aussi d'enrichir les cohortes lorsqu'elles sont appariées à d'autres données administratives, de fiabiliser les questions sur le recours aux soins dans les enquêtes et d'alléger les questionnaires sur ces volets. Elles ne remettent pas pour autant en cause le besoin d'enquêtes. La Drees conduira sa prochaine enquête Santé en population générale en 2019. L'appariement entre ces enquêtes et les données médico-administratives se révèle très riche. Il a été engagé voilà quelques années, mais il n'a pas encore été poussé au bout. Il est devenu un objectif prioritaire désormais.

4) [Le big data en santé](#)

La ministre de la Santé a missionné des groupes de travail associant chercheurs et représentants du secteur privé fin 2015. Leurs travaux donneront lieu à un rapport qui paraîtra dans les prochains mois. Plusieurs centaines de contributions ont également été recueillies dans le cadre de la consultation faire-simple.gouv.fr. Un atelier citoyen a été lancé par le SGMAP en mai-juin pour appréhender les enjeux et les risques. Enfin, un colloque sera organisé le 4 juillet par la ministre et la Drees, qui a également mis en place un groupe de travail méthodologique en partenariat avec l'Insee pour explorer des méthodologies statistiques un peu nouvelles.

Un projet est en cours entre la CNAM et un laboratoire de l'École Polytechnique. La Drees souhaiterait par ailleurs travailler sur le PMSI pour tenter de définir des réseaux de soins dans la logique du parcours de soins, comme le font les États-Unis ou le Canada. Par ailleurs, la Drees

pourrait investir de nouvelles sources de données. L'Agence nationale de sécurité et du médicament (ANSM) a lancé des projets avec des chercheurs pour développer la surveillance sanitaire du médicament à partir des requêtes sur les moteurs de recherche. Les nouvelles sources de données (objets connectés, moteurs de recherche, réseaux sociaux) offrent de nombreux éléments liés à la santé et les acteurs privés investissent massivement dans ce domaine. Il serait donc dommage que le secteur public reste totalement à l'écart de ce mouvement.

Des questions se posent sur les finalités du big-data. Sept familles d'usages ont été identifiées à travers les groupes de travail en lien avec la CNIL. Une question très lourde se pose aussi autour du recueil du consentement, notamment si l'on veut ouvrir des données hospitalières et des dossiers médicaux et les croiser avec des données externes. La sécurité des données ou des traitements soulève également des interrogations. Les experts débattent notamment sur la nécessité d'éviter la constitution de grands entrepôts et de fonctionner avec des interfaces et des données non situées physiquement. Le comité national d'éthique a été saisi, car de nombreuses questions se posent en la matière sur les risques d'accroissement ou de réduction des inégalités, autour de la médecine prédictive et de la liberté pour le patient de ne pas avoir accès à certaines informations. Des questions autour de l'assurance se posent aussi. L'assurance repose sur la mutualisation. Or le big data renforce l'individualisation et pourrait donc favoriser la segmentation. Des éléments échappent au dispositif public, notamment l'assurance emprunteur. Le big data soulève enfin des questions récurrentes sur la qualité des données, l'opposition entre corrélation et causalité, la transparence des algorithmes, etc.

La séance est levée à 18 heures 30.

La prochaine séance de l'ASP aura lieu le 12 octobre.

Autorité de la statistique publique
Séance du 22 juin 2016
ANNEXE

ÉTAIENT PRESENTS

Monsieur Dominique BUREAU
Président

Monsieur Denis BADRE
Ancien Sénateur

Monsieur Abdeldjellil BOUZIDI
Economiste, Université Sorbonne Nouvelle Paris 3

Monsieur Bruno DURIEUX
Ancien ministre
Président du comité national des conseillers du commerce extérieur de la France

Monsieur Éric DUBOIS
Conseiller Maître à la Cour des comptes

Monsieur Jean GAEREMYNCK
Président du comité du secret statistique
Membre du Conseil d'État

Madame Anne-Marie BROCAS
Inspectrice générale des affaires sociales

Madame Patricia BLANCARD
Membre du Conseil économique social et environnemental

EXCUSÉS

Madame Véronique HESPEL
Inspectrice générale des finances

Monsieur Philippe CUNEO
Chef de l'Inspection générale de l'Insee

ASSISTAIENT EGALEMENT À LA SEANCE

Madame Claudine GASNIER
Rapporteur de l'Autorité de la statistique publique

Madame Béatrice GIMARD
Secrétaire de l'Autorité de la statistique publique