

L'EVALUATION *EX ANTE* ET *EX POST* DES POLITIQUES SOCIALES : EST-ON DANS UNE IMPASSE ?

Mireille Elbaum

EDITORIAL BOARD

Chair: Xavier Ragot (Sciences Po, OFCE)

Members: Jérôme Creel (Sciences Po, OFCE), **Eric Heyer** (Sciences Po, OFCE), **Sarah Guillou** (Sciences Po, OFCE), **Xavier Timbeau** (Sciences Po, OFCE), **Anne Epaulard** (Sciences Po, OFCE).

CONTACT US

OFCE
10 place de Catalogne | 75014 Paris | France
Tél. +33 1 44 18 54 24
www.ofce.fr

WORKING PAPER CITATION

This Working Paper:
Mireille Elbaum,
L'évaluation *ex ante* et *ex post* des politiques sociales : est-on dans une impasse ?
Sciences Po OFCE Working Paper, n° 29/2025.
Downloaded from URL: www.ofce.sciences-po.fr/pdf/dtravail/WP2025-29.pdf
DOI - ISSN

ABOUT THE AUTHORS

Mireille Elbaum, Chercheuse affiliée à l'OFCE-SciencesPo, ancienne membre de l'Inspection générale des affaires sociales.

Email Address: elbaummireille@gmail.com

ABSTRACT

L'expérience française en matière d'évaluation des politiques sociales connaît actuellement des limites majeures. Un état des lieux des pratiques d'évaluation et de la façon elles sont mobilisées au service de l'action publique, au confluent du monde de la recherche et des administrations, permet de les mettre en évidence. En premier lieu, les politiques sociales sont fortement dépendantes d'incertitudes macroéconomiques de nature plurielle qui tendent à compliquer la cohérence entre prévisions à court, et moyen et long terme. Elles rendraient nécessaire un débat transparent et élargi autour de leurs hypothèses, qui, en dépit du développement des instruments de prévision et de micro-simulation et de l'existence d'instances comme les Hauts Conseils, est actuellement insuffisant. En termes d'évaluation ex-post, l'un des problèmes concerne le recours excessif aux « indicateurs de performance », parfois considérés comme « plus efficaces » que les démarches évaluatives. S'ils peuvent améliorer le suivi des politiques, leur prolifération a des effets pervers et ne peut se substituer à des évaluations approfondies, comportant une dimension qualitative. Par ailleurs, l'évaluation des politiques sociales est aussi confrontée en France à un manque de clarté sur ses objectifs et à une multiplicité d'acteurs porteurs d'approches disparates. En résultent des problèmes de coordination et de lisibilité, et souvent une faible association des acteurs sociaux. Enfin, les expérimentations sociales, censées éclairer la généralisation des politiques, se sont multipliées depuis 2003, mais leurs méthodes et leur impact restent discutables, sachant qu'elles sont loin de garantir la pertinence et la reproductibilité de leurs résultats à grande échelle, ce qui devrait conduire à repenser leur rôle. Du côté des travaux académiques, il faut certes souligner l'important renouveau des méthodes quantitatives d'évaluation fondées sur l'utilisation de données individuelles, notamment pour les politiques d'emploi et d'insertion. Ces méthodes nécessitent toutefois du temps et des analyses complémentaires, dont les exigences entrent en contradiction la forte accélération des débats politique et médiatique. Le risque est alors que les démarches évaluatives, originellement à visée démocratique, soient de fait remplacées par le recours à des indicateurs mal maîtrisés, retraités de façon plus ou moins transparente et jugés suffisants pour apprécier et orienter l'action publique.

KEYWORDS

Prévisions, évaluation, politiques sociales, indicateurs, expérimentations sociales, méthodes évaluatives.

JEL

C10, C93, D78, E61, H11, H53, H68, I38, J11.

L'évaluation *ex ante* et *ex post* des politiques sociales : est-on dans une impasse ?

Mireille ELBAUM¹

Alors que l'idée « d'évaluation » est aujourd'hui banalisée dans les discours sur les politiques publiques et présentée comme le moyen de fournir les clés de politiques « efficaces » et permettant une maîtrise optimale des finances publiques, la pratique actuelle de ces exercices semble, *ex ante* comme *ex post*, pour partie dans une impasse lorsque l'on examine leur cadre d'analyse, leur mode de conduite, leurs apports concrets et leur appropriation effective par les acteurs du débat social.

Il est clair que la plurisémié de la notion d'évaluation et la multiplicité des opérateurs qui en revendiquent la pratique participent à ces limites, mais on peut tenter de caractériser plus avant les manifestations et les causes des difficultés actuelles, qui posent une vraie question « d'utilité citoyenne ».

Les réflexions qui suivent, centrées sur les politiques sociales, ont un caractère personnel et reprennent en partie le texte d'une conférence prononcée à l'université Paris I en décembre 2024 dans le cadre du master « chargé(e)s d'études et de recherche économiques et sociales ²».

I) Une élaboration des politiques sociales peu ouverte, au-delà des « chiffrages budgétaires », à la prise en compte des évaluations précédentes ou en cours

L'expérience de l'élaboration de textes budgétaires et des réformes récentes ou en discussion offre un terrain d'observation des conditions d'élaboration des politiques sociales qui met en doute la place de l'évaluation dans ces processus, qu'il s'agisse des enseignements des évaluations *ex post*, ou de ceux d'études *ex ante* allant au-delà des seuls chiffrages budgétaires.

¹ Chercheuse affiliée à l'Observatoire français de conjonctures économiques, ancienne membre de l'Inspection générale des affaires sociales.

Adresse e-mail : elbaummireille@gmail.com

² Les constats formulés et les références citées dans ce document de travail sont antérieurs à septembre 2025 et destinés à être le cas échéant complétés dans une version ultérieure.

Les raisons de ces difficultés sont plurielles et tiennent autant à l'environnement économique incertain de ces travaux qu'à leurs cadres de partage et de discussion collective.

1.1) Un cadre macroéconomique brouillé dont la cohérence temporelle semble de plus en plus difficile à envisager

Les politiques sociales, qu'il s'agisse des mesures pour l'emploi, des prestations de solidarité, des retraites ou même des politiques de santé sont extrêmement sensibles au contexte démographique et macroéconomique, qui encadre à court (chômage, insertion) comme à long terme (santé, retraites) leur pertinence, leur efficacité et les conditions de leur financement. Les incertitudes sont aujourd'hui patentées en la matière, et les travaux conduits par les économistes semblent difficilement à même de les éclairer.

En termes conjoncturels, après les fluctuations induites depuis 2020 par l'épidémie de Covid puis la sortie de crise sanitaire, les perspectives économiques apparaissent de plus en plus liées à l'évolution de la situation politique internationale : conséquences de la guerre en Ukraine sur les approvisionnements, revirements de la politique américaine en termes de politique de défense et de droits de douane, incertitudes liées aux guerres au Moyen Orient, mais aussi et surtout, pour la France, les réactions à ce nouveau contexte de l'Union et des pays européens, en termes de politiques industrielle, douanière et budgétaire. À cet égard, et même si les outils de prévision à court terme se perfectionnent en continu grâce à l'accès des statisticiens à des données nouvelles (données de caisse, ...) et à des innovations techniques (*nowcasting*...), le cadrage économique à un an et plus qui préside à l'élaboration des textes budgétaires nécessiterait pour les éclairer la « mise sur table » de modèles de prévision et de simulation intégrés à l'échelle européenne, voire mondiale (pour ce qui est de l'impact des tarifs douaniers³).

À moyen-long terme, comme le montre la succession des rapports du Conseil d'orientation des retraites (COR)⁴, l'une des deux variables clés pour les projections relatives aux politiques sociales est en outre la productivité du travail, qui concentre également débats et incertitudes. Le lien entre les dépenses de retraite, et plus largement de protection sociale, rapportées au PIB et la productivité résulte en effet du mode d'indexation sur les prix adopté depuis la réforme Balladur de 1993, qui a définitivement désindexé les retraites du revenu

³ Voir par exemple Antoine Bouët et Vincent Vicard, « *Liberation Day ? Une première estimation des effets des Reciprocal tariffs* », Présentation à la presse, avril 2025.

⁴ Le dernier en date actualise les hypothèses retenues en se fondant sur celles adoptées par la Cour des comptes lors de la « mission flash » effectuée à la demande du gouvernement. Voir Conseil d'orientation des retraites, *Évolutions et perspectives des retraites en France*, Rapport annuel, juin 2025 et Cour des comptes, *Situation financière et perspectives du système de retraites*, Communication au Premier ministre, février 2025.

moyen des actifs, dans le but d’opérer, en « période de croisière », une réduction du niveau relatif des retraites sans affichage explicite⁵.

Les questions portent d’abord, et peut-être avant tout, sur le contenu même de la notion de productivité et sa mesure au niveau national ou européen, dans un contexte de numérisation et de mondialisation accélérées des économies, jointes à « la plateformesisation » de certains services et à la domination des multinationales qui optimisent la « localisation fiscale » de leurs activités et les redevances facturées à leurs filiales. Les débats de politique économique portent quant à eux sur les causes du ralentissement des gains de productivité observé depuis les années 1980 et sur la capacité des pays européens à agir pour les surmonter : temps de latence dans la mise en œuvre des évolutions technologiques, freins institutionnels et humains à l’adoption des innovations, constitution de monopoles empêchant l’arrivée de concurrents parmi les entreprises qui diffusent les nouvelles technologies, « épuisement » de l’impact possible des dernières vagues technologies sur l’emploi, dont le niveau reste « incompressible » pour certains services de proximité⁶.

Les « nouvelles » incertitudes débattues aujourd’hui, qui accroissent les « fourchettes » des hypothèses à prendre en compte, concernent par ailleurs l’impact des politiques de transition écologique, qui, si elles étaient mises en œuvre, pourraient induire une chute significative des gains de productivité d’ici cinq à dix ans, suivie éventuellement d’une remontée à plus long terme⁷. De son côté, l’implantation à grande échelle des algorithmes issus de l’intelligence artificielle peut faire anticiper, en fonction des choix organisationnels et des processus de contrôle mis en place au sein des entreprises, des gains de productivité inédits concernant des emplois relevant du « tertiaire supérieur ». Leur ampleur potentielle est néanmoins très difficile à appréhender, d’autant que des complémentarités peuvent aussi être envisagées avec les emplois existants et, qu’à l’échelle de l’ensemble de l’économie, les destructions d’emploi peuvent être contrebalancées par le développement de nouvelles activités et de nouveaux métiers⁸.

⁵ Voir notamment Didier Blanchet, Mathieu Chabaud, Tom Olivia et Jean Rubin, « Quelles règles d’indexation des retraites pour diminuer la sensibilité du système à la croissance ? », *Insee Analyses*, n°109, avril 2025 et Didier Blanchet, Mathieu Chabaud, Tom Olivia et Jean Rubin, « Quels effets budgétaires et redistributifs des règles d’indexation des retraites ? », Documents de travail Insee, n°2025-08, avril 2025.

⁶ Voir notamment, sur le cas français, Philippe Askenazy, Émilie Cupillard, Guillaume Houriez, Yves Jauneau et Dorian Roucher, « À la recherche des gains de productivité perdus depuis la crise sanitaire », *Blog de l’Insee*, juillet 2024.

⁷ Voir notamment Jean Pisani-Ferry et Selma Mahfouz, *Les incidences économiques de l’action pour le climat*, Rapport à la Première ministre, mai 2023.

⁸ Voir par exemple Jan Hatzius, Joseph Briggs, Devesh Kodnani, “*The Potentially Large Effects of Artificial Intelligence on Economic Growth*” <https://www.gspublishing.com>, Goldman Sachs, mars 2023, Daron

À l'horizon d'une vingtaine d'années, les perspectives démographiques deviennent également déterminantes. À cet égard, l'accentuation récente de la baisse de la natalité, après le rebond de 2021, pourrait amorcer un tournant structurel rapprochant progressivement la France des comportements de fécondité d'autres nations européennes (pays scandinaves inclus) et développées⁹. La confirmation de ce phénomène obérerait alors une originalité et un avantage majeur pour la soutenabilité du modèle économique et social français. La dynamique des migrations semble aujourd'hui encore plus incertaine, avec d'un côté des politiques restrictives affichées (mais pas toujours pratiquées) par certains pays européens, et de l'autre des besoins de main d'œuvre durablement avérés dans certains secteurs, métiers ou qualifications, joints à des flux possiblement aussi forts qu'imprévisibles liés aux troubles politiques et internationaux, et à de possibles « nouvelles concurrences » entre pays en ce domaine.

Une difficulté majeure réside enfin du passage du court, au moyen et au long terme dans les prévisions ou les projections réalisées, qui utilisent des cadres analytiques et des instruments différents. D'une part, des situations économiques contrastées peuvent perdurer sur les différents « marchés » de biens et de services, avec pour certains des tensions durables sur l'offre liées à des pénuries de matières premières ou de main d'œuvre, et pour d'autres des déficits de demande, ce qui peut conduire les politiques économiques à portée « générale » à jouer « à contre-emploi ».

En outre, si les prévisions officielles de court terme sont généralement révisées au fur et à mesure de l'évolution du consensus des organismes de prévision nationaux et internationaux, les prévisions de moyen terme utilisées sont généralement « normées », et résultent des hypothèses retenues par le gouvernement dans le cadre du « Plan budgétaire et structurel à moyen terme (PMST) » transmis à la Commission européenne à l'automne, et le cas échéant actualisé par le « rapport d'avancement » remis au printemps. Leur « raccordement » doit alors être opéré, sur la base d'explications et d'échéances plus ou moins plausibles, avec les tendances de long terme par ailleurs retenues dans des enceintes comme le COR, ce qui fait apparaître à l'évidence des problèmes de transition et de cadencement.

Ces sujets relatifs aux options possibles et au « réalisme » des scénarios à moyen terme, ainsi que leur cohérence économique et sociale avec des perspectives de long terme où la

Acemoglu, « *The Simple Macroeconomics of AI* », *Paper prepared for Economic Policy*, avril 2024 ; Philippe Aghion et Anne Bouverot, *IA : notre ambition pour la France*, Rapport de la Commission de l'intelligence artificielle, mars 2024.

⁹ Hélène Thélot, « Bilan démographique 2024, En 2024, la fécondité continue de diminuer, l'espérance de vie se stabilise », *Insee Première*, n°2033, janvier 2025.

démographie et les migrations jouent un rôle central, mériteraient d’être explicitement posés et débattus de façon partagée. Le cadre de ce débat peut être discuté¹⁰, mais il devrait en tout cas dépasser celui des seules projections sur les retraites et intégrer à la fois l’ensemble de la protection sociale (santé, dépendance, famille) et d’autres dépenses essentielles (éducation, adaptation/atténuation du changement climatique...) à la pérennité d’un modèle solidaire.

1.2) Des difficultés pour la construction des réformes et des politiques sociales dans un cadre raisonné

Cet ensemble d’incertitudes, aujourd’hui importantes et qui portent y compris sur les méthodes, pèse largement sur la possibilité de concevoir et de faire partager au « corps social » des cheminements viables et crédibles du court au moyen, puis au long terme, et, du côté des décideurs, à « cadencer » de façon satisfaisante les politiques économiques et sociales.

Le cadre d’élaboration de ces politiques se prête en outre assez mal à la prise en compte des évaluations réalisées ou en cours. Deux exemples peuvent en être donnés.

Tout d’abord, autour des modèles de micro-simulation, qui sont d’un apport précieux pour l’évaluation, *ex ante* ou *ex post*, des politiques sociales, et sont désormais partagés entre plusieurs organismes, voire ouverts aux utilisateurs. Un modèle « statique » comme Ines¹¹ permet ainsi, en identifiant les gagnants et les perdants d’une modification de la législation socio-fiscale, d’appréhender les effets redistributifs de ces politiques, même si cela reste le plus souvent sous l’hypothèse, difficile à lever, de « comportements constants ». Un modèle de micro-simulation « dynamique » comme Destinie¹² permet quant à lui de simuler, au niveau des individus ou des ménages, l’impact des politiques de retraite dans différentes hypothèses de législation et de comportement de liquidation, en éclairant le devenir des disparités de situation par sexe, niveau de rémunération ou génération. Or, si l’utilisation des propriétés redistributives de ces modèles est au cœur des études évaluatives *ex post* régulièrement conduites sur les politiques socio-fiscales et de retraites¹³, c’est un aspect si ce n’est ignoré, du moins beaucoup moins mis en avant, sauf par certains chercheurs, lors de la

¹⁰ Éric Heyer, Mathieu Plane, Xavier Ragot, Raul Sampognaro et Xavier Timbeau, « Quelles trajectoires pour les finances publiques de la France ? », *Policy Brief*, n°146, juillet 2025.

¹¹ Insee, Le modèle de microsimulation Ines, *Insee Méthodes*, février 2025.

¹² Insee, Le modèle de microsimulation Destinie 2, *Insee Méthodes*, septembre 2018.

¹³ Voir par exemple Sarah Abdouni, Gabriel Buresi, Fabien Delmas, Gwénaëlle Dumont, Lauriane Ramuzat, et Jules Tapin, « Les mesures socio-fiscales de 2023 : la non-reconduction des mesures exceptionnelles de 2022 diminue le revenu disponible des ménages modestes », *France Portrait social*, édition 2024, novembre 2025 ; Drees, « La redistribution diminue d’un tiers le taux de pauvreté », *Communiqué de presse*, juin 2025.

préparation des évolutions législatives, qui ont tendance à privilégier des estimations *ex ante* prioritairement centrées sur le « chiffrage budgétaire » des coûts associés.

En matière de prestations sociales, une variable clé est par ailleurs régulièrement oubliée, voire sciemment omise, des estimations *ex ante* présentées par les pouvoirs publics : il s'agit des « taux de recours » ou de « non recours »¹⁴, qui, comme l'ont montré les expériences de la CMU-C et du RSA activité, fournissent un critère central d'évaluation de la pertinence d'une mesure sociale¹⁵. Cette ignorance apparente peut refléter une difficulté objective à envisager les réactions d'adhésion des publics cibles, mais elle peut aussi être motivée par le désir politique d'afficher des résultats significatifs tout en tablant, de fait, sur des coûts plus réduits.

Enfin, les cadres mis en place pour la concertation et la discussion autour des politiques sociales, ne permettent, et c'est une limite, qu'un partage et une prise en compte partiels des travaux d'étude et d'évaluation aux stades préalables à la prise de décision.

Concernant les prévisions macroéconomiques et de finances publiques, une part notable des discussions et confrontations techniques entre organismes de prévision ont désormais lieu dans le cadre du Haut conseil des finances publiques (HCFP), chargé d'apprécier le réalisme des prévisions associées aux projets de lois financières et de vérifier la cohérence de la trajectoire des finances publiques avec les engagements pris par le gouvernement français au niveau européen. Dans la mesure où il ne donne que des avis sur le scénario gouvernemental et ne produit lui-même ni prévision ni chiffrages, certains économistes ont préconisé¹⁶, ainsi que son président à l'occasion des erreurs de prévision intervenues en 2024 en matière de finances publiques, de lui confier une mission élargie à l'élaboration de ce cadre macroéconomique et à l'évaluation *ex ante* du coût des réformes ou des mesures nouvelles. La pertinence du cadre relativement étroit et proche de la Cour des comptes qu'est HCFP¹⁷ peut toutefois être discutée pour l'organisation d'une confrontation large des prévisions macroéconomiques à différentes échéances. Il pourrait lui être préféré le retour à une formule

¹⁴ Définis comme la proportion des personnes qui accèdent ou n'accèdent pas aux prestations dont elles pourraient théoriquement bénéficier. Voir notamment Lucie Gonzalez et Emmanuelle Nauze-Fichet (coord.), Le non-recours aux prestations sociales- Mise en perspective et données disponibles, *Dossiers de la Drees*, n°57, juin 2020.

¹⁵ Philippe Warin, *Le non recours aux politiques sociales*, Presses universitaires de Grenoble, coll. "Libres cours Politique", 2016.

¹⁶ Adrien Auclert, Thomas Philippon et Xavier Ragot, « Quelle trajectoire pour les finances publiques françaises ? », *Les notes du conseil d'analyse économique*, n° 82, juillet 2024

¹⁷ Le HCFP, présidé par le Premier président de la Cour des comptes, comprend onze membres, dont quatre magistrats de la Cour des Comptes, le directeur général de l'Insee et cinq personnalités qualifiées.

proche de l'ancienne commission économique des comptes de la nation¹⁸, plus largement ouverte et moins focalisée sur la seule soutenabilité des finances publiques, sujet certes essentiel mais ne résumant pas tous les enjeux de politiques économique et sociale.

Concernant les politiques sanitaires et sociales proprement dites, les Hauts conseils institués dans les domaines de la santé publique et de la protection sociale¹⁹ ont des apports reconnus, grâce à des travaux qui associent partenaires sociaux, parlementaires et personnalités qualifiées et mobilisent simultanément les travaux des administrations économiques et des chercheurs, pour poser les problématiques et creuser, sous une variété d'angles et d'approches, les différents aspects des sujets examinés. Il existe en outre une capacité désormais rôdée à mener ou mobiliser, dans ce cadre de ces Hauts conseils ou au sein de France Stratégie, une série d'études évaluatives, voire des évaluations plus générales des réformes conduites dans ce champ (*cf. infra* point 2.2).

Ces travaux n'ont toutefois jusqu'ici guère été utilisés à leur juste valeur et parfois même mis en doute, comme le montre le fait d'avoir confié à la Cour des comptes en 2024-2025 une nouvelle évaluation prévisionnelle de la situation financière des régimes de retraites à l'horizon 2045, laquelle n'a d'ailleurs que peu modifié les concepts et les résultats émanant des projections du COR²⁰.

Par ailleurs, entre le moment de la réflexion et celui de l'évaluation, le « temps législatif » de la préparation et de la discussion des réformes perd de la transparence et occasionne une série de « brouillages », que l'on peut penser en partie évitables. En effet, les Hauts conseils suspendent habituellement l'examen des sujets traités pendant les débats parlementaires, alors que les études d'impact associées aux grands projets de loi économiques et sociaux sont de la seule compétence gouvernementale et, ainsi, en partie « normées ». Cela peut susciter des réactions de méfiance de la part des acteurs sociaux, ou même plus simplement des difficultés d'appropriation quant aux hypothèses et aux méthodes utilisées. Cela engendre aussi une absence de confrontation, excepté dans l'arène médiatique, entre ces estimations et celles réalisées concomitamment par des chercheurs, avec des écarts dont il est difficile d'apprécier les causes et la portée. Enfin, les études d'impact ne font pas, sauf

¹⁸ Celle-ci a connu deux « avatars » : la Commission des comptes et des budgets économiques de la nation, qui a été remplacée en 1999 par la Commission économique de la nation, elle-même supprimée en 2015.

¹⁹ Il s'agit du Haut conseil de la santé publique (HCSP), du Conseil d'orientation des retraites (COR), du Haut conseil pour l'avenir de l'assurance maladie (HCAAM), du Haut Conseil de la famille, de l'enfance et de l'âge (HCFEA), et du Haut Conseil du financement de la protection sociale (HCFiPS).

²⁰ Voir Cour des comptes, *Situation financière et perspectives du système de retraites*, *ibid.* et Conseil d'orientation des retraites, *Évolutions et perspectives des retraites en France*, *ibid.*

exception, l'objet d'une réactualisation lors des débats parlementaires, même lorsque des amendements importants et structurants sont intégrés aux réformes.

Le « conclave » sur les retraites organisé en 2025 a pu à cet égard marquer une inflexion, dans la mesure où les partenaires sociaux appelés par le gouvernement à négocier sur la réforme des retraites se sont trouvés alimentés par des simulations sur chacune des pistes envisagées, sachant que cela est habituellement le cas pour les régimes paritaires qu'ils gèrent (Agirc-Arrco, Unedic). Quel que soit l'avenir concernant le mode de « pilotage », il pourrait aussi, et plus généralement être utile de réfléchir à un « cadre de confiance » permettant l'analyse et la confrontation, avec des spécialistes et des chercheurs, des résultats des différents chiffrages et études effectués lors de la production d'un texte législatif et de sa discussion au Parlement, après ou avant que « la main soit prise » par les Hauts conseils dans les périodes préalables ou postérieures aux réformes.

II) L'évaluation des politiques publiques : un terme galvaudé, des objectifs parfois flous et des pratiques disparates

Dans le domaine des politiques publiques, l'idée d'évaluation est depuis les années 1990 promue de façon de plus en plus affirmée, donnant lieu à un « affichage » consensuel, et à une hypothèse sous-jacente : celle que sa pratique pourrait donner aux pouvoirs publics le moyen de faire des choix « scientifiquement éclairés » que le discours politique serait impuissant à proposer. On peut certes convenir que l'évaluation des politiques publiques constitue une démarche indispensable pour en resituer et en « décrypter les enjeux »²¹ en montrer les effets non attendus, et surtout en analyser l'impact eu égard à leur environnement, leurs acteurs et leurs usagers. L'appréciation qu'elle vise à apporter sur l'action publique peut être ainsi lue non seulement comme une aide à la décision politique mais aussi comme une démarche explicative envers les citoyens, qui conduit certains à en revendiquer l'usage comme un élément de la vie démocratique²².

L'évaluation des politiques publiques s'est à cet égard beaucoup développée dans le domaine de la protection sociale, et plus largement des politiques sanitaires et sociales, avec des progrès incontestables. Cela s'est toutefois accompagné d'une tendance discutable à invoquer à tout propos la notion d'évaluation, à propos d'un corpus de bilans, rapports ou

²¹ Jérôme Minonzio, « L'évaluation des politiques familiales : discours sur la méthode ou mot d'ordre gestionnaire ? », *Informations sociales*, n°150, 2008/6.

²² Nasser Mansouri - Guilani, *Promouvoir la culture de l'évaluation des politiques publiques*, Avis et rapport du Conseil économique, social et environnemental, septembre 2015.

études à la portée et aux méthodes disparates, établis par des intervenants divers, qui pourrait sans clarification nuire à la viabilité même de la démarche.

2.1) L'approche évaluative questionnée en amont par le recours « pathologique » aux indicateurs

En premier lieu, le recours « tous azimuts » aux indicateurs dit « de performance » a pu pendant un temps faire penser que les démarches évaluatives allaient être mises au deuxième plan. Depuis une vingtaine d'années en effet, l'appréciation des résultats des politiques publiques (« stratégies », programmes, « plans », et surtout lois financières annuelles suite à la LOLF...), s'est focalisée sur une batterie limitée « d'indicateurs synthétiques », comme le meilleur moyen d'en apprécier les « performances ».

L'utilisation d'indicateurs dans la sphère publique est loin nouvelle, à commencer par les principales données qui structurent de façon permanente le débat économique et social (PIB, taux de chômage, risque et seuil de pauvreté, ...), tout en étant matière à discussion sur leur contenu et leurs imperfections. Ces indicateurs sont d'ailleurs devenus des enjeux dans les débats conduits avec la société civile, qu'il s'agisse de mettre en évidence et de critiquer les limites du PIB²³, ou d'enrichir la batterie des indicateurs utilisés dans la sphère économique par la prise en compte de dimensions sociales et environnementales²⁴.

Le développement systématique de l'usage des indicateurs fait cependant surtout référence à un concept de « performance » inspiré des théories du management et de la « nouvelle gestion » publics. Ils sont supposés dans cette perspective servir à mesurer la réalisation « d'objectifs », que les politiques et les organisations publiques doivent se voir assigner de la façon la plus claire possible, et permettre, à eux seuls, une appréciation « quantitativement objectivée » des résultats obtenus. La mise en place d'indicateurs va alors souvent de pair avec l'affichage d'objectifs chiffrés, sans que le passage entre les deux n'ait d'ailleurs forcément été pensé.

Ces démarches ont fait l'objet d'importants débats portant au premier chef sur leurs fondements. Pour certains auteurs, elles viseraient à assurer une plus grande transparence sur les objectifs et les résultats de l'action publique vis-à-vis du Parlement, des acteurs sociaux,

²³ Joseph Stiglitz, Amartya Sen et Jean-Paul Fitoussi, *Rapport de la Commission sur la mesure des performances économiques et du progrès social*, Éditions Odile Jacob, 2009.

²⁴ Voir notamment Conseil national de l'information statistique, *La déclinaison française des indicateurs de suivi des objectifs de développement durable*, Rapport du groupe de travail du Cnis, 2018, Insee, *Les nouveaux indicateurs de richesse – Indicateurs statistiques*, Chiffres détaillés, janvier 2025 et Insee, *Indicateurs pour le suivi national des objectifs de développement durable*, Chiffres détaillés, février 2025.

des contribuables et des citoyens, selon un principe dit « d'*accountability* »²⁵. Selon d'autres points de vue, elles traduiraient au contraire la volonté de limiter les interventions publiques à la plus stricte « utilité », et le recours systématique aux indicateurs serait la traduction d'une « gouvernance par les nombres » et servirait de support à une contractualisation de type « marchand » entre les organismes administratifs et ceux qui en ont la responsabilité politique ou la tutelle.²⁶

Ces réflexions montrent à la fois les apports potentiels du recours aux indicateurs pour le suivi et l'appréciation des programmes sociaux, mais aussi ses limites majeures. D'un côté, il permet de s'interroger sur les missions prioritaires assignées à l'action publique, d'inciter au développement d'outils de suivi statistique, de susciter d'éventuelles alertes ou inflexions et de contraindre à la diffusion, interne et externe, d'une information plus transparente.

Des « effets indésirables » majeurs sont cependant à déplorer, au premier rang desquels la prolifération de batteries d'indicateurs en tous sens, laquelle se mue parfois en véritable « pathologie ». Elle s'accompagne du risque de polariser l'attention sur la seule « facette » de l'action publique éclairée par les indicateurs retenus, ainsi que sur les activités ou les publics les plus « rentables » pour se rapprocher des résultats espérés. S'y ajoutent des problèmes de maîtrise, de cohérence et d'articulation entre des indicateurs suivis à des dates ou dans des champs différents, et plus fondamentalement, le fait que ces indicateurs ne mesurent par définition que des résultats déjà anticipés, attendus et définis, alors que l'intérêt de l'action publique est de pouvoir prendre en compte de nouvelles questions stratégiques et des impacts ou problèmes non attendus.

Du point de vue des organismes publics, peu de réflexions ont en outre lieu sur le passage, souvent fait de façon « automatique » entre la notion d'indicateur et celle « d'objectif », ainsi que sur la pertinence d'une déclinaison automatique de ces objectifs à des niveaux très désagrégés des organisations ou des territoires. Les acteurs ou les personnels en charge de ces domaines d'action disent alors de plus en plus souvent avoir à « rapporter » sur des données dont ils ne reconnaissent ni la pertinence ni la validité pour apprécier le contenu de leur travail et la réussite de leurs actions. À cet égard, la systématisation du recours aux objectifs et indicateurs dits « de performance » comme unique critère de jugement doit de plus être

²⁵ Edward Arkwright., Christian de Boissieu, Jean-Hervé Lorenzi, Julien Samson, *Économie politique de la Lolf*, Rapport du Conseil d'analyse économique, La Documentation française, 2007.

²⁶ Alain Supiot, *La Gouvernance par les nombres : cours au Collège de France 2012-2014*, Fayard, coll. « Poids et mesures du monde, 2015.

considérée comme relevant d'une « obsession métrique » qui peut conduire à des erreurs d'appréciation, voire à des effets pervers obérant l'intérêt de ces approches²⁷.

La « politique des indicateurs » ne saurait donc se substituer à l'éclairage des réalités sociales par des études et des enquêtes, non plus qu'à des processus construits d'évaluation dont, au niveau des politiques sociales comme des organisations qui les mettent en œuvre, la production d'indicateurs (y compris lorsqu'elle est baptisée « rapport d'évaluation » comme dans le projet de loi de financement de la sécurité sociale - PLFSS), ne peut au mieux servir que d'alerte ou de première amorce.

2.2) Des démarches d'évaluation marquées par la fragmentation des intervenants et l'hétérogénéité des objectifs et des méthodes

L'évaluation des politiques sociales est aujourd'hui en France, au-delà d'un objet commun affiché consistant à rechercher si les moyens juridiques, administratifs ou financiers mis en œuvre permettent de produire les effets attendus et d'atteindre les buts assignés, marquée par la polysémie et le flou sur ses propres objectifs, la multiplicité des organismes qui s'en réclament, et, au bout du compte, le caractère disparate des pratiques observées en termes de contenu et de méthodes.

Si l'évaluation porte de façon centrale sur l'efficacité (les effets obtenus eu égard aux objectifs recherchés et la qualité du service rendu) et l'efficience (le rapport entre les ressources mobilisées et les résultats obtenus) des politiques qu'elle considère, une définition plus complète devrait conduire à y inclure l'appréciation de leur cohérence, de leur pertinence et de la diversité de leurs impacts sur la société. Elle devrait à cet égard relever d'une démarche pluraliste, mobilisant des données à la fois quantitatives et qualitatives, associant experts et intervenants de terrain. Elle aurait dans ce cadre vocation à confronter des points de vue issus de plusieurs disciplines des sciences économiques et sociales, et à replacer l'action publique au sein d'un système d'acteurs, prenant en compte l'appréciation des usagers et de leurs représentants²⁸.

L'évaluation des politiques sociales ne devrait donc pas se réduire à la production de statistiques et d'études évaluatives, même si cet aspect est central lorsque l'on cherche à produire des évaluations « d'impact » selon des méthodes rigoureuses²⁹. Elle implique une

²⁷ Gaël de Peretti, « *Le diktat des indicateurs* », La vie des idées, 2019

²⁸ Bernard Perret, *L'évaluation des politiques publiques*, La Découverte, coll. « Repères », 3e éd., 2014.

²⁹ Voir notamment CAE, « Évaluation des politiques publiques », *Les notes du Conseil d'analyse économique*, n°1, février 2013 et Rozenn Desplat et Marc Ferracci, *Comment évaluer l'impact des politiques publiques ? Un guide à l'usage des décideurs et des praticiens*, Collection Évaluation, France Stratégie, septembre 2016.

approche spécifique, reposant sur un questionnaire et une analyse d'ensemble, dans un cadre institutionnel assurant la déontologie, l'indépendance et liberté de diffusion des résultats obtenus³⁰.

Or, notamment en matière de politiques sanitaires et sociales, on observe aujourd'hui un foisonnement d'initiatives et de structures se réclamant de l'évaluation, qui révèle un processus d'institutionnalisation fragmenté et inachevé. Historiquement, après l'étape importante du dispositif interministériel institué par le gouvernement Rocard en 1990, le système a été simplifié et resserré en 1998, puis suspendu en 2002. Un déplacement est ensuite intervenu, avec les exercices successifs dits de « réforme de l'État » (révision générale de politiques publiques -RGPP- en 2009, puis modernisation de l'action publique –MAP- entre 2012 et 2017), mais ces approches se sont davantage centrées sur des processus d'audit orientés vers la maîtrise des dépenses publiques que sur le développement des démarches d'évaluation proprement dites³¹. La mission d'évaluation confiée par la révision constitutionnelle de 2008 au Parlement, avec l'assistance de la Cour des comptes n'a de ce point de vue guère levé l'ambiguïté.

Au bout du compte, l'évaluation des politiques sociales donne aujourd'hui lieu à un foisonnement d'opérations conduites selon des objectifs et des modalités divers, posant des questions de coordination et de cohérence, et reflétant un « agenda » inachevé et des difficultés d'institutionnalisation³².

Parmi la variété de structures qui « revendiquent » des missions d'évaluation dans les domaines sanitaires et sociaux, figurent ainsi notamment :

— du côté de la recherche, des structures spécialisées, le plus souvent à dominante d'économie quantitative et parfois soutenues *via* différents canaux par les pouvoirs publics : Centre de recherche en économie et statistique (CREST), Institut des politiques publiques (IPP), Fédération de recherche Travail, emploi et politiques publiques (TEPP), Laboratoire interdisciplinaire d'évaluation des politiques publiques (LIEPP), Observatoire français des conjonctures économiques (OFCE) ...; les financements dévolus à l'évaluation restent toutefois limités et n'intègrent pas forcément de « programme socle » d'un montant

³⁰ Patrice Duran, « L'évaluation des politiques publiques : les sciences sociales comme sciences de gouvernement », *Idées économiques et sociales*, n°193, 2018/3.

³¹ Antoine Bozio et Lucile Romanello, « Évaluation des politiques publiques : le bilan contrasté du quinquennat », *Les notes de l'IPP*, n°25, mars 2017.

³² Rozenn Desplat et Clément Lacouette-Fougère, « L'évaluation des politiques publiques en France », *Document de travail de France Stratégie*, 2019/13.

suffisant³³ ; ils restent avant tout ciblés sur les promoteurs d'approches quantitatives, alors que d'autres chercheurs ou organismes, tels le Centre d'études de l'emploi et du travail (CEET), l'Institut de recherches économiques et sociales (Ires) ou certaines universités sont aussi parties prenantes des processus d'évaluation, notamment du point de vue des acteurs ou des bénéficiaires des politiques ;

— au sein des différents ministères, les corps d'inspection (Inspection générale des affaires sociales, Inspection générale des finances, parfois de façon conjointe), et les services statistiques ministériels, mais aussi les principaux opérateurs publics (France Travail, agences sanitaires, caisses de sécurité sociale), qui prennent directement des initiatives en ce domaine, parfois prévues dans leurs processus de contractualisation avec l'État ;

— au niveau interministériel, France Stratégie, héritier à la suite du Commissariat au plan de qui « héberge » certaines commissions ou d'instances d'évaluation demandant des moyens importants, ainsi que le réseau des Hauts conseils qu'il héberge (*cf. supra* point 1.2) ;

— la Cour des comptes, qui, pour le compte ou non du Parlement, se réclame de plus en plus de l'objectif d'évaluation dans l'orientation de ses travaux ;

— sous l'égide du Parlement, l'Office parlementaire d'évaluation des politiques de santé, la Mission d'évaluation et de contrôle (MEC) instituée au sein de la Commission des finances de l'Assemblée nationale, complétée, pour la protection sociale, par une mission particulière créée en 2004, la MECSS et, depuis 2009, par un comité d'évaluation et de contrôle qui coordonne les opérations d'un programme directement déclenché par l'Assemblée nationale ;

— du côté des acteurs sociaux, le Conseil économique, social et environnemental (CESE) qui a notamment vocation, depuis sa réforme du printemps 2021, à intervenir en intégrant en tant que de besoin une « participation citoyenne ».

Comparé à celui qui prévaut dans d'autres pays, ce dispositif apparaît marqué par l'absence de coordination des différents programmes au niveau national, tout en restant d'ampleur assez limitée, principalement orienté vers les spécialistes et avec de faibles retombées sur les processus de décision³⁴. Il n'échappe pas pour autant à la prédominance d'administrations d'État encore relativement peu ouvertes aux universitaires et aux chercheurs, ce qui laisse ouverte la question de son « ancrage institutionnel », et notamment

³³ Jean-Luc Tavernier et Nicolas Véron, *Analyse économique et évaluation des politiques publiques*, Rapport à la première ministre, novembre 2023.

³⁴ Adam Baiz, Pierre-Henri Bono, Sandra Debu, Rozenn Desplatz, Clément Lacouette-Fougère et Alain Trannoy, « Vingt ans d'évaluations d'impact en France et à l'étranger. Analyse comparée des pratiques dans six pays », Document de travail de France Stratégie, 2019/16.

d'une « maison commune »³⁵ assurant à la fois la crédibilité scientifique des évaluations et l'association de l'ensemble des acteurs intéressés³⁶.

Cette profusion et cette fragmentation sont particulièrement visibles dans le domaine des politiques sanitaires et sociales, et notamment des politiques d'emploi et d'insertion (s'agissant par exemple des mesures d'allègement du coût du travail, du RSA ou des mesures en faveur des jeunes et des chômeurs de longue durée). Elles se traduisent par un foisonnement « concurrentiel » d'approches qui pose des problèmes de cohérence et de lisibilité, dans la mesure où la multiplicité des intervenants institutionnels aboutit à des disparités de champs et de méthodes qui suscitent une série d'interrogations, voire de confusions :

- difficulté à distinguer les opérations d'évaluation des audits et des contrôles ;
- absence de protocoles et de méthodologies partagées, un bilan d'exécution ou une enquête de satisfaction pouvant çà et là être baptisés évaluation ;
- questions sur l'indépendance des évaluations réalisées sous l'égide d'acteurs directement responsables de la mise en œuvre des politiques ;
- multiplicité d'approches et de résultats qui peuvent être pour partie contradictoires, ne font pas toujours l'objet de synthèses comparatives et entretiennent une impression de confusion et d'opacité ;
- absence fréquente d'association des acteurs sociaux à une démarche qui se voulait initialement participative et démocratique ;
- carences manifestes concernant l'évaluation de dispositifs moins placés sous le feu des priorités politiques ou des méthodes expérimentales, comme les couvertures collectives en matière d'invalidité et de prévoyance ;

Certaines des études les plus mises en avant sont au bout du compte souvent polarisées sur une ou deux variables d'intérêt « phares » (par exemple, le taux de retour à l'emploi), sachant qu'elles ne peuvent résumer toutes les dimensions de l'évaluation, et ne devraient pas faire oublier l'ensemble de ses autres aspects. Sans dispositif spécifique de coordination, les synthèses conclusives peuvent en outre se révéler difficiles, ce qui impose des démarches de clarification, tant sur le « vocabulaire » que sur la portée des travaux et des résultats considérés (*cf. infra* point 3.4).

³⁵ Jean-Luc Tavernier et Nicolas Véron, *ibid.*

³⁶ Voir notamment Rozenn Desplat et Clément Lacouette-Fougère, *ibid.* et Patrice Duran, *ibid.*

2.3) La « mantra » des expérimentations en débat

Attachée à la problématique de l'évaluation, l'idée d'élaborer les politiques sociales sur la base « d'expérimentations » a connu un développement certain suite à la réforme constitutionnelle de mars 2003 autorisant l'expérimentation normative, mais aussi dans le cadre de l'article 72 de la Constitution qui permet, au plan local, des dérogations expérimentales aux dispositifs nationaux³⁷. La thèse générale avancée était que, si ces expérimentations étaient suivies d'évaluations menées selon des modalités « scientifiquement valides », elles permettraient d'éclairer l'opportunité d'une part, les conditions d'autre part, d'une généralisation de dispositifs ayant *a priori* vocation à être implantés à l'échelle nationale³⁸.

Les « expérimentations sociales » se sont notamment étendues depuis 2008, avec, d'une part, l'expérimentation du RSA dans une trentaine de départements dans la période précédant sa généralisation³⁹, et, d'autre part, le lancement du fonds d'expérimentation pour la jeunesse (FEJ). Tandis que les appels à projets lancés par ce dernier se sont progressivement taris, les expérimentations ont depuis lors notamment concerné, dans le domaine social, la garantie jeunes, une expérimentation territoriale visant à résorber le chômage de longue durée (ETCLD) suite à une initiative promue par ATD-Quart Monde⁴⁰, le renforcement de l'accompagnement de certains chômeurs ou le programme « Un chez-soi d'abord » en faveur du logement des personnes sans-abri présentant des troubles psychiques ou des addictions. Les expérimentations des années récentes ont également porté sur des projets territoriaux visant à « éliminer » le chômage de longue durée⁴¹ ou le non-recours aux prestations sociales (expérimentations « Territoires zéro chômeur de longue durée » et « Territoires zéro non recours »). Elles ont continué à se focaliser de façon répétée sur le RSA avec, d'une part, l'expérimentation lancée en 2021-22 de la recentralisation de l'attribution et du financement de l'allocation, et, d'autre part, celle qui a précédé en 2024 la généralisation de l'accompagnement et de la mise sous condition d'activité du RSA dans le contexte de la création de « France Travail ».

³⁷ Bernard Gomel et Évelyne Serverin, « L'expérimentation des politiques publiques dans tous ses états », *Informations sociales*, n°174, 2012/6.

³⁸ François Legendre « Le recours croissant aux expérimentations sociales en économie », *Informations sociales*, n°174 2012/6.

³⁹ Guillaume Allègre, « L'expérimentation du RSA et ses ambiguïtés », *Informations sociales*, n°174, 2012/6.

⁴⁰ Anne Perrot, Geneviève Lallemand-Kirche et Florence Allot, *L'évaluation économique de l'expérimentation visant à résorber le chômage de longue durée (ETCLD)*, Rapport Igas-IGF, octobre 2019.

⁴¹ Voir Olivier Bouba Olga (président), *Expérimentation Territoires zéro chômeur de longue durée*, Rapport final du comité scientifique, avril 2021 et Cour des comptes, *L'expérimentation « Territoire zéro chômeur de longue durée »*, Rapport, juin 2025.

Dans le secteur de la santé, la loi de financement de la sécurité sociale pour 2018 a introduit, en son article 51, un dispositif permettant d'expérimenter de nouvelles organisations reposant sur des modes de financement dérogatoires, sous condition que ces organisations contribuent à améliorer le parcours des patients, l'efficacité du système de santé, l'accès aux soins ou la pertinence de la prescription des produits de santé. Le dispositif, en vigueur depuis avril 2018, a fait l'objet d'un bilan à cinq ans, montrant un ralentissement du nombre de nouvelles expérimentations autorisées en 2022 et 2023⁴², sachant que d'autres expérimentations sanitaires ont également pris place, concernant par exemple les « salles de consommation à moindre risque » (SCMR) ou le cannabis thérapeutique.

Ces initiatives se sont souvent accompagnées de dispositifs d'études comprenant des indicateurs de suivi, des enquêtes statistiques auprès des bénéficiaires et non bénéficiaires des dispositifs et/ou des monographies visant à appréhender les comportements des acteurs. Elles ont pu donner lieu à différents rapports d'évaluation, le cas échéant issus de comités scientifiques ou de corps d'inspection.

Néanmoins, au-delà même de leur contenu intrinsèque, on peut s'interroger sur ce recours aux politiques d'expérimentation et sur son utilité réelle pour l'action publique. Certains économistes et chercheurs estiment que les expérimentations n'ont pas donné le fruit potentiel attendu parce qu'elles n'ont pas été conduites et évaluées de façon suffisamment rigoureuse : cela aurait nécessité des conditions de mise en œuvre plus précisément encadrées pour rapprocher le plus possible d'une sélection aléatoire des bénéficiaires (encadré 1). Cela aurait aussi nécessité, de façon évidente, que les décisions de modification ou de généralisation des dispositifs concernés ne soient pas prises avant la mise à disposition des évaluations correspondantes, et parfois même avant la fin des expérimentations elles-mêmes.

La pertinence des expérimentations sociales peut par ailleurs être mise en doute au regard de l'objectif de généralisation mis en avant à leur propos. D'un point de vue technique, l'évaluation d'expérimentations circonscrites ne garantit pas que les avantages différentiels donnés aux bénéficiaires, par exemple à l'occasion de mesures d'insertion, restent les mêmes dans un contexte plus large, ni que leurs comportements demeurent identiques en l'absence de la mobilisation exceptionnelle à laquelle l'expérimentation a donné lieu (encadré 1).

⁴² Ministère de la santé et de la prévention, *Rapport au Parlement sur les expérimentations innovantes en santé, Article 51 de la loi de financement de la sécurité sociale pour 2018*, 2023.

ENCADRÉ 1

Les expérimentations contrôlées en matière de politiques sociales : principes et usages

Les expérimentations contrôlées sont en théorie le moyen de procéder à une évaluation réellement aléatoire des dispositifs publics, en dispensant de procéder à des corrections forcément imparfaites des biais de sélectivité liés aux caractéristiques observables et inobservables des bénéficiaires (cf. *infra* point 3.1)⁴³.

Elles nécessitent un partenariat entre expérimentateurs et évaluateurs et des conditions de mise en œuvre précises et rigoureuses. Si plusieurs méthodes existent pour « randomiser » les publics faisant l'objet d'un « traitement » par les dispositifs expérimentés, toutes comportent une forme d'affectation aléatoire (par tirage au sort) des bénéficiaires à ces programmes, qui peut s'avérer plus lourde et complexe, et en tout cas plus difficile à maîtriser que le recours à des observations existantes.

En dépit de la richesse de ces opérations, il peut être difficile d'en dégager des résultats significatifs, et surtout d'en tirer de réels enseignements sur l'impact d'une extension à l'échelle nationale. La validité de leur interprétation est d'abord soumise à un ensemble de conditions souvent difficiles à réaliser sous la responsabilité d'acteurs locaux ou déconcentrés, et dont le non-respect ne peut être que partiellement corrigé par des méthodes statistiques : sélection incomplètement aléatoire des territoires ou des groupes « de traitement » et « de contrôle », problèmes de significativité des impacts estimés⁴⁴. De façon plus fondamentale, le passage entre l'évaluation de ces expérimentations et leur généralisation potentielle est loin d'être immédiat⁴⁵ :

- les comportements des individus participant à l'étude peuvent être influencés par le fait même qu'il s'agit d'une expérimentation (effet Hawthorne), ce qui n'est pas sans poser problème eu égard à un objectif de généralisation ;
- cette généralisation rendrait permanentes des mesures temporaires, dont c'est précisément le caractère exceptionnel qui a pu contribuer à l'incitation, donnant lieu à une mobilisation particulière des opérateurs et à une affectation fréquente de ressources supplémentaires, guère transposables à l'échelle générale ;
- la généralisation amenuise enfin, en le « diluant », l'avantage relatif donné aux populations cibles et fait passer d'un équilibre partiel à un équilibre global, sachant que les activités aidées dans le cadre de l'expérimentation ont pu exercer un impact positif en concurrençant celles d'autres acteurs économiques (effet d'aubaine et question du « bouclage » économique).

⁴³ Voir notamment Yannick L'Horty et Pascal Petit, « Évaluation aléatoire et expérimentations sociales », *Revue française d'économie*, vol. XXVI, n°1, 2011 ; Sylvain Chabé-Ferret, Laura Dupont-Courtade et Nicolas Treich, « Évaluation des politiques publiques : expérimentation randomisée et méthodes quasi-expérimentales », *Économie et prévision*, n°211-212, 2017/2.

⁴⁴ Agnès Labrousse et Philippe Zamora, « Expérimentations de terrain et politiques publiques du travail et de l'emploi, Apports récents et mise en perspective », *Travail et Emploi*, n°135 (2013), janvier 2014.

⁴⁵ Voir notamment Guillaume Allègre, « L'expérimentation sociale des incitations financières à l'emploi : questions méthodologiques et leçons des expériences nord-américaines », *Document de travail de l'OFCE*, n°22, juillet 2008 et Rozenn Desplat et Marc Ferracci, *ibid.*

Plus fondamentalement, la décision de s'en remettre à une expérimentation est en général prise dans un contexte politique particulier, soit qu'il s'agisse d'une mesure dont le principe n'est pas réellement accepté, soit que l'on vise une première avancée concernant des dispositifs coûteux dont le financement n'est pas immédiatement programmé. Ce caractère de « processus politique » des expérimentations signifie que leur objectif réel est souvent autre que celui affiché⁴⁶. Plutôt que rechercher une généralisation peu probable, il peut s'agir de soutenir des équipes ou des projets porteurs de changements souhaitables, sans modifier autrement qu'à la marge le fonctionnement de base du système. Le recours aux expérimentations peut donc être un moyen d'offrir aux acteurs des capacités d'initiative par rapport aux politiques nationales : le reconnaître impliquerait de ne pas centrer le processus d'évaluation sur la perspective d'une éventuelle généralisation, mais plus directement sur l'accompagnement et la pérennisation de ces expériences.

Au bout du compte, les objectifs, contenus et modalités, institutionnels comme scientifiques, de l'évaluation des politiques sanitaires et sociales, doivent donc donner lieu à des réflexions qui dépassent largement le tropisme des indicateurs, la « mantra » des expérimentations, et même le « tout quantitatif » de certaines études, en dépit du fait que leurs méthodes se soient beaucoup perfectionnées (*cf. infra* point 3.1). Les idées de transversalité et de diversité des regards qui sous-tendent depuis l'origine la notion d'évaluation devraient dans cette perspective retrouver toute leur valeur.

III) Une tension croissante entre la sophistication des méthodes d'évaluation et l'accélération du débat social, sous contrainte politique

Au-delà des problèmes liés au cadre économique et institutionnel des évaluations, la lucidité impose enfin de tenir compte de l'obstacle majeur qu'introduit, en termes d'appropriation, la divergence croissante entre la sophistication des méthodes mobilisées et l'accélération du débat social.

⁴⁶ Bernard Gomel et Évelyne Serverin, « L'expérimentation sociales aléatoire en France en trois questions », *Travail et Emploi*, n°135 (2013), janvier 2014.

3.1.) Un renouveau des méthodes quantitatives largement focalisé sur les études évaluatives des politiques d'emploi et d'insertion ...

Sans revenir au cas spécifique des expérimentations (*cf. supra* point 2.3), les progrès des méthodes d'analyse économétrique se sont fortement focalisés sur l'évaluation des politiques publiques, et en son sein sur celle des politiques sociales, d'emploi ou d'insertion.

Il s'agit plus précisément d'études quantitatives à dimension évaluative, dont le développement a mobilisé un ensemble de chercheurs en économie en partie à des fins de « reconnaissance par leurs pairs », et qui ont pour objectif revendiqué de comparer dans une perspective longitudinale la situation des bénéficiaires d'aides publiques (individus ou entreprises) à celle d'échantillons « témoins » de non bénéficiaires, en corrigeant les « biais de sélectivité » dus au fait que l'accès à ces aides n'est pas aléatoire (encadré 2).

ENCADRÉ 2

Les méthodes économétriques pour l'évaluation de politiques sociales

L'un des aspects majeurs de l'évaluation des politiques sociales consiste à rechercher si le passage par un dispositif (allocation, aide à l'emploi, formation...) a un effet sur le devenir à plus ou moins long terme de ses bénéficiaires (individus, ménages ou entreprises), dont on suppose le comportement modifié. Ces approches reposent sur la mobilisation de données individuelles permettant de suivre leurs trajectoires en les comparant à des « échantillons témoins » de non-bénéficiaires, au regard des hypothèses de comportement attendues.

Ces approches, particulièrement utilisées dans le champ des politiques d'emploi et d'insertion, se heurtent au fait que l'accès aux dispositifs donne lieu à des « biais de sélectivité », qui sont à la fois liés aux caractéristiques observables des bénéficiaires (sexe, âge, diplôme, région, parcours antérieur...), mais aussi à des caractéristiques « inobservables » (par exemple les relations sociales ou les capacités des personnes ou, pour les entreprises, leur propension à l'innovation), dont la non prise en compte conduirait à une évaluation biaisée⁴⁷.

L'effacement de ces biais implique de recourir soit à des « expérimentations contrôlées » de type aléatoire (*cf. supra* point 3.1), soit à des méthodes d'estimation spécifiques dites « quasi-expérimentales », qui reposent au premier chef sur la constitution de « groupes de contrôle » (groupes témoins de non bénéficiaires) dont les caractéristiques doivent être les plus proches possibles de celles des « groupes de traitement » (les bénéficiaires).

Parmi ces méthodes, on peut citer :

- le recours à des « variables instrumentales » expliquant la participation à un programme mais n'influant pas sur ses résultats ;

⁴⁷ Voir notamment Denis Fougère, « Les méthodes économétriques d'évaluation », *Revue française des affaires sociales*, n°1-2, juillet 2010 ; Sylvain Chabé-Ferret, Laura Dupont-Courtade et Nicolas Treich, *ibid.*

- l'estimation « par appariement » d'échantillons « traités » et « non traités », cet appariement s'opérant soit sur leurs caractéristiques observables, soit sur un « score de propension » résumant l'effet des variables dites « de conditionnement » ;
- l'estimation « par différence de différences », qui compare les évolutions dans le temps des « variables d'intérêt » pour les groupes traités et non traités, en supposant que les effets temporels jouent sur eux de la même façon ;
- le recours à des modèles Tobit de sélection endogène, qui modélisent conjointement la probabilité d'accéder à un dispositif (avec des résidus incluant l'incidence des variables inobservées), et l'impact de celui-ci⁴⁸.

Ces méthodes nécessitant des données et des périodes d'observation adéquates, leur utilisation a été largement favorisée par le développement de bases de données impliquant fréquemment la statistique publique. Elles résultent d'une part de recueils systématiques d'informations sur les bénéficiaires des grands programmes ou dispositifs publics, et d'autre part de la possibilité de constituer des panels longitudinaux de ménages, de travailleurs ou d'entreprises, à partir de diverses sources permanentes de données fiscales, sociales ou sanitaires (au travers du système national des données de santé - SNDS), susceptibles d'être appariées avec les précédentes dans des conditions préservant l'anonymat et la sécurité des données individuelles. Cela a par exemple été le cas pour les minima sociaux ou l'allocation personnalisée d'autonomie, ou, avec une optique différente, pour des dispositifs d'aide aux entreprises comme le crédit d'impôt compétitivité-emploi (CICE) ou, dans leurs différents avatars, les exonérations de charges sur les bas salaires.

3.2) ...mais qui implique du temps, des synthèses interprétatives et des compléments d'analyse

Pour intéressantes qu'elles soient, ces approches méthodologiques doivent impérativement s'inscrire dans un temps assez long, tenant compte de la durée pertinente d'observation du « devenir » des personnes ou des entreprises concernées (parfois plusieurs années), ainsi que des différentes étapes indispensables au recueil des informations, à leur appariement avec les bases de données adéquates et, le cas échéant, à la sollicitation d'équipes de recherche.

⁴⁸ Voir notamment Pauline Givord, « Méthodes économétriques pour l'évaluation des politiques publiques », *Document de travail de la Direction des Études et Synthèses économiques de l'Insee*, G2010 / 08 ; Denis Fougère et Nicolas Jacquemet, « Inférence causale et évaluation d'impact », *Économie et statistique*, n°510-511-512, décembre 2019.

Ces études sont par ailleurs souvent réalisées, y compris pour l'évaluation d'un même dispositif et lorsque l'initiative a été lancée de façon coordonnée, avec des champs, des périodes d'observation et des méthodes de modélisation différents. Des besoins de mise en perspective ou de « réconciliation » de ces analyses économétriques parfois sophistiquées dont les conclusions peuvent en partie diverger apparaissent alors souvent, nécessitant des travaux de synthèse qui peuvent aboutir à des « fourchettes d'estimation » assez larges, par exemple en ce qui concerne les créations d'emplois induites par les mesures d'allègement du coût du travail⁴⁹, voire à la persistance d'interrogations ouvertes comme celles relatives à l'existence de « trappes à bas salaires »⁵⁰.

Les résultats de ces approches économétriques doivent en outre, au-delà du chiffre « emblématique » des « effets emploi » associés aux différents dispositifs, être confrontés d'autres sources de complexité, essentielles à prendre en compte. Toujours en matière de coût du travail, il semble par exemple que l'usage de l'avantage financier liés aux allègements de charges connaisse d'importantes variations en fonction de la situation des entreprises en termes de contraintes financières, d'environnement concurrentiel, de difficultés de recrutement ou même de position dans la « chaîne de valeurs », selon que ces entreprises sont donneuses d'ordre ou sous-traitantes⁵¹.

Et il importe aussi, ce que ces modélisations ne peuvent directement faire, d'expliquer pourquoi une incitation financière du même ordre peut être diversement mobilisée, voire induire des effets différents, selon la « catégorisation » des dispositifs, leurs conditions de mise en œuvre et la « perception » qu'en ont les bénéficiaires. Ces aspects ont montré leur importance lors de l'expérience du CICE, crédit d'impôt que les directions des ressources humaines des entreprises n'ont pas toujours pris en compte au même titre que les exonérations de cotisations sociales, ou, du côté des prestations de solidarité, pour le RSA activité, dont l'assimilation à un dispositif d'aide sociale a rebuté nombre de travailleurs modestes, alors que la prime d'activité est davantage apparue comme une prestation sociale de droit commun.

⁴⁹ Voir par exemple Yannick L'Horty, Philippe Martin, Thierry Mayer, « Baisse de charges : stop ou encore ? », *Les notes du conseil d'analyse économique*, n°49, janvier 2019 ; Igas, « Présentation et analyse des évaluations portant sur les « allègements généraux » », in : IGF-Igas, *Évaluation de l'efficacité des mesures de réduction ou d'exonération de cotisations ou de contributions de sécurité sociale prévue par la loi organique n° 2022-354 du 14 mars 2022 relative aux lois de financement de la sécurité sociale*, mars 2023.

⁵⁰ Antoine Bozio et Étienne Wasmer, *Les politiques d'exonérations de cotisations sociales : une inflexion nécessaire*, Rapport final, octobre 2024.

⁵¹ Jérôme Gautié et Frédéric Lerais, *Politiques d'exonération sur les bas salaires : usages et effets potentiels*, Rapport final, Agence d'objectifs CFDT - Ires, janvier 2024.

3.3) Un débat politique dont l'accélération et la simplification percutent la conduite des évaluations et font perdre son sens à leur appropriation

Face à ce temps nécessaire et à cette complexité accrue des évaluations, directement liée ou non au renouvellement des méthodes quantitatives, on ne peut en regard qu'être frappé par l'accélération du débat médiatique, liée aux exigences combinées de l'information en continu et des réseaux sociaux, et auxquels les décideurs politiques sont souvent amenés à répondre « dans l'instant ». Cette accélération s'accompagne d'une « caricaturisation » des messages diffusés pour retenir l'attention, qui se traduit par l'oubli des *caveat*, des constats d'hétérogénéité et des nuances associées à toute bonne étude scientifique, et *a fortiori*, à des évaluations de politiques faisant appel à des registres pluriels.

Même lorsque les données d'études ne sont ni temporellement décalées ni d'une technicité insurmontable, un problème de communication existe à l'évidence autour de leurs résultats, sachant que le débat politique repose lui-même davantage sur les reprises médiatisées de ces travaux que sur leur contenu effectif. L'exemple évoqué précédemment de la pertinence des mécanismes d'indexation en matière de retraites en est une illustration, qui a conduit à la polarisation du débat sur la question plus « facile » de « l'âge de départ ».

Il existe donc une dissociation de plus en plus prononcée entre le temps et les conditions nécessaires aux processus d'évaluation et ceux qui prévalent dans les débats politique et médiatique. Cela se traduit dans les « logiques politico-administratives »⁵² par des décisions de transformation répétée, voire frénétique, des dispositifs, parfois à de simples fins « d'affichage » ou, pour le moins, sur des bases peu étayées. Cette tendance de fond percute les évaluations en cours, rend illusoire leur appropriation et nuit aux fondements mêmes de la démarche d'étude et d'évaluation, à savoir l'utilité qu'elle peut avoir, au-delà des seuls chiffrages budgétaires, pour les acteurs sociaux et les décideurs publics.

Ces constats doivent donc imposer une modestie certaine aux promoteurs des travaux académiques, mais ils impliquent aussi, au-delà de la prolifération d'appels incantatoires à « l'évaluation », une réflexion partagée renouvelée sur sa définition, ses méthodes, ses modalités et ses « arènes ». On court sinon le risque que ce type de démarche originellement à visée démocratique soit remplacé par le recours à des indicateurs mal maîtrisés issus de données brutes, retraités de façon plus ou moins transparente par l'intelligence artificielle, et, qui plus est, que ces éléments et les « *punchlines* » que l'on peut en déduire soient en eux-mêmes jugés suffisants pour apprécier et orienter l'action publique.

⁵² Patrice Duran, « Évaluation des politiques publiques : les leçons de l'expérience », *Revue française d'administration publique*, n°177, 2021/1.



SciencesPo

ABOUT OFCE

The Paris-based Observatoire français des conjonctures économiques (OFCE), or French Economic Observatory is an independent and publicly-funded centre whose activities focus on economic research, forecasting and the evaluation of public policy.

Its 1981 founding charter established it as part of the French Fondation nationale des sciences politiques (Sciences Po) and gave it the mission is to “ensure that the fruits of scientific rigour and academic independence serve the public debate about the economy”. The OFCE fulfils this mission by conducting theoretical and empirical studies, taking part in international scientific networks, and assuring a regular presence in the media through close cooperation with the French and European public authorities. The work of the OFCE covers most fields of economic analysis, from macroeconomics, growth, social welfare programmes, taxation and employment policy to sustainable development, competition, innovation and regulatory affairs.

ABOUT SCIENCES PO

Sciences Po is an institution of higher education and research in the humanities and social sciences. Its work in law, economics, history, political science and sociology is pursued through ten research units and several crosscutting programmes.

Its research community includes over two hundred twenty members and three hundred fifty PhD candidates. Recognized internationally, their work covers a wide range of topics including education, democracies, urban development, globalization and public health.

One of Sciences Po's key objectives is to make a significant contribution to methodological, epistemological and theoretical advances in the humanities and social sciences. Sciences Po's mission is also to share the results of its research with the international research community, students, and more broadly, society as a whole.

PARTNERSHIP

SciencesPo