



Notes et études socio-économiques

CENTRE D'ÉTUDES ET DE PROSPECTIVE

n° 36 – Juin 2012



Anne Le Roy, Guillaume Millot

- L'utilisation de la donnée chiffrée dans le pilotage et l'évaluation des politiques publiques
Le cas des politiques de développement rural

NESE n° 36, juin 2012, pp. 103-121

CENTRE D'ÉTUDES ET DE PROSPECTIVE

SERVICE DE LA STATISTIQUE ET DE LA PROSPECTIVE

Présentation

Notes et Études Socio-Économiques est une revue du ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt, publiée par son Centre d'Études et de Prospective. Cette revue technique à comité de rédaction se donne pour double objectif de valoriser des travaux conduits en interne ou des études commanditées par le ministère mais également de participer au débat d'idées en relayant des contributions d'experts extérieurs. Veillant à la rigueur des analyses et du traitement des données, elle s'adresse à un lectorat à la recherche d'éclairages complets et solides sur des sujets bien délimités. D'une périodicité de deux numéros par an, la revue existe en version papier et en version électronique.

Les articles et propos présentés dans cette revue n'engagent que leurs auteurs.

Directrice de la publication :

Fabienne Rosenwald, MAAF-SG-SSP, Chef du Service de la Statistique et de la Prospective

Rédacteur en chef :

Bruno Hérault, MAAF-SG-SSP, Chef du Centre d'Études et de Prospective

Secrétaire de rédaction :

Pierre Claquin, MAAF-SG-SSP-CEP, Chargé de mission

Comité de rédaction

Jean-Claude Teurlay, MAAF-SG-SSP, Adjoint au chef du SSP

Martin Bortzmeyer, MEDDTL-CGDD, Chef de bureau

Patrick Aigrain, FranceAgriMer, Direction Marché Études et Prospective

Frédéric Courleux, MAAF-SG-SSP-CEP, Chef du BEAE

Bruno Hérault, MAAF-SG-SSP, Chef du Centre d'Études et de Prospective

Aurélie Darpeix, MAAF-DGPAAT, Chargé de mission au BPCEC

Aurélien Daubaire, MEFI-DGT, chef du BEA (POLSEC 4)

Nathanaël Pingault, MAAF-DGPAAT, Chef du BSECC

Jean-Luc Pujol, INRA, Directeur Mission d'anticipation Recherche Société et Développement durable

Sylvain Rousset, IRSTEA, Ingénieur Chercheur

Julien Vert, MAAF-SG-SSP-CEP, Chef du BPSIE, Centre d'Études et de Prospective

Sébastien Treyer, IDDRI, Directeur des programmes

Tanocrède Voituriez, CIRAD, IDDRI

Pascale Pollet, MAAF-SG-SSP, Sous-directrice de la SDSSR

Composition : SSP - ANCD

Impression : SSP - BSS

Dépôt légal : à parution

ISSN : 2259-4841

Renseignements et diffusion : voir page 4 de couverture

L'utilisation de la donnée chiffrée dans le pilotage et l'évaluation des politiques publiques. Le cas des politiques de développement rural

Anne Le Roy¹, Guillaume Millot²

Résumé

Le recours au chiffre semble créer de la rigueur, de l'objectivité, de la connaissance et faciliter les comparaisons. Par conséquent, une évaluation de politique publique sans donnée chiffrée paraît à certains difficilement concevable. Toutefois, l'objectivité et la précision peuvent n'être qu'une impression puisque les chiffres sont des constructions basées sur une modélisation de la réalité. La simplification de cette dernière au travers d'un chiffre peut cacher des subtilités quant à la description des mécanismes d'action des politiques publiques. Si l'utilisation de données quantifiées est légitime en évaluation, toutes formes de données chiffrées et toutes utilisations de celles-ci ne sont pas équivalentes. D'où la question centrale de cet article : quelle place accorder à la donnée chiffrée dans une démarche évaluative ? Cette réflexion porte sur l'évaluation vue comme un outil de production de connaissances et de compréhension des politiques, de leur mise en œuvre et de leurs effets. Basé sur une analyse des pratiques d'évaluation des politiques de développement rural conduites sous la responsabilité du ministère de l'Agriculture, cet article vise à mettre en lumière les conditions optimales de production et d'utilisation de données en vue de leur utilisation dans des évaluations.

Mots clés

Évaluation, politique publique, développement rural, donnée chiffrée, quantification

Le texte ci-après ne représente pas nécessairement les positions officielles du ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt. Il n'engage que ses auteurs.

1. Maître de conférences en sciences économiques - Centre de Recherche en Économie de Grenoble (CREG).

2. Guillaume Millot était chargé de mission « évaluation territoriale et sociale » au Centre d'Études et de Prospective (CEP) du Ministère en charge de l'agriculture au moment de la rédaction de cet article.

Introduction

L'utilisation de données chiffrées crée une connaissance à laquelle est associée, assez naturellement, une idée de rigueur. Cette image de rationalité confère un caractère rassurant au chiffre qui tend alors à rendre l'argument qu'il soutient quasi-incontestable. La construction de connaissances, voire la démonstration, par la quantification, est ainsi régulièrement utilisée dans le pilotage, l'évaluation et la communication de l'action publique. Pour autant, l'impression d'objectivité et de précision émanant du chiffre peut faire illusion dans la mesure où celui-ci est une construction reposant sur une représentation plus ou moins schématique et plus ou moins consensuelle de la réalité. L'aspect concret d'un chiffre ne doit pas faire illusion : il est toujours le résultat d'une conceptualisation. Derrière toute donnée chiffrée se cache une théorie qui peut, un jour, s'avérer obsolète ou se trouver mise en défaut suite à l'évolution du contexte socio-économique ou à une progression de la connaissance scientifique. Dans le même temps, l'évaluation, en tant qu'outil de production de connaissances et de compréhension des politiques, est difficilement envisageable sans un recours à la donnée chiffrée. Pour autant, les utilisations qui en sont faites en évaluation sont multiples et n'ont pas toutes le même potentiel de création de connaissances. De plus, la conception du recours à la donnée chiffrée dans le cadre d'une évaluation tend à se centrer sur la seule notion d'indicateur. Il apparaît donc essentiel d'apporter un éclairage sur les différentes pratiques qui peuvent avoir cours en matière de mobilisation de données quantifiées en évaluation.

Pour ce faire, notre réflexion s'est appuyée sur les pratiques évaluatives des politiques de développement rural menées en France par le ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt (MAAF). En effet, ces politiques représentent un vivier important d'expériences en matière de mobilisation de la donnée chiffrée dans l'évaluation³.

Dans un premier temps nous dresserons un état des lieux des sources de données mobilisées et des modes de mise en forme et d'analyse de ces données (1). Cette première étape mettra en évidence deux conceptions de l'évaluation, qui divergent fortement par leur niveau de prise en compte de l'interaction entre politique et contexte d'intervention. La compréhension du contexte d'une intervention publique se faisant également *via* la mobilisation de données chiffrées, c'est plus largement la place de la donnée dans le pilotage de l'action publique qu'il convient d'interroger avant de formuler quelques recommandations visant la production de ces données (2).

1. Mobilisation et utilisation de la donnée chiffrée dans l'évaluation des politiques de développement rural

Dans cette première partie, nous essaierons de répondre successivement aux questions suivantes : sur quelles sources de données l'analyse et l'évaluation des politiques de développement rural s'appuient-elles (1.1) ? De quelles analyses et de quels traitements ces données peuvent-elles faire l'objet (1.2.) ? Comment ces données chiffrées s'inscrivent-elles dans une démarche évaluative répondant à un questionnement évaluatif explicite (1.3) ?

3. L'illustration des propos reposera principalement sur les expériences d'évaluations de types *ex post* et à mi-parcours. Les évaluations *ex ante* relèvent d'une démarche différente de chiffrage *a priori* d'objectifs à atteindre. Le questionnement à mettre en œuvre et les usages de la donnée chiffrée diffèrent donc des autres types d'évaluation.

1.1. Les sources de données

Un panorama des données régulièrement mobilisées dans le cadre des pratiques d'évaluation des politiques de développement rural au MAAF permet d'en identifier trois sources : (1.1.1.) des données d'enquêtes produites par les services de statistique publique, (1.1.2.) des données de suivi administratif et budgétaire et (1.1.3.) des données d'enquêtes réalisées par les prestataires de l'évaluation. Plus que les institutions chargées de la production de ces données, ce sont les modes de production de ces données et l'articulation de cette production avec l'activité d'évaluation qui permettent de tracer les grandes lignes de partage entre ces catégories. Certaines données sont produites spécialement à l'occasion de travaux d'évaluation, tandis que d'autres préexistent et assurent des fonctions supplémentaires. De plus, certaines données visent l'exhaustivité, tandis que d'autres sont produites à partir d'un échantillonnage.

1.1.1. Les données d'enquêtes produites par les services de statistique publique

Les données fournies par ces services sont produites indépendamment des politiques mises en œuvre. Elles visent à suivre l'évolution de la situation socio-économique et environnementale du pays et des différents secteurs d'activités⁴. Elles sont donc utiles pour caractériser les contextes de l'action publique. Deux modes de production de ces données existent : les enquêtes par sondage qui reposent sur le tirage d'un échantillon – enquête « structures », enquête « pratiques culturelles », enquête « bâtiments d'élevage », réseau d'informations comptables agricole (RICA) – et les recensements exhaustifs tels que le recensement agricole. Les recensements sont intéressants dans la mesure où ils couvrent l'ensemble des populations de bénéficiaires et de non-bénéficiaires. En revanche, selon la fréquence de leur mise à jour, ils peuvent devenir obsolètes dans le cadre de l'évaluation de programmes courts. De plus, un éventuel calage des calendriers (production statistique et évaluation de politique) est difficilement envisageable car une même source statistique peut alimenter l'évaluation de plusieurs politiques dont les calendriers de conception et de mise en œuvre ne sont pas en phase. Enfin, si un recensement est exhaustif vis-à-vis des personnes ou des structures soumises à enquêtes, ce qui assure par la suite des tirages d'échantillons de qualité pour les enquêtes par sondage et une connaissance générale des profils de l'ensemble des bénéficiaires potentiels d'une politique, tous les champs thématiques couverts par ce type d'enquêtes ne font pas l'objet du même niveau d'approfondissement.

Basées sur des échantillons plus restreints et couvrant des champs thématiques clairement définis, les enquêtes par sondage peuvent faire l'objet de mises à jour plus fréquentes et permettent de recueillir des informations précises sur le fonctionnement des systèmes observés. En revanche, si elles fournissent des données globalement représentatives de l'ensemble de la population, l'échantillon est parfois trop restreint et ne permet pas d'avoir une bonne représentativité des sous-populations « bénéficiaires » et « non-bénéficiaires ». Il est à noter qu'une partie des enquêtes par sondage est pilotée au niveau européen (enquête structure, RICA) et que les États membres n'ont pas toute latitude pour adapter les critères d'échantillonnage aux spécificités nationales. Par exemple, le seuil relatif à la dimension économique des exploitations à partir duquel l'échantillonnage du RICA est conçu laisse de côté de nombreuses exploitations de taille très modeste, notamment dans les nouveaux États membres de l'Union européenne. Dans le cadre d'un chantier évaluatif, les données de recensement permettent d'avoir une bonne vision générale de la situation en début et en

4. Nous faisons ici référence au secteur agricole élargi à sa transformation par les industries agroalimentaires en aval, deux domaines largement couverts par les enquêtes statistiques du MAAF. Le secteur forestier ainsi que les autres secteurs de l'économie rurale comme les services ou l'artisanat sont moins décrits par la statistique agricole mais peuvent être couverts par d'autres sources statistiques.

fin de période de programmation, alors que les données d'enquêtes par sondages permettent d'apporter les premiers éléments de réponse relatifs aux évolutions constatées et réellement imputables à la politique évaluée.

1.1.2. Les données de suivi

Les données de suivi sont avant tout construites pour aider au pilotage et s'assurer du bon déroulement de l'action publique tant sur le plan budgétaire (suivi des engagements et des paiements, consommation des crédits bien répartie sur l'ensemble de la période de programmation, compte rendu vis-à-vis de la Commission européenne de la consommation des enveloppes budgétaires selon le programme établi), que sur le plan administratif (vérification des critères d'éligibilité aux aides), sans oublier les aspects stratégiques (par exemple, données relatives à l'âge ou au sexe des bénéficiaires, en lien avec les préoccupations européennes d'égalité des chances). Par construction, il s'agit de données relatives à la gestion et aux réalisations concrètes des différentes mesures dans une logique de suivi. Elles ne peuvent donc pas, seules, répondre directement aux besoins de l'évaluation qui concerne également l'impact global d'un programme. La production de ces données est concomitante à la réalisation de la politique. La possibilité de les utiliser ultérieurement dans le cadre d'une évaluation n'est pas toujours clairement identifiée par les acteurs de la chaîne administrative et par les bénéficiaires. Il arrive donc qu'elles ne soient pas systématiquement saisies lors de la constitution ou de l'instruction des dossiers de demande d'aide, soit parce que leur production n'a pas été suffisamment anticipée, soit parce qu'elle fait l'objet d'une priorité moindre par rapport à des tâches de gestion ou de paiement des aides. Ces données peuvent parfois s'avérer incomplètes lors de leur mobilisation au cours d'une évaluation. Leur production est assurée par les institutions et services chargés de la mise en œuvre de cette politique mais également par d'autres intervenants sur le même secteur d'activité : Mutualité Sociale Agricole, Agence Bio, Crédit Agricole, Groupama, etc. Ces données ne concernent que les bénéficiaires et c'est leur principal défaut. Les non-bénéficiaires ou les porteurs de dossiers rejetés sont rarement intégrés dans ces bases. Ces données ne couvrent pas davantage les questions relatives aux effets indirects ou aux impacts de la politique.

1.1.3. Les données d'enquêtes réalisées par les prestataires de l'évaluation

Lors de la réalisation d'un chantier d'évaluation, en plus de mobiliser les données existantes issues du suivi et des statistiques publiques, les évaluateurs sont régulièrement amenés à produire des données nouvelles. Ces données sont produites au moyen d'enquêtes réalisées en face à face avec des bénéficiaires, des non-bénéficiaires et des agents chargés de la mise en œuvre des mesures évaluées, mais aussi grâce à des questionnaires envoyés par courrier ou par email. Dans ce cas, la question de la représentativité des échantillons se pose et la liberté pour les destinataires de ne pas répondre aux questionnaires peut introduire un biais significatif.

Les données ainsi produites sont en général de nature qualitative. Toutefois, si le temps et les moyens dont dispose l'évaluateur le permettent, il est possible de produire des données quantifiées, en lien plus ou moins direct avec les éléments fournis par l'analyse qualitative. Les données à caractère qualitatif permettent d'aborder des thématiques non couvertes par la statistique habituelle et d'approfondir certains points de nature sociologique, comme la perception de la politique évaluée par les publics cibles ou les jeux d'acteurs. Enfin, les données d'enquêtes réalisées par les évaluateurs permettent d'affiner les résultats obtenus à partir des données statistiques et de suivi, en mettant au jour des mécanismes par lesquels la politique a produit ses effets, attendus ou inattendus.

Si produire des données est une étape nécessaire, elle n'est pas pour autant suffisante. Il importe de donner du sens aux chiffres en recourant à des traitements et analyses appropriés.

1.2. Le traitement et l'analyse des données mobilisées

La mobilisation de données quantitatives dans le cadre d'une démarche d'évaluation ne se résume pas à un recours à des données statistiques et ne peut être réduite au remplissage de tableaux d'indicateurs. Elle implique une diversité de traitements et de mises en forme, pour la plupart non spécifiques à l'évaluation et utilisés dans le cadre d'autres démarches d'analyse scientifique ou politique (1.2.1.). Elle montre l'importance de la définition des périmètres géographiques, temporels et thématiques des analyses conduites (1.2.2.).

1.2.1. Diversité des traitements des données

La principale boîte à outils utilisée en évaluation de politiques est la statistique descriptive. De simples moyennes et écart-types peuvent faire apparaître bien des résultats par le biais de comparaisons entre échantillons. Le choix des échantillons sur lesquels ces statistiques seront calculées dépend de la question posée : il peut par exemple reposer sur des considérations d'ordre géographique pour des comparaisons entre zones de plaine et de montagne sur la question de l'Indemnité Compensatoire de Handicap Naturel, de taille d'exploitation entre petites, moyennes et grandes, ou encore d'orientation de production, pour comparer les effets d'une mesure d'investissement entre élevage et culture. La présentation de ces statistiques descriptives peut faire appel à de nombreuses mises en formes graphiques (tableaux, courbes d'évolution, histogrammes, etc.) et elle doit s'accompagner de commentaires apportant une aide à l'interprétation des chiffres présentés.

Le recours aux outils d'analyse spatialisée est possible mais n'est pas systématique en évaluation. Il dépend de l'étendue et de la diversité du territoire concerné par la politique évaluée ainsi que du champ thématique couvert par cette politique. Ceci est d'autant plus vrai que l'action en question est territorialisée. La cartographie permet alors d'identifier certains phénomènes spatiaux liés entre eux et permet d'établir graphiquement des corrélations. Toutefois, elle ne dispense pas du besoin d'établir les liens de causalité, s'ils existent, susceptibles d'expliquer les corrélations ainsi observées. Le deuxième pilier de la PAC en est un bon exemple. En effet, cette politique accompagne les évolutions des structures d'exploitation en partie induites ou rendues nécessaires dans le cadre de la spécialisation des régions. Naturellement, ceci implique une prise en compte de la diversité des contextes territoriaux dans sa conception, sa mise en œuvre et son évaluation. L'existence de dispositifs zonés (zones défavorisées, zones Natura 2000 et plus récemment les périmètres de captage prioritaires définis suite au Grenelle de l'environnement) au sein de cette politique appuie également ce constat. La nécessité de fournir des analyses spatiales de la mise en œuvre des politiques de développement rural a conduit à la naissance de l'observatoire des politiques communautaires de développement rural (ODR), sous la triple impulsion du MAAF⁵, de l'agence de services et de paiement (ASP) et de l'Institut national de la recherche agronomique (INRA).

La réalisation d'évaluations implique une capacité à anticiper les effets d'une politique. Cette démarche peut reposer sur des outils de modélisation économique et sur le calcul

5. Plus particulièrement des services en charge de la statistique, de l'évaluation et de la gestion et du règlement européen de développement rural.

économique qui comprend les analyses coûts-avantages, les analyses coûts-efficacité et les analyses multicritères. L'intérêt de ces outils d'évaluation réside dans leur capacité à objectiver les coûts et les avantages pour une société, liés à la mise en place d'une politique. En effet, ces outils incluent tous une étape de quantification des conséquences de la politique, qu'ils fassent ou non l'objet d'une monétisation. Ils mobilisent un volume significatif de données chiffrées, relatives au contexte et aux effets identifiés *a priori* ou *a posteriori* (Le Roy, 2009). En revanche, le manque de disponibilité de certaines données⁶ ou la difficulté technique de l'estimation monétaire⁷ peut conduire à négliger la prise en compte de certains coûts ou avantages dans le calcul. Cela entraîne un biais significatif dans le résultat obtenu.

Par ailleurs, l'évaluation des politiques publiques peut faire appel à des outils mathématiques d'estimation des effets propres, dérivés de la recherche médicale. Ces outils, peu utilisés pour l'évaluation des politiques de développement rural, ont toutefois été testés « grandeur nature » pour estimer les effets propres de certaines mesures agro-environnementales (MAE)⁸. S'ils présentent l'avantage d'offrir une quantification fine de la proportion des effets observés réellement attribuables à la politique observée, ils sont toutefois très gourmands en données chiffrées. Les outils d'estimation des effets propres sont intéressants dans le cadre de mesures simples dont les mécanismes d'action socio-économiques ont été révélés. Ils peuvent alors être facilement modélisés et soumis à des tests économétriques. En revanche, ils peuvent rencontrer des difficultés liées à l'échantillonnage réalisé *a posteriori*⁹. Le recours à l'évaluation expérimentale permet de lever cette difficulté. Toutefois, ces deux outils, estimation *ex post* des effets propres et démarche expérimentale, peuvent révéler un lien de causalité mais n'expliquent pas les mécanismes sociaux, économiques et environnementaux à l'origine de cette causalité (Labrousse, 2010). De plus, ils trouvent une limite supplémentaire dans la nature de l'action évaluée. En effet, ils ne permettent pas l'évaluation de politiques et de programmes complexes, multifactoriels et multicausaux, ou encore d'actions intégrant des effets de diffusion (impacts sur les non-bénéficiaires).

Cette étape relative au traitement du chiffre dans le cadre d'une démarche évaluative fait émerger trois questions qui affecteront les données à mobiliser et leurs modalités de traitement : la première relative à l'échelle géographique, la seconde liée à la dimension temporelle des phénomènes étudiés et la dernière concernant le champ thématique couvert.

1.2.2. Échelles spatiale et temporelle, champs thématiques couverts

À quelle échelle géographique l'analyse des données doit-elle être réalisée ? L'analyse des effets d'une politique au niveau du bénéficiaire, qui mobilise des données microéconomiques, produit une information différente de l'analyse des effets à un niveau plus agrégé. L'évaluation *ex post* du programme de développement rural national (période de programmation 2000-2006), portant sur la mesure d'aides à l'investissement dans les industries agroalimentaires, a révélé des effets contradictoires selon que l'analyse se situait au niveau

6. Par exemple, l'absence de données consolidées mesurant l'impact des produits phytosanitaires sur la santé des professionnels, ne permet de prendre en compte l'effet potentiel d'une mesure visant à réduire leur utilisation à des fins de santé publique.

7. La difficulté à chiffrer sous forme monétaire la valeur de la biodiversité rend donc quasiment impossible l'estimation monétaire des impacts des mesures agroenvironnementales.

8. L'évaluation des effets propres des MAE a fait l'objet d'une convention d'appui méthodologique avec le Centre clermontois de l'IRSTEA (Chabé-Ferret et Subervie, 2009).

9. L'évaluation des effets propres des MAE, et plus particulièrement de la prime herbagère agri-environnementale (PHAE), a effectivement été confrontée à ce problème. En raison de la disproportion entre les populations de bénéficiaires et de non-bénéficiaires, la démarche d'appariement n'a pu être menée qu'en tronquant largement l'échantillon des bénéficiaires. L'impact n'a été, au final, pris en compte que sur une partie des bénéficiaires, limitant de ce fait la portée du résultat.

de l'entreprise ou d'un territoire plus large. En effet, l'investissement réalisé s'est traduit dans certaines entreprises aidées par des créations d'emplois et des gains de parts de marché. Dans le même temps, certaines entreprises concurrentes, voisines et non aidées, ont connu une trajectoire inverse, illustrant l'effet d'éviction d'une action publique. Il en résulte un bilan nul à l'échelle du territoire englobant ces entreprises (Leenhardt, 2009).

De façon similaire, la question du temps de l'action publique émerge fréquemment : à quel pas de temps devons-nous raisonner afin de comprendre et mesurer les transformations générées par l'intervention publique ? Les horizons temporels peuvent différer assez nettement selon que l'on se place du point de vue du porteur de projet, du responsable politique ou administratif ou encore de l'évaluateur. De plus, les calendriers des uns et des autres peuvent être en décalage entre eux, mais aussi avec le temps nécessaire pour que l'intervention publique produise ses effets et que leur quantification soit possible. Par exemple, la continuité entre les périodes de programmation de la politique européenne de développement rural et les calendriers d'évaluation associés ne laissent pas le temps aux effets d'un programme de prendre toute leur ampleur avant leur évaluation finale et le lancement du programme suivant. Si la donnée de suivi est généralement disponible en fin de période de programmation, la donnée statistique ne l'est pas forcément.

L'inventaire des sources de données à utiliser dans le cadre d'une évaluation nous confronte aussi à la question du champ thématique couvert par les données à rassembler, qui peut dépasser le secteur d'intervention de la politique concernée. Ainsi, une politique sectorielle agricole a des impacts en aval, depuis les industries agroalimentaires et autres secteurs d'activité utilisant des produits agricoles comme facteurs de production, jusqu'aux consommateurs. Les impacts du deuxième pilier de la PAC au niveau des consommateurs n'ont pas, à notre connaissance, été étudiés à ce jour. Une telle étude ferait appel à des données qui dépassent largement le domaine agricole pour inclure la consommation des ménages. De même, les aides à la diversification des activités agricoles vers l'artisanat ou l'agrotourisme, parfois montrées du doigt comme des sources de distorsion de concurrence puisque favorisant les agriculteurs au détriment des artisans et des professionnels du tourisme, peuvent faire appel à des sources de données non agricoles, si l'on veut mettre en évidence des effets d'éviction potentiels.

1.3. Le recours au chiffre dans les démarches évaluatives

Si le rôle de la donnée dans les démarches évaluatives doit être explicité (1.3.1.), il ne faut pas oublier qu'elle n'est pas une fin en soi mais un outil d'aide à la démonstration (1.3.2.).

1.3.1. La délicate mobilisation des données quantifiées dans l'évaluation des politiques

Au regard des pratiques de conception et d'évaluation des politiques (figures 1.a et 1.b), dans quelle mesure les valeurs prises par les indicateurs résultent-elles de la double influence du contexte et de la politique évaluée ?

Figure 1.a - Indicateurs classiques de suivi et d'évaluation dans le cycle d'une politique publique (CCSE)

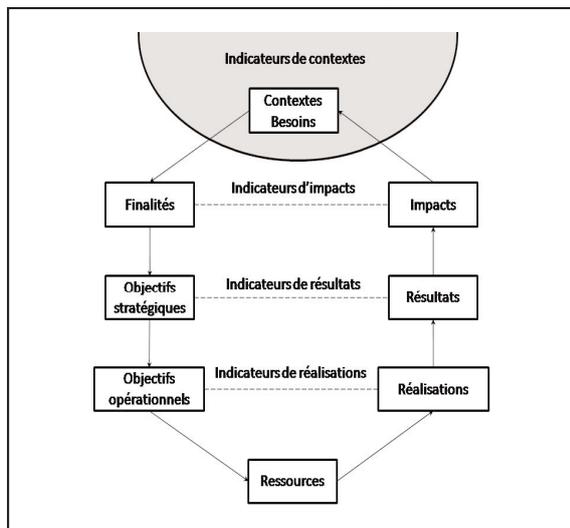
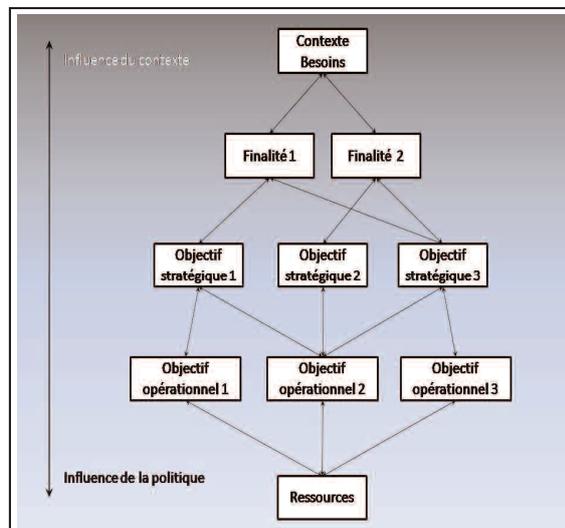


Figure 1.b - Influences du contexte et de la politique sur la donnée chiffrée



De par son statut de politique européenne, le deuxième pilier de la PAC s'accompagne d'un Cadre Commun de Suivi et d'Évaluation (CCSE) qui définit un ensemble de questions évaluatives et un tableau d'indicateurs de différents niveaux : indicateurs de moyens, de réalisations, de résultats et d'impacts, indicateurs de références et indicateurs de contexte. Ce cadre d'évaluation préconise également la formalisation d'une théorie d'action¹⁰, parfois appelée « graphe d'objectifs de la politique », avant de procéder à son évaluation. Outre la connaissance créée par cet outil, il est un élément du référentiel d'action de la politique, tout comme l'ensemble des questions évaluatives. Ce référentiel constitue une liste de vérification des impacts de la politique et permet de s'assurer que les effets – réalisations, résultats et impacts – correspondent bien aux objectifs initialement fixés pour cette politique (cf. figure 1.a). De cette correspondance entre objectifs et effets de la politique, il découle que les indicateurs préconisés pour l'évaluation dans le CCSE peuvent aussi servir à fixer ces objectifs. À chaque niveau d'objectif/effets peut être théoriquement associé un système d'indicateurs le quantifiant. Cette approche du suivi et de l'évaluation se traduit dans la pratique par le souci de présenter des tableaux d'indicateurs, dans une démarche où la forme, c'est-à-dire le tableau préconstruit et rempli, peut prendre le dessus sur le fond, à savoir l'information potentiellement contenue dans ces indicateurs et la question évaluative à laquelle ils sont censés apporter un éclairage.

Cette représentation du cycle d'une politique publique (figure 1.a) réduit la complexité de la réalité de l'action publique. Le système de quantification qui lui est associé se trouve de fait entaché de défauts. D'une part, la représentation linéaire de la théorie d'action ne permet pas de rendre compte de la dimension « programme » d'une politique composée de nombreuses mesures qui peuvent être combinées pour répondre de façon articulée à un ensemble d'objectifs de niveaux hiérarchiques divers (approche systémique et non linéaire). L'hégémonie de l'indicateur dans l'évaluation séquence et découpe artificiellement la politique publique en programmes, actions et sous-actions. La complexité des politiques de développement rural est alors mieux appréhendée par la figure 1.b. De plus, cette linéarité de la conception des politiques tend à négliger les interactions entre contexte et politique, présentes à chaque niveau de la hiérarchie des objectifs. À l'inverse, la figure 1.b illustre bien la double influence du contexte et de la politique sur la valeur prise par les données chiffrées.

10. Ce terme désigne la chaîne de causalité reliant les actions mises en œuvre à leurs impacts attendus.

La construction même du système de quantification et d'indicateurs ne peut ignorer cet état de faits. À titre d'illustration, comment interpréter le nombre d'hectares cultivés en agriculture biologique, utilisé comme indicateur de réalisation par le CCSE ? Sachant que la production biologique existait et évoluait avant que ne soit mise en place une politique publique d'incitation, il est impossible d'imputer l'intégralité du résultat à la seule mesure d'incitation à la conversion au mode de production biologique, plutôt qu'à l'évolution plus général du contexte d'essor de ces productions. L'influence du contexte passé et présent affecte la valeur d'un indicateur de réalisation théoriquement destiné à isoler l'effet imputable à l'action publique.

Cette interaction constante entre politique et contexte se traduit jusque dans la production de données statistiques. À titre d'exemple, la modernisation des exploitations agricoles, au sortir de la guerre, a conduit à la disparition progressive de la traite manuelle, rendant finalement obsolète la question du recensement agricole relative à l'existence d'une installation de traite mécanique sur l'exploitation.

Pour clarifier la description du système d'indicateurs du CCSE, dans son état actuel, il est essentiel de le regarder à la lumière de la figure 1.b. Cela conduit à distinguer :

- des indicateurs consubstantiels à la politique (moyens et réalisations), qui mesurent des effets directs dont le lien de causalité avec la politique peut être établi relativement facilement. Ces indicateurs relèvent plus du suivi que de l'évaluation.

- des indicateurs davantage liés au contexte (référence, impacts et résultats) qui peuvent contribuer à la mesure des effets indirects de la politique et des modifications de contexte, à une échelle plus agrégée. Ces indicateurs, et plus particulièrement les indicateurs de résultats et d'impacts, font aujourd'hui l'objet de discussions au niveau européen. En effet, la notion même d'indicateur d'impact est remise en question car les impacts d'une politique ne sont pas directement observables et la mise au jour de ses effets indirects passe par l'analyse fine de la dite politique.

Par ailleurs, vu de l'extérieur, le chiffre apparaît comme au centre du processus d'évaluation des politiques. Cette vision déformée de l'évaluation est due à l'accent mis sur la notation et au recours de plus en plus systématique aux indicateurs de performance et de résultats. Si ces notions d'indicateurs, qui se situent principalement au niveau microéconomique, ont un rôle à jouer dans le domaine des politiques publiques, ils ne sont pas les seuls à intervenir dans la formulation d'un jugement évaluatif. Ils sont complétés par une recherche de quantification des impacts au niveau macroéconomique, tenant compte d'effets de système, d'éviction ou de diffusion. La question qui se pose en évaluation est bien celle des effets réellement imputables à la politique, sachant que les effets d'aubaine et d'éviction sont inévitables. Une des conséquences liées à la conception linéaire des politiques présentée en figure 1.a se traduit ainsi par la tentation de vouloir déduire les impacts macroéconomiques des politiques à partir des observations faites au niveau microéconomique. Cette tentation conduit la Commission européenne à chercher l'agrégation des indicateurs d'impacts à l'échelle de l'Union, parfois au mépris des règles régissant l'agrégation de données relatives aux phénomènes sociaux, économiques et environnementaux. Par exemple, l'indicateur d'impact relatif à la biodiversité ne peut être agrégé à l'échelle européenne, la biodiversité étant une notion fortement dépendante de l'échelle d'observation.

1.3.2. La donnée chiffrée : un outil d'aide à la démonstration

Si des éléments de quantification de l'efficacité de la politique importent en évaluation, la compréhension des mécanismes permettant cette efficacité est primordiale afin de juger de sa pertinence et de sa reproductibilité dans d'autres situations. Dans cette perspective, la démarche d'évaluation fondée sur un questionnement évaluatif s'impose. Or, on assiste actuellement

Encadré 1- Exemple de mobilisation de données chiffrées dans la construction d'une réponse à une question évaluative : le cas de l'aide à l'installation

Lors de l'évaluation *ex post* du PDRN, les évaluateurs en charge de la question des structures d'exploitation devaient juger de l'efficacité de l'aide à l'installation dans l'activité agricole. Pour cela, ils ont eu recours à une démarche de mobilisation progressive de données statistiques.

Tout d'abord, ils ont mis à plat, à l'aide des données de suivi et des données de la MSA, l'évolution du nombre d'installations annuelles totales, aidées et non aidées, sur la période 2000-2006. La comparaison avec les chiffres des années 1990 a mis en évidence le rôle complémentaire de l'aide au départ en pré-traité dans les dynamiques d'installation. La comparaison entre zones de montagne et zones de plaine a d'abord montré une baisse moins marquée du rythme des installations dans les premières. À l'aide des données d'en-

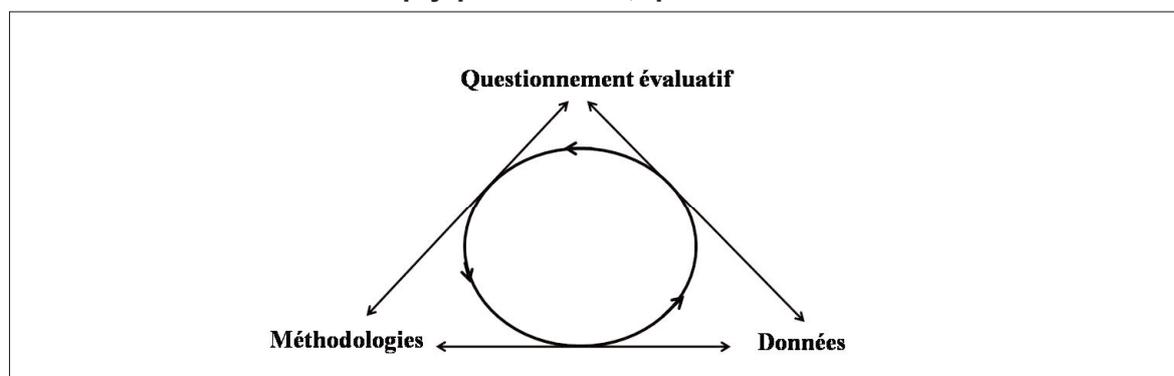
quête « structure », les évaluateurs ont vérifié que cette baisse moins rapide n'était pas due à une moindre restructuration des exploitations de montagnes. Afin d'identifier les freins à l'installation, les évaluateurs se sont par ailleurs livrés à une analyse approfondie des coûts d'installation en les détaillant par poste : foncier, bâtiment, cheptel, équipement, etc. Ils ont montré que l'augmentation des coûts d'installation était principalement due au coût des bâtiments. En établissant le ratio entre montant de l'aide et coût d'installation, ils ont également montré que celui-ci était plus grand en zones de montagne qu'en zones de plaine. En mobilisant des données d'enquêtes du Centre national pour l'aménagement des structures des exploitations agricoles (CNASEA), les évaluateurs ont également rappelé la pérennité plus grande des installations aidées par rapport aux installations non aidées.

à un phénomène de normalisation des pratiques d'évaluation des politiques de développement rural, où la donnée chiffrée prime même sur le questionnement évaluatif. Certes la donnée chiffrée et les indicateurs doivent être mobilisés en vue de produire un jugement et d'étayer l'argumentation du résultat final, mais la formulation d'un jugement évaluatif est avant tout de nature qualitative et ne nécessite pas systématiquement d'être quantifiée.

L'expérience acquise au sein des services compétents du MAAF (Centre d'études et de prospective) en matière d'assistance à maîtrise d'ouvrage sur les chantiers d'évaluation, met en avant trois points fondamentaux : 1) le questionnement évaluatif, 2) les données quantitatives et 3) les méthodes. Ce triptyque¹¹ est schématisé en figure 2.

Figure 2 - Le triangle évaluatif

L'évaluation : le triptyque données, questionnement et méthodes



11. Ce triangle n'est pas propre à la démarche évaluative. Il est également partagé par la majorité des disciplines scientifiques.

Il n'y a pas d'unicité dans les relations entre données et méthodes, entre données et questionnements et entre méthodes et questionnements. Aucune bijection ne peut être établie. À chaque chantier d'évaluation correspond un questionnaire nouveau, des données nouvelles et des méthodes nouvelles s'appuyant sur différentes disciplines scientifiques, parmi lesquelles l'économie tient une place importante.

Aussi, mobiliser des données dans une évaluation c'est avant tout réaliser un aller-retour entre les données (brutes ou mises en forme), la méthode retenue et le questionnaire évaluatif, en prenant en compte les contraintes qui lient méthodes, production et traitement des données. Une question évaluative appelle en première lecture une batterie de données pour y répondre. Les résultats issus de l'analyse de ces données sont susceptibles de soulever une chaîne de questions dont les réponses permettent de lever des ambiguïtés causales : les changements observés sont-ils dus à la politique évaluée et/ou à un changement structurel ou conjoncturel de contexte ? Dans quelles proportions ? Répondre à ces questions ne peut se faire sans mobiliser des données chiffrées permettant de construire un argumentaire solidement étayé, susceptible de créer une connaissance opérationnelle et théorique permettant d'accroître l'utilité des évaluations ainsi conduites. Cela nous amène à interroger plus largement, la place du chiffre dans la construction de connaissances et le pilotage de l'action publique

2. Comment améliorer les interactions entre la donnée chiffrée et la démarche évaluative ?

L'évaluation étant « une démarche collective de construction de jugements pratiques » (Perret, 2009), son déroulement s'apparente à un processus mobilisant des informations chiffrées et non-chiffrées en vue de produire un jugement socialement légitime (Millot, 2010). Dans cette optique, il importe non seulement de proposer une mesure de l'impact des actions publiques mais aussi de révéler les mécanismes qui les génèrent. La construction de ces données, dont il faut d'abord rappeler qu'elles sont le fruit d'un processus cognitif (2.1.) s'inscrit dans une démarche évaluative afin qu'elle soit productrice de connaissances (2.2.).

2.1. La construction de données chiffrées : un processus cognitif

La diversité du vocabulaire et des usages ne doit pas faire oublier qu'un chiffre est le fruit d'une co-production impliquant divers acteurs (1.1.1.). Il résulte toujours d'un processus cognitif que l'utilisateur doit maîtriser pour éviter tout mésusage de la donnée (1.1.2.).

2.1.1. Le chiffre : des usages divers, mais toujours une construction

Que l'on parle de donnée, de variable, de statistique, de descripteur ou d'indicateur, il s'agit de transformations de la réalité sous forme de « mise en statistiques ». Qu'elles soient brutes et/ou élaborées, ces données, combinées entre elles, informent sur le contexte, quantifient un effet ou une cause, démontrent ou réfutent un lien de causalité tout en fournissant des intuitions et en révélant des faits qui interpellent. La mise à jour de liens de causalité, leur quantification et la révélation de faits surprenants font bien partie des objectifs de la démarche d'évaluation des politiques publiques.

Le chiffre n'est pas, en effet, doté d'une seule fonction pour un seul usage. Il est au contraire multidimensionnel, ce qui doit conduire tout utilisateur à préciser, pour chaque usage, la fonction qu'il entend activer. De cette façon, les données finalement sélectionnées et utilisées seront en phase avec les fonctions recherchées. La production des données est donc fréquemment soumise à la nécessité de devoir répondre à un maximum de besoins parmi ceux formulés par l'ensemble des utilisateurs potentiels. Dès lors, préciser la posture adoptée est indispensable et doit être explicite, faute de quoi la donnée ainsi élaborée pourrait ne pas produire l'information recherchée.

Dans le cadre d'une analyse socio-éco-environnementale, il s'agit d'articuler un ensemble d'éléments et d'outils quantitatifs et qualitatifs afin d'accroître les capacités d'analyse des mécanismes d'une politique (Le Roy, Offredi, 2011). La production de données ne peut donc résulter d'une procédure, c'est-à-dire de l'application d'un protocole prédéfini et standardisé, mais d'un processus collectif relevant de conventions entre acteurs pour lesquelles il n'existe pas de consensus *a priori* (Gadrey, 2006) et reflétant ainsi le caractère réflexif de la mesure (Perret, B. 2011). En bref, quantifier c'est « convenir pour mesurer » (Desrosières, 2008). La question des zonages illustre le besoin de dépasser les procédures et cadres existants pour faire émerger de nouvelles conventions à partir desquelles la quantification peut déboucher sur des mesures ayant du « sens », donc une utilité : les zones prioritaires, ou à enjeux spécifiques comme les espaces Natura 2000, correspondent rarement aux subdivisions géographiques administratives prises en compte par les statisticiens. Il s'agit dès lors de mettre en place un dispositif d'observation à des échelles suffisamment fines pour pouvoir les agréger en territoires pertinents. Cela revient à mettre en place des outils suffisamment souples pour répondre à des besoins d'observation territoriale différenciés.

Il est donc nécessaire de ne pas déconnecter les indicateurs ainsi construits du système étudié et des cadres conceptuels à partir desquels ce système est élaboré. Faute de quoi le risque de mésusage s'accroît.

2.1.2. Un usage avisé de la donnée conditionné par la maîtrise de sa construction

L'évolution des contextes institutionnels et décisionnels (construction européenne et décentralisation, formes renouvelées de l'action publique) est à l'origine de la multiplication et de la diversification des demandes de données chiffrées. Dans le même temps, les objectifs de rationalisation des politiques, la pratique du *benchmarking* et la culture du résultat, couplés à la montée en puissance d'une forme de contractualisation des projets politiques, favorisent la multiplication des indicateurs, ce qui n'est pas sans risque, notamment celui du recours au chiffre comme simple habillage ou comme outil marketing destiné à attirer des financements devenus plus rares. Le rôle de l'État dans le fonctionnement de l'économie évoluant, les demandes de données chiffrées changent et gagnent en importance.

Or, les concepts, méthodes et outils valables au niveau national peuvent perdre de leur pertinence à l'échelle locale et inversement ; d'autant que les moyens matériels des statistiques nationales ne s'adaptent pas nécessairement aux changements d'échelle. Par ailleurs, les conventions ayant permis de faire consensus n'ont souvent de sens qu'au niveau où elles ont été pensées (Gadrey, 2003). De même, la validité temporelle d'un indicateur peut poser problème au fil de l'évolution du contexte de la société, des connaissances scientifiques et des objectifs des politiques. Et la façon dont la donnée est appropriée peut évoluer dans le temps, conduisant à la détourner des usages initiaux pour lesquels elle a été créée. L'utilisation qui a été faite du PIB, comme instrument de mesure de la performance économique d'une région, illustre tout à fait ce type de « dérive ». À l'origine, dans les années 1950, destiné à alimenter en données les modèles macroéconomiques keynésiens et faire

des prévisions dans le cadre des plans de l'après-guerre, le PIB mesure la production marchande et celle des administrations au niveau national. Mais, victime de son succès, il s'impose peu à peu dans les représentations, pour un autre usage : la mesure du « bien-être » d'un pays ou la richesse d'une région. Or, cet agrégat ne peut rendre ce service car il n'a pas été conçu pour cela. Au final, les critiques dont a fait l'objet le PIB, ou plutôt les usages qui en sont faits, débouchent sur des réflexions dont l'objectif est la construction de nouveaux indicateurs pour prendre en compte la richesse créée mais non valorisée par la sphère strictement marchande (travail domestique par exemple) ainsi que les externalités (atteintes à l'environnement et inégalités sociales)¹². *In fine*, cela revient à faire émerger de nouvelles conventions à partir desquelles de nouveaux processus de quantification sont mis en place afin de mesurer la richesse et le bien-être social et environnemental ; sachant que ce qui « compte est ce qui est compté » (Viveret, 2002).

Pour chaque réalité que l'on cherche à appréhender par un indicateur, plusieurs angles d'approche sont possibles. Prenons l'exemple de l'analyse microéconomique de l'activité d'une exploitation agricole. Afin de comparer des exploitations entre elles, cette activité doit-elle être rapportée à la Surface Agricole Utile (SAU), au cheptel, au volume produit, à la valeur ajoutée réalisée, aux nombres d'emplois induits, etc. ? Chacune de ces options produit une information différente, éclairant une même situation sous plusieurs angles. Elles sont par conséquent susceptibles d'élargir le champ des possibles en matière d'intervention publique. Le choix d'un dénominateur est par ailleurs souvent révélateur d'un modèle agricole implicite (modèle capitalistique ou reposant sur la main-d'œuvre, modèle de production-expédition ou transformation-commercialisation). Leur multiplicité permet en revanche de confronter les modèles d'agriculture, ou bien d'identifier des objectifs contradictoires, ou complémentaires, mais aussi d'interroger la place de l'activité agricole au sein de cette société. Dans l'évaluation des politiques de développement rural, cette confrontation des différents dénominateurs est régulièrement réalisée car il ne faut pas oublier que nos politiques publiques cherchent parfois à atteindre plusieurs objectifs qui ne sont pas toujours cohérents. Les données produites doivent donc permettre la construction d'un ensemble modulable d'indicateurs permettant d'éclairer cette complexité et non de la nier. Cet éclairage sous divers angles de la réalité est notamment prôné par le « réalisme critique » de Lawson et Lewis (1997) et la *realistic evaluation* de Pawson et Tilley (2007).

2.2. Pour une construction de données inscrite dans une démarche évaluative¹³

Ces considérations et la nécessité d'un regard critique sur la données chiffrée nous amènent à ré-interroger la statistique publique afin de proposer des améliorations touchant à la fois son contenu, sa construction et ses usages.

2.2.1. Des recours différents à la donnée, que l'évaluation peut contribuer à concilier

L'exigence de renouvellement des données statistiques s'accroît avec l'évolution des demandes et, plus fondamentalement, avec l'émergence d'un nouveau système de valeurs

12. Un inventaire des critiques faites au PIB ainsi que des propositions alternatives existantes a été fait par J. Gadrey et F. Jany-Catrice, 2007.

13. L'objectif n'est pas de soumettre la production de statistiques à l'influence du politique mais de quantifier des phénomènes socio-économiques à ce jour peu, ou pas, révélés par des données chiffrées. De fait, l'évaluation et la statistique partagent une déontologie commune d'indépendance, d'impartialité et de transparence. Des réflexions sont à mener sur la façon de concilier au mieux la production de différentes données en adoptant des postures claires dans le cadre de démarches évaluatives productrices de connaissances.

et une nouvelle hiérarchie des « éléments qui comptent » (dont font partie les attentes écologiques et sociétales), conformément à l'hypothèse soutenue par J. Gadrey et F. Jany-Catrice (2003 et 2007) dans le prolongement des travaux de D. Méda (1999) et du rapport de P. Viveret (2002).

Si toute connaissance tend naturellement à s'objectiver et à se préciser à partir de mesures quantifiées, l'appréhension de l'importance du chiffre et la manière de s'en saisir ne sont pas les mêmes selon que l'on se situe du côté des politiques, des chercheurs ou des citoyens. En effet, le chiffre constitue une assise rigoureuse pour les décideurs politiques, tant qu'il peut leur fournir des données empiriques « utiles » à la prise de décision (recensements pour déterminer les recettes fiscales ou le nombre d'exploitations et d'hectares en zone de montagne pour calibrer l'enveloppe allouée à l'ICHN, connaître les évolutions de la population agricole pour assurer l'équilibre du régime agricole des retraites, etc.). La donnée chiffrée peut permettre d'échapper, au moins partiellement, au flou et au subjectif des comparaisons dans le temps et l'espace. Mais la mobilisation d'un avis d'expert reste essentielle pour faire face aux tentations d'agrégation de données partielles ou de construction rapide d'indicateurs, simples ou composés, pour les besoins de décideurs pressés, ou de médias familiers de chiffres « signifiants » et peu nombreux (Jany-Catrice F. et Méda D., 2010). Pour le chercheur en revanche, le recours au chiffre vise avant tout à confirmer ou infirmer une thèse structurée à partir d'un débat théorique et non de considérations politiques.

Face à cette pluralité d'usages possibles du chiffre¹⁴, rappelons qu'aucun n'est scientifiquement supérieur aux autres. Il importe en effet d'encourager des « modalités » de co-production permettant de rapprocher davantage évaluation, production de données et analyse de l'intervention politique, conformément à notre triangle évaluatif (figure 2). Il s'agit d'améliorer les conditions de production et d'utilisation de données chiffrées afin que leur mobilisation dans l'évaluation des politiques ne se résume pas à une simple mesure de niveaux d'objectifs atteints mais contribue réellement à l'amélioration des politiques publiques par les connaissances qu'elles contribuent à produire.

Ainsi, la quantification des impacts environnementaux dans le domaine agricole pourrait permettre la mise en place d'un « pont » entre les ministères concernés en vue de développer des méthodes et des indicateurs communs, contribuant ainsi au développement d'une vision commune sur les pratiques agroenvironnementales.

Reste à poser la question du « volume » de données à produire. Si un éclairage multiple de la réalité, au travers de différents indicateurs construits pour véhiculer des informations spécifiques, est indispensable, inversement, un excès de données peut être contre-productif face à la concurrence des indicateurs synthétiques, plus « séduisants » et faciles. De plus, mettre à disposition des décideurs une batterie d'indicateurs revient à leur laisser la liberté de choisir l'indicateur qui leur paraît pertinent pour fonder leur jugement. Cela revient à faire l'hypothèse que ces acteurs sont capables de savoir ce qui compte et ce qu'il faut compter, ce qui n'est pas toujours vérifié. Un accompagnement soigné du choix ou de l'utilisation des bons indicateurs peut s'avérer nécessaire, dans la mesure où les démarches encadrant le recours à des données chiffrées co-construites permettent d'améliorer l'utilité socio-économique des démarches évaluatives (Offredi, 2010).

Au final, la construction de telles données ne doit pas être juste un moyen d'introduire du réalisme dans la démarche évaluative. Elles doivent produire une information ayant du

14. Il peut y avoir un recours aux chiffres à des fins de controverses (rapport Stern), d'interpellation (rapport sur la dette publique de 2005), de plaidoyer (projet mené au Conseil régional du Nord-Pas-de-Calais ayant débouché sur un indicateur de santé), d'appui à l'évaluation des besoins sociaux pour se positionner dans le temps et l'espace, etc.

sens, permettant d'alimenter une réflexion sur les mécanismes sur lesquels les interventions publiques s'appuient. À titre d'exemple, des suivis de cohorte détaillés sur des échantillons restreints pourraient permettre de mieux identifier les facteurs de réussite ou d'échec d'une mesure d'aide à l'investissement, que ce soit au niveau des exploitations des IAA ou des industries forestières.

2.2.2. En guise de recommandations opérationnelles

À terme, il est possible d'imaginer que les politiques deviennent un des objets d'étude de la statistique publique. Cela reviendrait à interroger de façon plus systématique le lien entre les évolutions observées dans la société et les politiques en place, l'objectif n'étant pas d'inféoder la statistique publique à la sphère politique, mais d'utiliser au mieux l'appareil statistique pour aider à l'évaluation, donc au pilotage de l'action publique.

Dans un premier temps, il convient de systématiser la mutualisation de la collecte de données en continuant le rapprochement déjà initié des données administratives de suivi et des données statistiques. Le pré-remplissage des questionnaires du recensement agricole de 2010 à l'aide des données administratives issues des déclarations PAC des exploitants agricoles et des demandes d'aides du deuxième pilier représente, outre la simplification et le gain de temps qui en découle lors des entretiens d'enquête, un premier pas dans ce sens. En effet, l'introduction dans les recensements et enquêtes statistiques de variables relatives au statut de bénéficiaire ou de non-bénéficiaire d'une exploitation vis-à-vis d'aides publiques est intéressante et sera notamment utile lors de la recherche d'éventuels effets d'aubaine.

L'impossibilité matérielle pour la statistique publique de produire par anticipation l'ensemble de la donnée nécessaire à l'évaluation d'une politique – montant des aides, ensemble de critères d'échantillonnage – implique que la production, par les évaluateurs, de données dédiées doit se poursuivre. Toutefois, la qualité de la donnée recueillie au moyen de ce circuit doit pouvoir faire l'objet d'une amélioration, notamment au niveau de la représentativité des échantillons de bénéficiaires et de non-bénéficiaires consultés lors d'une évaluation. Un appui technique des services de statistique publique au tirage des échantillons participerait à cette amélioration. Une autre solution aux biais de représentativité des réponses effectives aux enquêtes menées par des évaluateurs réside dans le redressement des résultats de ces enquêtes à partir de données consolidées sur l'échantillonnage statistique ayant servi à leur production.

L'indicateur met en évidence un phénomène, mais il le fait souvent de façon indirecte et détournée. Aussi un même phénomène peut avoir différentes traductions d'un pays à l'autre ou d'une région à l'autre. Par exemple, les navettes domicile-travail caractéristiques des espaces périurbains peuvent se traduire en temps ou en distance de parcours selon les sensibilités culturelles. Aussi il convient de recentrer le questionnement sur le sens de ce qu'on mesure et non sur la valeur prise par la mesure. Dans un premier temps, les producteurs de statistiques peuvent éclairer sur le sens des données qu'ils produisent. Par la suite, il convient de poursuivre collectivement l'effort d'identification des phénomènes importants à mesurer, en formulant les questions évaluatives qui correspondent et en proposant des alternatives dans leur mode de mesure en n'imposant pas un indicateur unique. Recentrer les indicateurs sur le sens de ce qu'ils mesurent, donc sur les phénomènes, permettra par exemple de mieux faire apparaître des grands types d'espaces ruraux comparables d'un État membre à un autre, et ce malgré des différences de niveau de développement parfois importantes. La comparabilité des politiques sera donc améliorée. À titre d'illustration, la définition des espaces ruraux actuellement retenue au niveau européen, pour le pilotage et la suivi de la politique de développement rural est très statique et normalisée. Elle pourrait être

améliorée si elle prenait en compte le lien fonctionnel entre ville et campagne, notamment au travers des dynamiques de navette domicile-travail et d'accès aux services.

Plus qu'un phénomène, c'est souvent un ensemble de phénomènes en interaction qu'il est nécessaire de mesurer au cours d'une démarche évaluative. La quantification introduit un fractionnement de systèmes complexes en systèmes élémentaires trop simples donc trop partiels et ne reflétant pas la réalité de la situation (Fouquet, 2010). Aussi, il convient de s'assurer que les systèmes d'indicateurs proposés permettent de bien conserver cette dimension multiple de la réalité et qu'ils ne la réduisent pas à une vision standardisée et aseptisée, donc erronée. Mesurer la valeur ajoutée d'un investissement ne se fait pas qu'en termes monétaires, mais également en termes humains et environnementaux. La triangulation¹⁵ n'est pas uniquement dans la confrontation de données de natures différentes, mais également dans l'éclairage qui est fait de la réalité par les statistiques.

La question de l'échelle de production de la donnée statistique est également essentielle. Le développement des approches locales plaide pour un besoin de représentativité fort à une échelle fine des données produites, et ce pour réaliser la nécessaire agrégation sur des périmètres variables selon les enjeux, du local au global. En effet, le caractère européen du deuxième pilier de la PAC, allié à ses dimensions partenariale et subsidiaire, rend nécessaire une mesure de ses effets à différentes échelles.

Par les connaissances qu'elle produit, l'évaluation des politiques publiques met au jour des phénomènes émergents dont une description plus approfondie repose nécessairement sur un recours à la statistique. Or, du fait de leur caractère novateur, ces phénomènes peuvent ne pas être suffisamment couverts par les enquêtes existantes. La possibilité de procéder à des surpondérations¹⁶ est une piste intéressante pour créer collectivement de la connaissance sur les systèmes agricoles et ruraux, et leurs réactions aux interventions publiques. Plus généralement et de manière plus systématique, le renforcement d'une démarche d'évaluation *in itinere*, avec un programme de travail conjoint entre évaluateurs et producteurs de statistiques permettrait de favoriser les synergies entre ces deux disciplines. La programmation pluriannuelle des politiques européennes de développement rural est un cas propice à une définition anticipée d'un tel programme commun.

Il apparaît également nécessaire d'investir plus largement dans la production ou la collecte de données statistiques relatives aux espaces ruraux pris dans leur définition la plus large et pas uniquement dans leur dimension agricole.

Enfin, il nous semble important de mener un travail de stabilisation des critères sur lesquels sélectionner des échantillons en vue de conduire des analyses comparatives des dynamiques rurales et des effets des politiques sur ces dynamiques. Cela reviendrait à se donner les moyens de répondre collectivement aux questions suivantes : quelles sont les dynamiques importantes à l'œuvre dans les espaces ruraux ? Quels sont les effets des actions politiques sur ces dynamiques ?

15. La triangulation est un des principes de base de l'évaluation qui consiste à recouper différentes sources d'information, quantitatives et qualitatives, afin de mieux étayer les jugements formulés.

16. Surreprésentation volontaire d'un sous-échantillon.

Conclusion

Le regard porté sur la mobilisation de la donnée chiffrée dans l'évaluation des politiques de développement rural met en évidence une diversité des sources de données et de leurs utilisations. Tantôt appelée pour décrire un contexte, tantôt pour mesurer un effet, la démarche de quantification est omniprésente dans le travail de l'évaluateur qui s'appuie largement sur des données préexistantes à l'évaluation. Toutefois, la mobilisation de données chiffrées est loin de se résumer à une simplification de la réalité au travers d'un ou deux indicateurs clés.

En effet, l'approche de l'évaluation centrée sur le questionnement évaluatif met l'accent sur le sens de l'action publique en interaction avec son contexte et appelle une pluralité de points de vue sur les questions à traiter. Cet éclairage multiple de la réalité trouve un fondement théorique dans les écoles du « réalisme critique » et de la *realistic evaluation*. Il se traduit, dans la pratique, par la mobilisation d'une batterie de données chiffrées. Chacune de ces données apporte une part d'information sur l'objet étudié dans la construction d'un jugement évaluatif plus global qui se révèle être, au final, la synthèse d'une observation combinant différents points de vue. Lorsque la donnée chiffrée mobilisée dans cette démarche préexiste à l'évaluation, il convient de rappeler les conditions et les conventions qui ont présidé à sa construction, afin d'éviter toute utilisation erronée. En revanche, compte tenu des évolutions rapides des modalités d'intervention de la puissance publique, notamment au travers des approches territorialisées, et de l'importance croissante de notions telles que le développement durable, certains consensus ayant servi à la construction des indicateurs aujourd'hui utilisés peuvent se trouver dépassés, voire obsolètes. Dans ce contexte, il apparaît donc nécessaire de les rénover, ce qui implique la mise au point de nouvelles modalités d'observation de l'interaction entre société, environnement et action publique au travers d'une plus grande prise en compte des spécificités de la démarche évaluative dans la production de données statistiques. Ainsi, la quantification des divers phénomènes à l'œuvre dans notre environnement, pris au sens large, pourra contribuer à la création de connaissances utiles à la compréhension des politiques publiques, donc à un meilleur pilotage de celle-ci.

Bibliographie

- Berriet-Sollic M., Le Roy A., Trouvé A., 2009, « Territorialiser la politique agricole pour plus de cohésion », *Revue d'Économie Rurale*, n° 313-314, pp. 129-146.
- Berriet-Sollic M., Chabé-Feret, S. (dir.), 2007, *Analyse et évaluation des impacts attendus de la politique de développement rural. Pour un pilotage plus efficace à l'échelle régionale*, INRA et régions Bourgogne et Auvergne.
- Bureau D., Mougéot M. Studer N., 2010, « Mesurer la performance de la gestion publique à la lumière de l'analyse économique », *Revue Française des affaires sociales*, n° 1-2, pp. 89-104

- Chabé-Ferret, S., Subervie, J., 2009, *Estimation des effets propres des mesures agro-environnementales du PDRN 2000-2006 sur les pratiques des agriculteurs*, Technical report, Cemagref UMR Metafort.
- Desrosières A., 2000, 2nd édition, *La politique des grands nombres, histoire de la raison statistiques*, La Découverte, Paris.
- Desrosières A., 2008, *L'argument statistique*, 2 tomes, Presses des Écoles des mines, Paris.
- Épices et Tercia, 2008, *Évaluation ex post du plan de développement rural national questions transversales et synthèse générale*.
- Fouquet A., 2010, « L'usage des statistiques : de l'aide à la décision à l'évaluation des politiques publiques », *Revue Française des affaires sociales*, n° 1-2, p. 309-322.
- Gadrey J., 2006, « Les conventions de richesse au cœur de la comptabilité nationale. Anciennes et nouvelles controverses », in Eymard-Duvernay F (Dir.) *L'économie des conventions, Méthodes et résultats*, Paris, La Découverte.
- Gadrey J. et Jany-Catrice F., 2007, *Les nouveaux indicateurs de richesse*, La Découverte, Paris.
- Gadrey J., 2008, « À la recherche de nouveaux indicateurs de bien-être et de santé sociale », in *Les statistiques publiques en débat, Regards sur l'actualité* n° 346, La Documentation française, Paris, pp. 53-62.
- Guibert B., 2008, « La production des statistiques publiques en France », in *Les statistiques publiques en débat, Regards sur l'actualité* n° 346, La Documentation française, Paris, pp.5-17.
- Jany-catrice F. et Méda D., 2010, « Les conditions sociales de la production du rapport Stiglitz sur les indicateurs de performances économiques et de progrès social », in Degraeve F. et al (dir.) *Les transformations et innovations économiques et sociale en Europe : quelle sorties de crises ?* Regards interdisciplinaires.
- Jany-Catrice F. et Zotti R., 2009, « La santé sociales des territoires. Un indicateur de santé sociale pour les régions françaises », *Futuribles*, n° 350, mars.
- Labrousse A., 2010, « Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective, d'outils de preuve et de gouvernement », *Revue de la régulation*, n° 7.
- Landel P. et Le Roy A., 2012, « Une analyse de la traduction régionale de la politique agricole rurale européenne : le cas de l'Émilie Romagne », *Revue d'Économie Régionale et Urbaine*, n° 1.
- Laurent C. et al., 2009, « Pourquoi s'intéresser à la notion d'evidence-based policy ? » *Revue Tiers-monde*, n° 200, p. 853-873.
- Lawson, T., 1997, *Economics and Reality*, London: Routledge.
- Leenhardt S., 2009, « Entre logique sectorielle et logique territoriale : quel positionnement pour les aides aux industries agroalimentaires ? », *Analyse N° 7*, MAAP, Centre d'études et de prospective, Paris.

- Le Roy A. et Offredi Cl., 2011, « La quantification au service de l'observation sociale local : à quelles conditions ? », *Revue Française de Socio-Économie*, premier semestre.
- Le Roy A., 2009, *Favoriser le rapprochement des pratiques du calcul économique et de l'évaluation pour une meilleure compréhension de l'action publique*, 9^e journée française de l'évaluation, SFE, Marseille 11-12 juin.
- Méda D., 2008, *Au-delà du PIB. Pour une autre mesure de la richesse*, Champs-Actuel, Paris.
- Millot G., 2010, *L'évaluation des politiques publiques, aide au pilotage de l'action ministérielle*, Centre d'étude et de prospective, *Analyse N° 19*, MAAP, Paris.
- Offredi C., 2004, « Penser le développement durable et l'évaluation » in *Dynamique de l'évaluation et du développement durable*, L'Harmattan, Paris.
- Offredi C. (dir.), 2010, *La notion d'Utilité sociale au défi de son identité dans l'évaluation des politiques publiques*, L'Harmattan, Paris.
- Perret B., 2008, « Évaluation et statistiques : une coopération difficile, » in *Les statistiques publiques en débat, Regards sur l'actualité n° 346*, La Documentation française, Paris, pp. 45-52.
- Perret B., 2011, *Pour une raison écologique*, Flammarion, Paris.
- Sanderson I., 2002, "Evaluation, policy learning and evidence-based policy making", *Public Administration*, vol. 80, n° 1, pp. 1-22.
- Tilley N., 2000, *Realistic Evaluation: An Overview*, Founding Conference of the Danish Evaluation Society
- Trosa S., 1992, « Le rôle de la méthode dans l'évaluation à travers l'expérience du Conseil Scientifique de l'Évaluation », *Politique et management public*, Vol X, n° 3, pp. 83-102.
- Viveret P., 2002, *Reconsidérer la richesse*, La tour d'Aigues, Ed. de L'Aube.

Recommandations aux auteurs

● Format

Les manuscrits sont présentés sous format Word ou Writer en police de taille 12. Ils ne dépassent pas 50 000 signes espaces inclus, y compris tableaux, graphiques, bibliographie et annexes.

Sur la première page du manuscrit doivent figurer :

- le titre de l'article ;
- le(s) nom(s) de(s) auteur(s) et leur(s) institution(s) ;
- le résumé de l'article (800 signes espaces compris) en français et en anglais ;
- trois à six mots-clés en français et en anglais.

Toutes les sources des chiffres cités doivent être précisées. Les sigles doivent être explicités. Lorsque l'article s'appuie sur une enquête, des traitements de données, etc., un encadré présentant la méthodologie est souhaité. Pour une meilleure lisibilité, les notes de bas de page doivent être limitées en nombre et en longueur.

Les références bibliographiques sont présentées ainsi :

- a** - Dans le texte ou les notes, chaque référence citée est constituée du nom de l'auteur et de l'année de publication entre parenthèses, renvoyant à la bibliographie en fin d'article. Par exemple : (Griffon, 2004).
- b** - À la fin de l'article, les références sont classées par ordre alphabétique d'auteurs et présentées selon les normes suivantes :
 - pour un ouvrage : nom de l'auteur, initiale du prénom, année, *Titre d'ouvrage*, ville, maison d'édition ;
 - pour un article : nom de l'auteur, initiale du prénom, année, « Titre d'article », *Revue*, n° de parution, mois, pages.

Seules les références explicitement citées ou mobilisées dans l'article sont reprises en fin d'article.

● Compléments pour mise en ligne de l'article

Dans la perspective de la publication de l'article sur le site internet du CEP et toujours selon leur convenance, les auteurs sont par ailleurs invités à :

- adresser le lien vers leur(es) page(s) personnelle(s) à caractère « institutionnelle(s) » s'ils en disposent et s'ils souhaitent la(les) communiquer ;
- communiquer une liste de références bibliographiques de leur choix utiles pour, contextualiser, compléter ou approfondir l'article proposé ;
- proposer une liste de lien vers des sites Internet pertinents pour se renseigner sur le sujet traité ;
- proposer, le cas échéant, des annexes complémentaires ou des développements utiles mais non essentiels (précisions méthodologiques, exemples, etc.) rédigés dans la phase de préparation de l'article mais qui n'ont pas vocation à intégrer la version livrée, limitée à 50 000 caractères. Ces compléments, s'ils sont publiables, viendront enrichir la version Internet de l'article.

● Procédure

Tout texte soumis est lu par au moins 3 membres du comité de rédaction. Deux fiches de lecture rédigées par un des membres du comité de rédaction et par un expert extérieur sont transmises aux auteurs. La décision de publication est prise collectivement par le comité de rédaction. Tout refus est argumenté.

Les manuscrits sont à envoyer, en version électronique uniquement, à :

- Pierre Claquin, secrétaire de rédaction : pierre.claquin@agriculture.gouv.fr
- Bruno Héroult, rédacteur en chef : bruno.herault@agriculture.gouv.fr

● Droits

En contrepartie de la publication, l'auteur cède à la revue *Notes et Études Socio-Économiques*, à titre exclusif, les droits de propriété pour le monde entier, en tous formats et sur tous supports, et notamment pour une diffusion, en l'état, adaptée ou traduite. À la condition qu'il demande l'accord préalable à la revue *Notes et Études Socio-Économiques*, l'auteur peut publier son article dans un livre dont il est l'auteur ou auquel il contribue à la condition de citer la source de première publication, c'est-à-dire la revue *Notes et Études Socio-Économiques*.

Notes et études socio-économiques

Tous les articles de *Notes et Études Socio-Économiques* sont téléchargeables gratuitement sur :

<http://www.agreste.agriculture.gouv.fr>

- Rubrique **Publications > Notes et études socio-économiques**

<http://agriculture.gouv.fr/centre-d-etudes-et-de-prospective>

- Rubrique **Publications du CEP > Notes et études socio-économiques**

Abonnement à l'alerte électronique en envoyant un message à l'adresse :

revue-nese@agriculture.gouv.fr avec le sujet « **abonnement** »

Notes et études socio-économiques

Ministère de l'Agriculture,

de l'Agroalimentaire et de la Forêt

Secrétariat Général

Service de la Statistique et de la Prospective

Centre d'études et de prospective

Renseignements :

Service de la Statistique et de la Prospective

Centre d'Études et de Prospective

12, rue Henri-Rol-Tanguy –

TSA 70007 –

93555 Montreuil sous Bois Cedex

tél. : 01.49.55.85.05

Diffusion :

Service de la Statistique et de la Prospective

Bureau des ventes – BP 32688

31326 – Castanet Tolosan cedex

Vente au numéro : agreste-ventes@agriculture.gouv.fr

fax : 05.61.28.93.66

Abonnement : tél. : 05.61.28.93.05