

N°s

472883 Ministre (INSEE)

472884 Ministre (INSEE)

472984 Ouvre boîte

10^{ème} et 9^{ème} chambres réunies

Séance du 17 mai 2024

Décision du 31 mai 2024

CONCLUSIONS

M. Laurent DOMINGO, Rapporteur public

L'association « Ouvre-boîte » poursuit son action d'ouverture des données publiques en s'intéressant aux data de l'INSEE et plus particulièrement celles permettant le calcul de l'indice des prix à la consommation (IPC).

Cet indice, publié mensuellement au Journal Officiel, est un instrument statistique qui permet de mesurer la variation moyenne des prix des produits de consommation des ménages, à partir d'un panier fixe de biens et de services, regroupant 1 600 familles de produits et actualisé chaque année. L'IPC rend ainsi compte de l'inflation des prix à la consommation et il sert de référence d'indexation dans de nombreux cas, dont le SMIC (article L. 3231-12 du code du travail).

Pour calculer l'IPC, l'INSEE se fonde sur les prix de vente, qu'il obtient par trois canaux : par des relevés dans des points de vente physiques, environ 30 000 répartis dans 99 agglomérations de plus de 2 000 habitants et en outre-mer ; par la collecte des prix de vente sur les sites de vente en ligne ; par transmission quotidienne des données de caisses des enseignes de la grande distribution (v. arrêté du 13 avril 2017 rendant obligatoire la transmission de données par voie électronique à des fins de statistique publique).

Ouvre-boîte a demandé, par deux fois, à l'INSEE de publier en ligne, sur le fondement du 4^o de l'article L. 311-9 du CRPA, 14 séries de documents administratifs comportant, selon l'association, les données nécessaires au calcul de l'IPC, que ce soit les données de base (produits et magasins), les données collectées (les prix) et les données de calcul (les pondérations par ex.)¹.

¹ 1) La liste des familles de produits et des produits qui les composent ; 2) La liste des 99 agglomérations et des 4 départements d'outre-mer dans lesquels sont effectués les relevés ; 3) La liste des 30.000 points de vente qui alimentent les relevés ; 4) Les données brutes de collecte ; 5) Les valeurs des 200.000 séries « produits précis dans un point de vente donné » et des 190.000 séries « tarif » collectée de façon centralisée ; 6) Les valeurs des 30.000 indices élémentaires (croisement variétés x agglomération) ; 7) La valeur des pondérations intervenant

A chacune de ces demandes, la CADA a rendu des avis (16 janvier 2020, 25 mars 2021) distinguant 3 groupes de documents : ceux couverts par le secret en matière statistique ; ceux pour lesquels la demande est imprécise ; ceux qui peuvent être communiqués et donc publiés en ligne.

Après sa deuxième demande, Ouvre-boîte a saisi le TA de Paris. Par un jugement du 9 février 2023, contre lequel tant le ministre de l'économie et des finances pour l'INSEE que l'association Ouvre-boîte ont formé un pourvoi, le tribunal :

- a prononcé un non-lieu à statuer sur les documents publiés par l'INSEE, à savoir la liste des 1 600 familles de produits et la liste des 99 agglomérations et des départements d'outre-mer enquêtés ;

- a jugé que les demandes de publication des documents n°s 10, 11, 13 et 14, qui visaient toutes autres sources de données ou tous autres documents utilisés pour le calcul de l'IPC, n'étaient pas suffisamment précises. Vous n'avez pas admis le pourvoi de l'association sur ces demandes (15 décembre 2023, n°472984) ;

- a jugé que le secret en matière statistique s'appliquait pour 5 séries de données (3, 4, 5, 8 et 9) : la liste des 30 000 magasins dans lesquels sont réalisés des relevés de prix, les produits précis qui font l'objet d'un relevé de prix en magasin ou en ligne, les prix collectés et les données issues des enquêtes « budget des familles » et « évaluation annuelle des dépenses de consommation des ménages ». Le pourvoi de l'association a été admis pour ces demandes (même décision).

- a jugé que la valeur des 30 000 indices élémentaires (6), la valeur des pondérations (7) et les codes sources utilisés (12) étaient communicables et devaient donc être publiés. L'Etat se pourvoit (avec une demande de sursis à exécution) en ce qui concerne les 30 000 indices élémentaires. L'INSEE a publié la valeur des pondérations et a annoncé qu'il publiera les codes sources.

dans le calcul de l'IPC ; 8) Les données extraites de l'enquête « Budget de famille » utilisées pour le calcul de l'IPC, anonymisées si besoin ; 9) Les données extraites des évaluations annuelles des dépenses de consommation des ménages réalisées par la Comptabilité Nationale utilisées pour le calcul de l'IPC (consommation finale effective des ménages restreinte au champ des biens et services faisant l'objet d'une production marchande) ; 10) L'ensemble des autres sources de données spécialisées intervenant dans le calcul de l'IPC ; 11) Toute la documentation utilisée par les économistes de l'INSEE pour le calcul de l'IPC et qui ne fait pas déjà l'objet d'une publication en ligne, cette documentation comprenant les traitements spécifiques (variétés saisonnières, évolution de prix pures, etc.) ; 12) L'ensemble des codes sources utilisés par les économistes de l'INSEE pour le calcul de l'IPC et qui comprend notamment les traitements spécifiques (variétés saisonnières, évolution de prix pures, etc.) ; 13) Tous les autres documents (code source, documentation, données, etc.) nécessaires à la reproductibilité des calculs de la 8ème génération de l'IPC, dont les résultats sont publiés mensuellement au Journal Officiel depuis janvier 2018 (IPC de l'ensemble des ménages, IPC hors tabac de l'ensemble des ménages, IPC hors tabac des ménages urbains dont le chef est ouvrier ou employé, IPC hors tabac des ménages du premier quintile de la distribution des niveaux de vie) ; 14) Tous les documents (code source, documentation, données, etc.) nécessaires au calcul de l'indice des prix à la consommation harmonisé (ICPH) depuis janvier 2018.

Si ce n'est le moyen tiré de ce que le tribunal a omis, dans le dispositif de son jugement, de rejeter le surplus des conclusions de l'association, l'essentiel du débat, quel que soit le pourvoi, porte sur le champ et les conditions d'application du secret en matière statistique.

I. Le secret en matière statistique est un secret protégé par la loi qui fait obstacle à la communication d'un document administratif (h du 2° de l'article L. 311-5). Ce secret est consacré par la loi n° 51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques.

On le trouve à l'article 6 qui protège les renseignements individuels ayant trait à la vie personnelle et familiale et, d'une manière générale, aux faits et comportements d'ordre privé et les renseignements individuels d'ordre économique ou financier figurant dans les questionnaires revêtus du visa prévu à l'article 2 de la loi : sauf dérogations, les premiers renseignements ne peuvent faire l'objet d'aucune communication de la part du service dépositaire avant l'expiration d'un délai de soixante-quinze ans suivant la date de réalisation de l'enquête ou d'un délai de vingt-cinq ans à compter de la date du décès de l'intéressé, si ce dernier délai est plus bref ; et les seconds d'aucune communication de la part du service dépositaire avant l'expiration d'un délai de vingt-cinq ans suivant la date de réalisation du recensement ou de l'enquête.

Ce secret est également prévu à l'article 7 bis de la loi de 1951, qui concerne les informations relatives aux personnes physiques et aux personnes morales, recueillies dans le cadre de sa mission, par une administration, une personne morale de droit public, ou une personne morale de droit privé gérant un service public et qui sont cédées, à des fins exclusives d'établissement de statistiques, à l'Institut national de la statistique et des études économiques ou aux services statistiques ministériels.

Au niveau européen, le règlement n° 223/2009 du 11 mars 2009 relatif aux statistiques européennes, qui sont produites dans le cadre du système statistique européen (SSE) qui est un partenariat entre Eurostat et les instituts nationaux de statistique (INS)², consacre également ce principe de secret statistique. Il signifie (article 2) que les données confidentielles relatives à des unités statistiques individuelles qui sont obtenues directement à des fins statistiques ou indirectement à partir de sources administratives ou autres doivent être protégées, et cela implique que l'utilisation à des fins non statistiques des données obtenues et la divulgation illicite de ces dernières soient interdites.

Ce principe assure la protection des données individualisées, qu'elles relèvent de la vie privée ou du domaine économique. Il participe, en protégeant les sources utilisées, de la fiabilité et donc de la qualité de la statistique.

II. Devant le tribunal, l'association Ouvre-Boîte soutenait que le secret statistique de l'article 6 de la loi de 1951 ne lui était pas opposable, car il ne vise que les renseignements figurant dans les questionnaires revêtus du visa prévu à l'article 2 de la loi, c'est-à-dire le visa

² Il existe au niveau européen un indice des prix à la consommation harmonisé (IPCH).

ministériel : or, selon l'association, l'IPC ne serait pas visé par le ministre et ne procéderait pas de questionnaires.

Bizarrement, pour les documents dont il a jugé qu'ils ne pouvaient être communiqués, le tribunal a écarté le moyen en retenant que le secret statistique ne s'applique pas uniquement aux enquêtes statistiques soumises au visa préalable du ministre, mais sans que l'on ne comprenne ce que le tribunal a entendu juger, car s'il est vrai que le secret statistique concerne aussi les données transférées par les administrations (article 7 bis), ce n'est cependant pas le sujet du litige.

De son côté, l'INSEE acquiesce à la réponse du tribunal car, se référant méticuleusement aux travaux parlementaires de la loi du 7 juin 1951, il considère que le secret de l'article 6 s'applique à toutes les données recueillies à des fins statistiques et non aux seules données recueillies dans le cadre d'enquêtes revêtues du visa ministériel. Cependant, la lettre de l'article 6 est claire et le visa ministériel est une condition d'application du secret statistique. Pour les autres statistiques, non visées par le ministre, ce sont les exceptions tirées de la protection de la vie privée ou du secret des affaires prévues par le CRPA (L. 311-6) qui s'appliqueront.

Ce visa ministériel, prévu à l'article 2 de la loi, est requis pour les enquêtes statistiques des services publics et il est accordé si l'enquête s'inscrit dans le cadre du programme des travaux statistiques³. Il se trouve que l'IPC fait partie du programme d'enquêtes statistiques et il est donc revêtu du visa ministériel⁴.

La réponse du tribunal est donc au moins insuffisamment motivée, certainement entachée d'une erreur de droit, mais il n'en demeure pas moins que le secret statistique s'applique bien à l'IPC en vertu de l'article 6 de la loi de 1951. Au besoin, vous pourrez substituer ce motif à celui retenu par le tribunal.

Car, contrairement à ce que soutient aussi l'association Ouvre-Boîte, il n'y a pas lieu de conférer au terme « questionnaire » figurant à l'article 6 une portée restrictive, qui le limiterait aux recueils d'informations sous forme de questions, d'autant plus que cet article 6 mentionne aussi, au titre des dérogations, les recensements et les enquêtes.

³ Une enquête statistique des services publics visée par le ministre fait d'abord l'objet d'un avis d'opportunité favorable d'un président d'une commission thématique du Conseil national de l'information statistique puis d'un avis de conformité favorable de la part du comité du label de la statistique publique. Elle est alors visée par le ministre. Elle est ensuite inscrite par Conseil national de l'information statistique au programme annuel d'enquêtes des services producteurs de la statistique publique, publié par arrêté ministériel (v. décret n° 2009-318 du 20 mars 2009 relatif au Conseil national de l'information statistique, au comité du secret statistique et au comité du label de la statistique publique).

⁴ Pour 2023, v. arrêté du 24 octobre 2022 portant approbation du programme d'enquêtes statistiques d'initiative nationale et régionale des services publics pour 2023 (enquêtes auprès des entreprises et des exploitations agricoles) ; pour 2024, v. arrêté du 7 novembre 2023 portant approbation du programme d'enquêtes statistiques d'initiative nationale et régionale des services publics pour 2024 (enquêtes auprès des entreprises et des exploitations agricoles).

Nous comprenons l'article 6 comme protégeant des « renseignements », on dirait aujourd'hui des données, collectés par l'INSEE, sur place ou à distance, et utilisés pour produire des statistiques dès lors qu'ils touchent à la vie privée ou qu'ils sont d'ordre économique ou financier, ce qui vise, pour employer le vocable européen, « l'unité statistique » qui peut être identifiée directement ou indirectement⁵.

III. Sur cette partie du jugement concernant les documents non-communicables, le tribunal a ensuite ajouté qu'il n'est pas sérieusement contesté que les documents demandés comportent des renseignements individuels ayant trait à des faits, des comportements d'ordre privé et des renseignements individuels d'ordre économique ou financiers. Cette appréciation n'est pas contestée en cassation.

Elle n'est au demeurant pas contestable pour les trois premiers documents : la liste des magasins, la liste des produits, les prix relevés ou collectés. Ce sont les données de base de la statistique. Ce sont des informations individualisées qui sont à la source des données économiques utilisées par l'INSEE et qui ne peuvent être révélées sans remettre en cause la fiabilité même de l'enquête, compte-tenu du risque de manipulation des prix qu'une telle divulgation permettrait⁶.

En revanche, les pièces du dossier ne permettent pas de saisir ce que recouvre exactement les données, telles qu'elles sont utilisées pour l'IPC, issues des enquêtes « budget des familles » et « évaluation annuelle des dépenses de consommation des ménages ». Nous comprenons seulement que ces données servent à déterminer les familles de produits choisies chaque année.

Or, l'association Ouvre-Boîte soutenait devant le tribunal et s'en plaint en cassation par des moyens d'insuffisance de motivation et d'erreur de droit, que quand bien même les documents dont elle demandait la publication seraient couverts par le secret statistique, ils pourraient être publiés après occultation des données relevant de ce secret, comme il est prévu à l'article L. 311-7.

Le tribunal n'a pas répondu à ce moyen. Cette omission est sans conséquence s'agissant de la liste des magasins, la liste des produits et les prix, car ces listes ne sont pas, en totalité, communicables, si bien qu'une demande tendant à leur occultation partielle est dépourvue d'objet, elle est inopérante. En revanche, s'agissant des données issues des enquêtes « budget des familles » et « évaluation annuelle des dépenses de consommation des ménages », cette

⁵ Le règlement précité de 2009 indique (articles 2 et 3) que le secret statistique porte sur les données confidentielles relatives à des unités statistiques individuelles, c'est-à-dire des données permettant l'identification, directe ou indirecte, d'une personne physique, d'un ménage, d'un opérateur économique ou d'une autre entreprise. Il précise que pour déterminer si une unité statistique est identifiable, il est tenu compte de tous les moyens appropriés qui pourraient raisonnablement être utilisés par un tiers pour identifier l'unité statistique.

⁶ Pour certains produits, qui représentent la plus grande majorité des ventes (par ex. la pâte à tartiner à la noisette), la publication de la liste des familles de produits peut donner une indication sur le produit, mais cette situation, qui est propre à un état du marché, est marginale.

omission nous apparaît devoir être censurée, faute pour le tribunal d'avoir vérifié qu'un tri des données était techniquement possible sans faire peser une charge déraisonnable sur l'INSEE.

Pour ces différents motifs, nous vous proposons d'annuler (quand bien même le tribunal n'a pas rejeté le surplus des conclusions qui doit donner lieu à l'annulation partielle proposée) le jugement attaqué en tant qu'il se prononce sur les documents 8 et 9.

IV. Pour ce qui est des documents dont le tribunal a jugé qu'ils devaient être publiés, seul le document 6 est en débat en cassation.

Il s'agit de la valeur des 30 000 (environ) indices élémentaires. Ces indices élémentaires sont la valeur des prix d'une variété de produits au niveau des agglomérations. Par ex., les prix de la moutarde, de la mayonnaise, de la sauce de soja, des cornichons, etc., à Montpellier, Perpignan, Narbonne, etc.

L'INSEE a refusé de publier cette liste de données, au motif qu'il n'est pas en mesure, faute de disposer de toutes les informations requises et des moyens nécessaires, de s'assurer qu'il est exclu que, par recoupement, à partir d'informations dont des tiers pourraient disposer, des informations couvertes par le secret statistique, que sont les points de vente enquêtés, les produits précis et les prix pratiqués, avec leur évolution, ne puissent être reconstituées.

Le risque de reconstitution est probablement nul à Paris ou Marseille ou dans toutes les autres grandes agglomérations, mais l'INSEE se préoccupait des plus petites agglomérations, de quelques milliers d'habitants, et il y en a dans sa liste des agglomérations ciblées pour l'IPC, où, selon la nature de la variété, peu courante, il serait aisé de déterminer la ou les enseignes enquêtées (ainsi que, par déduction l'une de l'autre, les prix pratiqués).

Le tribunal a sèchement répondu que l'INSEE n'apporte aucun commencement de preuve à l'appui de ces allégations alors qu'il est constant qu'il publie 257 indices élémentaires par mois.

Cette réponse fait l'objet de moyens d'IM, d'ED et de dénaturation, notamment. Ils nous apparaissent fondés.

Il est exact que l'INSEE n'a pas donné d'exemples précis et expliqué qu'en publiant la liste des 30 000 indices élémentaires, il serait possible, voire aisé, de savoir quelles enseignes sont enquêtées, voire quel est le niveau des prix pratiqués par une enseigne en particulier, à partir de la vente d'essence, d'aspirateurs ou de pizzas surgelées, à Langeac en Haute-Loire (3 700 habitants), à la Mure en Isère (5 000 habitants) ou à l'Ille-sur-Têt dans les Pyrénées-Orientales (5 500 habitants).

Mais l'argumentation de l'INSEE se plaçait sur plan méthodologique, celui du risque d'identification de données unitaires, sa probabilité dans certaines agglomérations, et son incapacité à déterminer exactement dans lesquelles sans déployer des moyens humains (et financiers) pour les seuls besoins de la publication, et sachant que les données utilisées (variétés et agglomérations) évoluent. Il appartenait au tribunal de se prononcer, dans le cadre

du caractère objectif de son office en matière de communication d'un document, sur la réalité de ce risque.

En outre, l'argument selon lequel l'INSEE publie déjà 257 indices élémentaires ne vaut pas. Le tribunal a peut-être cru que si l'INSEE en publie 257, il peut en publier 30 000 de plus. Mais précisément, les 257 indices publiés se situent à un niveau d'agrégation supérieur, qui est celui des sous-classes de la classification européenne des fonctions de consommation, qui ne permet pas la réidentification d'informations confidentielles.

Et c'est pour cette raison aussi, tirée de l'absence de risque de réidentification, que l'INSEE ne conteste pas en cassation l'injonction de publier la liste des pondérations par produits et par agglomérations utilisés dans le calcul de l'IPC et qu'elle a publié cette liste⁷ (l'IPC donnant par ailleurs lieu à la publication de nombreux jeux de données⁸).

Nous vous proposons donc d'annuler le jugement également en tant qu'il porte sur le document 6. Et compte-tenu des explications données par l'INSEE sur cette série de données et son incapacité à les trier précisément en fonction du risque d'identification de données unitaires, vous pourrez vous-même, sans renvoyer l'affaire au tribunal sur ce point, rejeter la demande de publication en ligne.

PCMNC :

- A l'annulation du jugement en tant qu'il se prononce sur le document 6 (pourvoi de l'Etat) et au rejet de la demande de publication de ce document (30 000 indices élémentaires ;
- A ce qu'il n'y ait pas lieu de statuer sur la demande de sursis à exécution ;
- A l'annulation du jugement en tant qu'il se prononce sur les documents 8 et 9 (pourvoi de l'association Ouvre-Boîte) et au renvoi de l'affaire au TA dans cette mesure ;
- Au rejet du surplus des conclusions.

⁷ <https://www.insee.fr/fr/metadonnees/source/indicateur/p1653/processus-statistique>

⁸ v. <https://www.insee.fr/fr/statistiques/series/102342213>