



Bruxelles, le **XXX**  
SWD(2014) **XXX**

**DOCUMENT DE TRAVAIL DES SERVICES DE LA COMMISSION**

**Méthodologie commune pour l'évaluation des aides d'État**

# **DOCUMENT DE TRAVAIL DES SERVICES DE LA COMMISSION**

## **Méthodologie commune pour l'évaluation des aides d'État**

### **Table des matières**

1	Introduction .....	2
2	Les objectifs de l'évaluation des aides d'État .....	3
3	Plan d'évaluation.....	5
3.1	Objectifs du régime d'aides à évaluer .....	5
3.2	Questions d'évaluation .....	5
3.3	Indicateurs de résultats .....	6
3.4	Méthodes: trouver une base de comparaison appropriée.....	7
3.5	Collecte des données: utilisation des meilleures sources disponibles .....	9
3.6	Calendrier de l'évaluation .....	11
3.7	Organe évaluateur: garantir indépendance et expertise .....	12
3.8	Publicité: faciliter la participation des parties prenantes .....	13
4	Critères de sélection des régimes d'aides d'État à évaluer.....	14
4.1	Régimes d'aides d'État de grande ampleur, y compris ceux relevant du règlement général d'exemption par catégorie .....	14
4.2	Régimes d'aides présentant des caractéristiques nouvelles.....	15
4.3	Régimes d'aides concernés par des changements importants prévisibles .....	15
4.4	Autres régimes d'aides .....	16
Annexe I:	Annexe technique concernant les méthodes d'identification de l'incidence causale .....	17
Annexe II:	Liste d'indicateurs de résultats envisageables .....	38
Annexe III:	Glossaire .....	43
Annexe IV:	Références .....	44

## 1 Introduction

Les aides d'État octroyées par les États membres visent à atteindre un large éventail d'objectifs fondamentaux, et notamment à réduire les disparités régionales au sein d'un même pays, à promouvoir les activités de recherche, de développement et d'innovation, et à favoriser un niveau élevé de protection environnementale.

Afin de déterminer les types d'aides compatibles avec le marché intérieur, les règles de l'UE applicables aux aides d'État s'appuient sur un système de contrôle ex ante: les régimes d'aides<sup>1</sup> sont autorisés sur la base de critères d'appréciation préalablement définis et en partant du principe que si lesdits régimes sont conformes à ces critères, leurs effets positifs l'emporteront sur leurs effets négatifs potentiels. En règle générale, l'appréciation des régimes d'aides se fait sans procéder à une évaluation suffisamment approfondie de leur incidence réelle, à terme, sur les marchés.

L'intérêt accordé jusqu'ici, lors de l'application des règles de l'UE en matière d'aides d'État, aux preuves ex post des résultats obtenus grâce aux fonds publics ou encore de l'incidence des aides d'État sur la concurrence n'a été que fort mesuré. Il n'en demeure pas moins capital que les décideurs, au niveau tant des États membres que de l'UE, prennent en considération les résultats concrets des aides d'État octroyées par le passé, de même que les leçons tirées des expériences antérieures. Non seulement pour veiller à ce que les régimes financés par des aides d'État gagnent en efficacité tout en créant moins de distorsions sur les marchés, mais aussi pour renforcer l'efficacité des régimes à venir, voire des règles futures relatives à l'octroi d'aides d'État.

Un certain nombre de pays évaluent déjà leurs mesures de subvention, même s'ils ne le font pas toujours à intervalles réguliers<sup>2</sup>. De même, les dépenses de l'UE (dont les financements au titre des Fonds structurels et d'investissement de l'UE tels que le FEDER, le FSE et le Feader) font l'objet, conformément aux règlements applicables et aux documents d'orientation publiés par la Commission<sup>3</sup>, d'évaluations ex ante, in itinere et ex post. Afin d'éviter tout doublon dans les évaluations menées par les États membres, le document d'orientation «Concepts et recommandations» concernant le suivi et l'évaluation précise que les exigences en matière

---

1 Les régimes d'aides représentent la majorité des aides octroyées: d'après les données du tableau de bord 2013, les régimes d'aides autorisés représentent 23 % de l'ensemble des mesures d'aides et 55 % du montant des aides. À ces régimes s'ajoute un autre éventail de régimes d'aides bénéficiant d'une exemption par catégorie, qui représente 63 % de l'ensemble des mesures d'aides et environ 32 % du montant des aides. Conformément au règlement n° 659/1999, on entend par «régime d'aides» «toute disposition sur la base de laquelle, sans qu'il soit besoin de mesures d'application supplémentaires, des aides peuvent être octroyées individuellement à des entreprises, définies d'une manière générale et abstraite dans ladite disposition et toute disposition sur la base de laquelle une aide non liée à un projet spécifique peut être octroyée à une ou plusieurs entreprises pour une période indéterminée et/ou pour un montant indéterminé».

2 Par exemple, dans plusieurs États membres, des rapports d'évaluation sur les aides d'État sont régulièrement établis à l'intention de la Cour des comptes ou du Parlement.

3 Le document d'orientation sur l'évaluation publié par la Commission pour la période de programmation 2014-2020 (disponible sur: [http://ec.europa.eu/regional\\_policy/information/evaluations/guidance\\_fr.cfm](http://ec.europa.eu/regional_policy/information/evaluations/guidance_fr.cfm)) présente dans le détail les concepts et recommandations applicables.

d'évaluation applicables pour les Fonds structurels et d'investissement de l'UE peuvent être satisfaites en pratiquant les évaluations imposées par les règles régissant les aides d'État.

L'initiative en faveur de la modernisation du contrôle des aides d'État<sup>4</sup> vise à concentrer les efforts de la Commission sur les régimes d'aides de grande ampleur, qui sont susceptibles d'avoir les incidences les plus importantes sur le marché intérieur. Parallèlement, l'analyse des affaires à caractère plus local et n'ayant qu'une incidence mineure ou plus limitée sur les échanges doit être simplifiée, notamment en laissant une plus grande marge de manœuvre aux États membres dans la mise à exécution de ces mesures d'aide et en élargissant le champ d'application du nouveau règlement général d'exemption par catégorie<sup>5</sup>. Afin, d'une part, de garantir que les effets positifs des aides d'État (tout en répondant à leur finalité première) continueront de l'emporter sur les effets négatifs susceptibles d'affecter la concurrence et les échanges et, d'autre part, de prévenir les distorsions de concurrence indues, cette simplification accrue devra s'accompagner d'une plus grande transparence, d'un meilleur contrôle du respect des règles régissant les aides d'État aux niveaux tant national qu'européen et d'une évaluation efficace<sup>6</sup>.

Le présent document définit une méthodologie commune pour l'évaluation des régimes d'aides d'État et vise à fournir des orientations aux pouvoirs publics concernés par la planification et la conduite de ces évaluations.

## **2 Les objectifs de l'évaluation des aides d'État**

L'objectif principal de l'évaluation des aides d'État réside dans l'appréciation des effets tant positifs que négatifs des régimes, à savoir la mise en balance de l'objectif d'intérêt public de ces aides avec leur incidence sur la concurrence et les échanges entre États membres. L'évaluation des aides d'État indique si et dans quelle mesure les objectifs initiaux du régime d'aides ont été atteints (appréciation des effets positifs) et permet de déterminer l'incidence du régime d'aides sur les marchés et la concurrence (appréciation des effets négatifs potentiels). Par conséquent, l'évaluation diffère, de par sa finalité, des deux exercices ex post actuellement réalisés par la Commission en ce qui concerne les régimes d'aides d'État, à savoir le contrôle<sup>7</sup> et l'établissement de rapports<sup>8</sup>.

---

4 Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions, Modernisation de la politique de l'Union européenne en matière d'aides d'État, COM(2012) 209 final du 8.5.2012.

5 Règlement (UE) n° 651/2014 de la Commission du 17 juin 2014 déclarant certaines catégories d'aides compatibles avec le marché intérieur en application des articles 107 et 108 du traité.

6 Voir également les conclusions du Conseil sur la réforme du contrôle des aides d'État du 13 novembre 2012.

7 Le contrôle exercé par la Commission consiste en un examen périodique de la légalité d'un échantillon de mesures d'aide d'État mises en œuvre par les États membres. Il vise à s'assurer que les États membres appliquent correctement les décisions prises par la Commission et se conforment aux dispositions juridiques applicables (à savoir celles énoncées dans le règlement général d'exemption par catégorie). La Commission évalue également, parmi un échantillon représentatif de cas, la conformité des aides avec les conditions et règles ex ante.

8 Les rapports annuels établis par les États membres ont pour objectif premier d'accroître la transparence des aides d'État octroyées par les États membres et de fournir, parallèlement, une source statistique fiable tant à des fins de contrôle que de formulation des politiques. Les données figurant dans les rapports annuels

L'évaluation des aides d'État doit, en particulier, permettre de juger de l'effet incitatif direct de l'aide sur son bénéficiaire (voir si l'aide a amené son bénéficiaire à préférer une autre voie et de quelle ampleur a été l'incidence de cette aide). Cette évaluation doit également rendre compte des effets globalement positifs et négatifs du régime d'aides non seulement sur l'accomplissement de l'objectif fondamental affiché mais aussi sur la concurrence et les échanges. Enfin, elle a pour vocation de juger de la proportionnalité et du caractère approprié de l'instrument d'aide choisi.

L'évaluation pourra, grâce à ces éléments d'appréciation, confirmer si les hypothèses qui sous-tendent l'autorisation du régime d'aides sont toujours d'actualité mais aussi faciliter l'élaboration des futurs régimes d'aides et règles régissant les aides d'État. Elle pourrait aussi être utilisée pour adapter les futures interventions des États, notamment en vue d'accroître l'efficacité et l'efficience des aides octroyées, de sorte que leurs effets positifs suffisent à justifier l'acceptation des distorsions de concurrence engendrées par l'intervention publique. Les améliorations susceptibles d'être apportées aux futurs régimes d'aides vont des adaptations apportées à leur conception, incluant la modification des critères de sélection et une évaluation plus poussée de l'effet incitatif, à des changements plus radicaux tels que la valorisation d'autres formes d'aide, la redéfinition des objectifs ou des bénéficiaires cibles ou encore le recours à d'autres options que les aides pour atteindre les objectifs fondamentaux fixés.

Il est essentiel de définir un calendrier d'évaluation adapté, c'est-à-dire offrant suffisamment de temps pour recueillir suffisamment d'éléments, tout en fournissant au plus tôt des résultats aux décideurs politiques, de façon à ce que les changements éventuels puissent être introduits en temps utile<sup>9</sup>. Aussi les évaluations des aides d'État devront-elles être considérées comme des évaluations *in itinere*, à mener pendant que le régime d'aides est opérationnel, plutôt que comme des évaluations *ex post*, menées une fois le régime expiré. Il conviendra également de prendre en compte les cas particuliers dans lesquels les effets des interventions ne sont pleinement perceptibles qu'à long terme et ceux dans lesquels les évaluations ne peuvent saisir et mesurer que leurs effets initiaux.

Enfin, l'évaluation des aides d'État devrait constituer un exercice d'apprentissage, tant pour la Commission que pour les États membres. Pour ce faire, elle devrait répondre à un certain nombre de normes de qualité minimales. La Commission devrait par conséquent s'assurer que des contrôles de qualité adéquats ont bien lieu, examiner en particulier le degré de fiabilité général de l'évaluation et détecter les éventuelles défaillances aux deux stades primordiaux que sont le plan d'évaluation et le rapport final. Si nécessaire, elle peut solliciter le concours d'experts externes indépendants lors du contrôle qualité de l'évaluation.

---

donnent avant tout des informations d'ordre quantitatif (par exemple sur la finalité des aides d'État et sur l'ampleur de l'enveloppe budgétaire). Quant à la Commission, elle se fonde sur les rapports établis par les États membres pour élaborer le tableau de bord des aides d'État.

9 Certaines lignes directrices concernent les aides d'État mentionnent une durée normale de quatre ans pour les régimes d'aides soumis à une évaluation.

La Commission pourrait également organiser, à l'intention des administrations nationales, des ateliers et des sessions de formation consacrés aux méthodes et techniques d'évaluation. De même, les expériences fructueuses et les bonnes pratiques des États membres pourraient être partagées et utilisées afin d'améliorer l'efficacité des régimes d'aides.

Les avantages de ces évaluations apparaîtront d'ici quelques années, lorsque les premiers rapports d'évaluation seront disponibles et leurs conclusions et recommandations diffusées. Ces rapports pourront alors être utilisés pour améliorer la conception des régimes d'aides ultérieurs et, si possible, les règles régissant les aides d'État. À moyen comme à long terme, les évaluations auront vocation à induire, progressivement, un changement d'approche plus fondamental à l'égard des aides d'État.

### **3 Plan d'évaluation**

Il est capital qu'un plan d'évaluation détaillé du régime d'aides d'État soit élaboré dès le départ, parallèlement à la conception dudit régime. L'approbation de ce plan d'évaluation par la Commission est un préalable pour garantir une égalité de traitement et sa mise en œuvre doit obéir à la plus grande rigueur.

Il est en effet communément admis que les évaluations sont plus efficaces dès lors qu'elles sont planifiées et préparées, car la collecte de données pertinentes s'en trouve facilitée. Une planification effectuée à un stade précoce a également toutes les chances de réduire sensiblement les ressources nécessaires à l'évaluation et, in fine, d'en améliorer la qualité.

En tout état de cause, le plan d'évaluation notifié par les États membres à la Commission, conformément aux règles applicables, devrait comporter au moins les éléments mentionnés ci-après.

#### **3.1 Objectifs du régime d'aides à évaluer**

La première étape de l'évaluation d'un régime d'aides consiste à définir clairement la «logique d'intervention» qui sous-tend le régime d'aides, autrement dit à décrire les besoins et les problèmes auxquels le régime entend remédier, les bénéficiaires cibles et les investissements visés, les objectifs généraux et spécifiques du régime ainsi que ses effets escomptés, sans oublier les principales hypothèses émises sur les facteurs externes susceptibles d'influer sur le régime.

#### **3.2 Questions d'évaluation**

Le plan d'évaluation doit définir la portée de l'évaluation, c'est-à-dire comporter des questions précises auxquelles il est possible d'apporter des réponses quantitatives, corroborées par les éléments de preuve nécessaires. Les questions d'évaluation doivent porter principalement sur l'incidence du régime d'aides d'État et être classées en fonction des trois niveaux suivants:

1. Incidence directe de l'aide sur les bénéficiaires:

- L'aide a-t-elle eu une incidence significative sur la ligne de conduite adoptée par les bénéficiaires de l'aide (effet incitatif)?
- L'aide a-t-elle eu un effet sur la situation des bénéficiaires (par exemple, a-t-elle amélioré ou altéré leur position concurrentielle, diminué ou aggravé leur risque de défaillance)?
- Dans quelle mesure l'aide a-t-elle eu les effets escomptés?
- L'incidence de l'aide a-t-elle varié en fonction des bénéficiaires (par exemple en fonction de leur taille, de leur localisation, de leur secteur d'activité)?

2. Incidence indirecte du régime d'aides:

- Le régime a-t-il eu des retombées sur l'activité d'autres entreprises ou sur d'autres régions géographiques? L'aide a-t-elle évincé les investissements consentis par d'autres concurrents ou détourné l'activité économique de sites voisins?
- Le régime d'aides a-t-il contribué à la réalisation de l'objectif fondamental concerné?
- Est-il possible de mesurer les effets cumulés du régime sur la concurrence et sur les échanges?

3. Proportionnalité et caractère approprié du régime d'aides:

- Le régime d'aides était-il proportionné au problème traité? Aurait-il été possible d'obtenir les mêmes effets au moyen d'une aide moins importante ou d'une autre forme d'aide (par exemple un prêt plutôt qu'une subvention)?
- L'instrument d'aide choisi était-il le plus efficace? Un autre instrument d'aide ou un autre type d'intervention aurait-il été plus approprié pour atteindre l'objectif recherché?

L'évaluation devrait évaluer, autant que faire se peut, l'incidence du régime d'aides aux trois niveaux, en abordant les questions se rapportant spécifiquement à ses objectifs. L'incidence directe du régime d'aides sur les bénéficiaires est généralement le type d'incidence qu'il est possible de mesurer avec le plus de fiabilité. Concrètement, la plupart des méthodes d'évaluation déjà élaborées ont été conçues de manière à apprécier ce type d'incidence. En outre, l'évaluation des effets directs des aides, parmi lesquels l'effet incitatif, revêt une importance capitale dans la mesure où elle apporte un éclairage précieux sur les différents types d'effets indirects et de distorsions prévisibles. En particulier, lorsque l'aide n'est assortie d'aucun effet incitatif, on est en droit de supposer qu'elle fausse la concurrence, dans la mesure où elle apporte aux bénéficiaires concernés des gains exceptionnels.

### **3.3 Indicateurs de résultats**

Les questions d'évaluation doivent conduire à choisir des indicateurs de résultats permettant de recueillir des informations chiffrées sur les résultats produits par le régime d'aides d'État. L'annexe II propose une liste indicative mais non exhaustive d'indicateurs de résultats couvrant l'incidence tant directe qu'indirecte des régimes d'aides, et notamment leurs effets potentiels sur la concurrence et les échanges. Le choix des indicateurs de résultats dépendra de l'objectif de l'aide à évaluer. Le plan d'évaluation doit indiquer en quoi les indicateurs retenus sont les plus adaptés pour mesurer l'incidence du régime d'aides concerné.

### **3.4 Méthodes: trouver une base de comparaison appropriée**

Les évaluations des aides d'État doivent être en mesure de déterminer l'incidence causale du régime en tant que tel, sans qu'elle soit faussée par d'autres variables susceptibles d'avoir eu un effet sur le résultat observé, comme par exemple le contexte macroéconomique général ou l'hétérogénéité des entreprises (différences en termes de taille, de lieu d'implantation, de situation financière ou de capacités de gestion). Le plan d'évaluation doit répertorier les principales méthodes à utiliser pour quantifier l'effet de l'aide et expliquer pourquoi ces méthodes sont les plus indiquées au regard du régime concerné.

L'incidence causale correspond à la différence entre le résultat obtenu grâce à l'aide et celui obtenu en l'absence d'aide. Alors que le résultat avec l'aide est observé auprès des entreprises bénéficiaires, le résultat obtenu en l'absence d'aide n'est mesuré qu'auprès d'entreprises non bénéficiaires. Par définition, nous n'observons pas ce qu'aurait été le résultat en l'absence d'aide pour les entreprises bénéficiaires de l'aide. Il est donc nécessaire, pour apprécier l'effet de l'aide sur les bénéficiaires, de construire un scénario contrefactuel (situation en l'absence d'intervention) s'appuyant sur les entreprises ou le groupe de contrôle se prêtant le mieux à la comparaison.

C'est de la qualité de ce groupe de contrôle que dépendra la validité de l'évaluation.

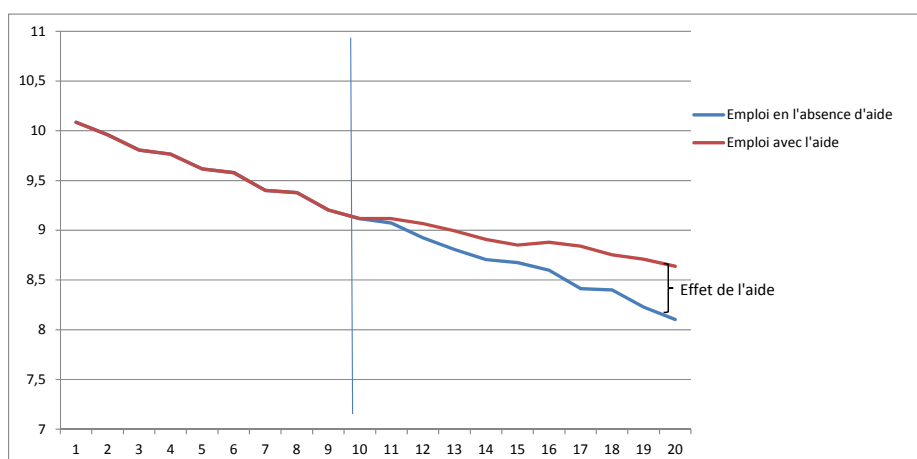
Les entreprises qui reçoivent une aide sont susceptibles de se trouver dans une situation très différente de celle des entreprises qui n'en bénéficient pas. Elles peuvent par exemple être confrontées à des conditions différentes sur le plan local en ce qui concerne l'offre et la demande, connaître des difficultés d'accès au crédit ou être plus ou moins performantes. Autant de facteurs qui peuvent avoir une incidence sur les performances ou le degré d'activité des entreprises, qu'elles reçoivent ou non une aide. Dès lors, une évaluation qui se bornerait à comparer des entreprises bénéficiaires à des entreprises non bénéficiaires risquerait de rendre compte d'abord et avant tout de cette réalité et non de l'effet proprement dit de l'aide. L'évaluation du régime d'aides ne peut donc pas se cantonner à une simple comparaison entre bénéficiaires et non-bénéficiaires, elle doit prendre en considération les caractéristiques – tant observables que non observables – des deux groupes d'entreprises.

Dans le cas des aides à finalité régionale, par exemple, les bénéficiaires situés dans des régions où les conditions du marché sont défavorables (c'est-à-dire où les marchés locaux de la production, du travail et des capitaux sont peu porteurs) enregistrent généralement de piètres performances par rapport aux non-bénéficiaires implantés dans des régions plus



prospères, ce qui ne reflète en aucun cas l'effet de l'aide en soi. Dès lors, l'objectif consiste à savoir si les entreprises bénéficiaires ont enregistré de meilleures performances qu'elles ne l'auraient fait en l'absence d'aide et non à déterminer si elles ont enregistré de meilleures performances que les non-bénéficiaires situés dans d'autres régions.

De même, les grandes tendances sectorielles doivent elles aussi être prises en compte pour quantifier l'effet des aides. En effet, le fait, par exemple, que les bénéficiaires d'aides à finalité régionale compriment leurs effectifs n'implique pas nécessairement que les aides n'ont pas été efficaces. À titre d'exemple, lorsque la conjoncture se dégrade au sein d'un secteur d'activité et que toutes les entreprises suppriment des emplois, les bénéficiaires des aides peuvent être amenés à dégraisser moins massivement qu'ils ne l'auraient fait en l'absence d'aide. Cet exemple est illustré dans le graphique ci-après, qui reproduit la courbe – tendancielle négative – du nombre d'emplois fournis par les entreprises bénéficiant d'aides, à la fois avant et après l'octroi de ces dernières. Il en ressort que cette courbe est moins négative une fois que l'entreprise a perçu une aide. L'écart entre la ligne tendancielle correspondant à l'absence d'aide et la ligne indicative du nombre d'emplois fournis par les entreprises après l'octroi des aides permet d'isoler l'incidence positive de ces dernières.



**Graphique 1: Incidence positive des aides en cas de tendance négative**

Identifier le groupe de contrôle devient difficile lorsque les non-bénéficiaires décident eux-mêmes de solliciter ou de ne pas solliciter d'aides. Si toutes les entreprises sont admissibles au bénéfice des aides (autrement dit toutes les entreprises qui présentent un projet et introduisent une demande d'aide la reçoivent de facto), les entreprises qui n'introduisent aucune demande d'aide seront a priori celles qui n'ont pas de projet. Les résultats enregistrés par les entreprises risquent d'indiquer que les entreprises non bénéficiaires ont affiché, en termes absolus comme relatifs, de piètres performances par rapport à celles ayant bénéficié des aides. Cependant, cette constatation peut s'expliquer par le seul fait que le premier groupe n'avait aucun projet à mettre en œuvre tandis que le second en avait, preuve que les dirigeants du premier groupe n'y voyaient aucun intérêt ou manquaient de créativité. Il est par conséquent essentiel que les entreprises du groupe de contrôle (entreprises n'ayant pas bénéficié des aides) fassent partie de ce groupe pour des motifs dépourvus d'effets sur les résultats mesurés. Or, cette condition risque de ne pas être remplie lorsque les entreprises se

sont auto-sélectionnées et ont décidé elles-mêmes, volontairement, de ne pas solliciter d'aides.

Toute différence systématique entre les bénéficiaires et les non-bénéficiaires d'aides d'État doit être convenablement prise en compte dans la conception de l'évaluation, de manière à ne pas fausser les résultats (biais de sélection). Plusieurs méthodes fiables ont été élaborées au cours des dernières décennies pour remédier à ce problème. Le choix de la méthode dépend du régime d'aides d'État concerné ainsi que des données disponibles. Cependant, chaque méthode a ses propres limites et n'acquiert sa validité que moyennant certaines hypothèses. Identifier et analyser ouvertement ces limites et ces hypothèses est donc indispensable pour garantir la crédibilité de l'étude.

Rendre le processus de sélection des bénéficiaires aléatoire constitue un moyen, parmi d'autres, de garantir que l'évaluation ne sera pas biaisée. Si les bénéficiaires d'aides sont entièrement sélectionnés de manière aléatoire, toute différence systématique constatée entre les performances des différentes entreprises pourra être attribuée aux aides octroyées. Cette méthode risque toutefois d'être difficile à appliquer, surtout dans le cas de régimes de grande ampleur existants. D'autres méthodes visent à utiliser les sources de variations exogènes dans l'environnement dans lequel les entreprises évoluent (autrement dit des variations qui ne sont déterminées ni par des paramètres ni par des variables dans le modèle), de façon à déterminer le lien de causalité<sup>10</sup>. L'annexe I du présent document d'orientation expose plus en détail les méthodes les plus pertinentes, en se penchant tout particulièrement sur les aspects pratiques de leur utilisation. Elle examine la façon dont chaque méthode détermine le lien de causalité, ce qui revêt une importance particulière dans le cadre des évaluations d'aides d'État pour lesquelles la conception ex ante vise à rendre possible une évaluation en bonne et due forme des effets des aides.

Pour conclure, l'incidence des aides multiples – qu'elles soient octroyées au titre d'un seul ou de plusieurs régimes ou qu'il s'agisse d'aides ad hoc – doit lui aussi être contrôlé. En effet, si les non-bénéficiaires d'un programme reçoivent des aides au titre d'autres programmes, ou si les bénéficiaires d'un programme reçoivent des aides supplémentaires au titre d'autres programmes, l'évaluation des effets du régime d'aides observé risque de s'en trouver faussée.

### **3.5 Collecte des données: utilisation des meilleures sources disponibles**

Il est indispensable de recueillir, tant sur les bénéficiaires des aides que sur le groupe de contrôle, des données cohérentes et en quantité suffisante. Déterminer le type de données nécessaires et obtenir l'accès aux sources de données fait partie intégrante du travail de planification de l'évaluation.

Pour garantir la qualité de l'évaluation, il est également indispensable de contrôler efficacement l'intervention ainsi que de recueillir et de traiter les données avec soin. Pour ce

---

<sup>10</sup> Les méthodes les plus couramment utilisées sont la méthode des doubles différences, la régression par discontinuité (RD) et les variables instrumentales.

faire, dès que le régime d'aides est autorisé, il convient de mettre en place un mécanisme permettant d'effectuer le suivi de l'intervention, ainsi que de collecter et de traiter les données requises. Les coûts de l'évaluation s'en trouveront considérablement réduits.

S'assurer de la bonne collecte des données requises tant sur les demandeurs que sur les bénéficiaires des aides constitue une étape essentielle dans l'élaboration du plan d'évaluation, en particulier si la disponibilité de ces données est susceptible de devenir l'un des critères d'admissibilité au bénéfice des aides.

Exception faite des données relatives aux demandes d'aides (y compris celles émanant des candidats éliminés lorsqu'elles sont disponibles), les sources d'information relatives aux bénéficiaires d'aides et au groupe de contrôle doivent être identiques, pour des raisons de comparabilité des données. Il est fort probable que ces données devront être recueillies auprès de sources multiples: bases de données contenant des informations sur les aides octroyées, registres des entreprises ou autres. L'évaluation devra éventuellement puiser dans des sources de données existantes, notamment les sources de données administratives (centres des impôts, registre du commerce, enquêtes sur l'innovation et office des brevets). Le plan d'évaluation devra donc passer en revue les sources de données existantes, décider si elles fournissent suffisamment d'informations pour procéder à l'évaluation et garantir l'accès à ces sources dans les délais requis.

Les données provenant de sources administratives telles que les instituts nationaux de statistique sont susceptibles d'être mises à la disposition des évaluateurs uniquement à certaines conditions, concernant notamment le respect de la vie privée et la confidentialité des données commerciales. Les conditions d'accès à ces données doivent être décrites dans le plan d'évaluation. L'autorité accordant l'accès aux données devra, si nécessaire, veiller à ce que les experts en charge de l'évaluation y aient bien accès.

Lorsque les données utilisées proviennent de sources différentes, il est essentiel de les recueillir dans un format qui permette d'apparier les variables correctement. Il pourra donc s'avérer nécessaire de trouver, dans chaque jeu de données utilisé, des identifiants uniques pour les différentes unités d'observation. Par exemple, les identifiants entreprise et usine devront être uniques dans tous les jeux de données; les adresses devront être recueillies dans un format idoine pour la géolocalisation, etc. L'origine exacte de l'identifiant peut différer d'un État membre à l'autre. L'identifiant peut par exemple être d'origine fiscale (comme le numéro de TVA) ou être fourni par les instituts de statistique [comme les numéros SIREN (identifiant d'entreprise) et SIRET (identifiant d'établissement) en France, tous deux fournis par l'Institut national de la statistique et des études économiques (INSEE)].

L'évaluation des aides d'État peut être complétée par des informations issues d'enquêtes sur les bénéficiaires et/ou d'entretiens avec les gestionnaires des régimes d'aides d'État. Sachant que les informations qualitatives ainsi recueillies sont, par nature, subjectives et que les réponses risquent de traduire davantage les intérêts stratégiques des bénéficiaires qu'une appréciation objective de l'effet des aides. Ce risque est particulièrement élevé lorsque l'interviewé part du principe qu'un témoignage favorable augmentera ses chances de recevoir

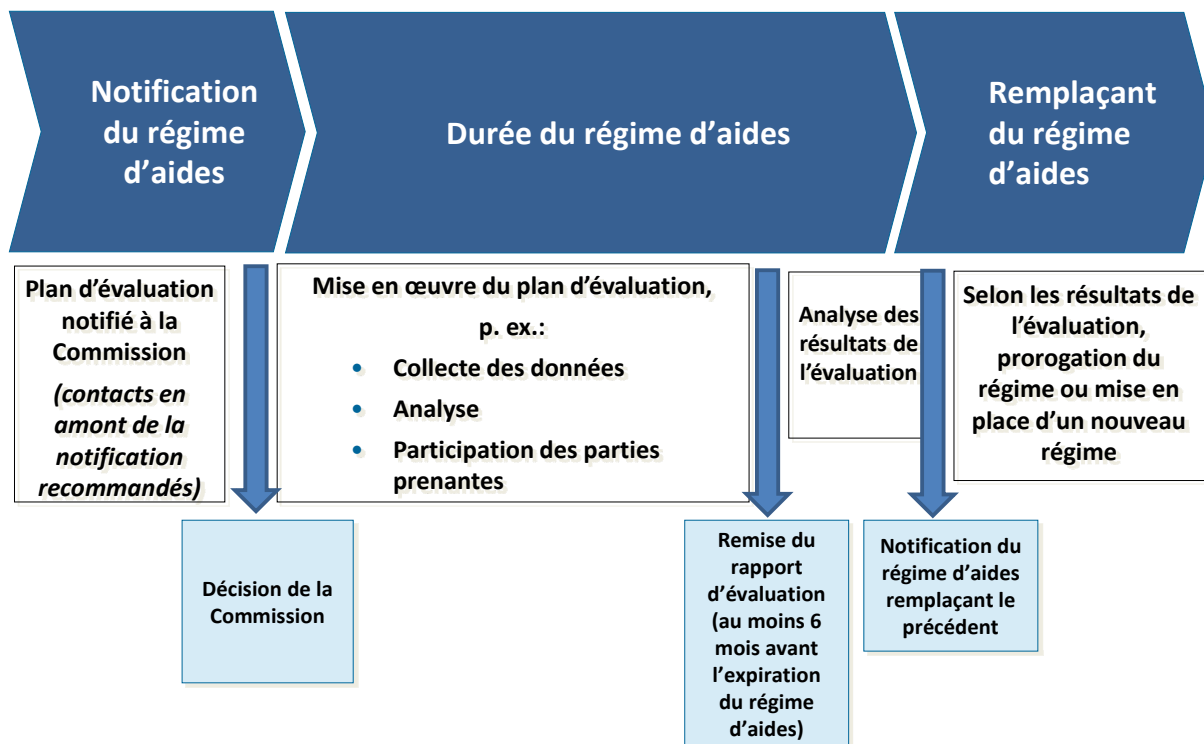
des aides à l'avenir. Il n'en demeure pas moins qu'à condition d'être traitées avec toute la prudence requise, les informations issues des exercices qualitatifs que sont les entretiens et les études de cas peuvent constituer une source d'informations complémentaires intéressante et faciliter l'interprétation des résultats de l'évaluation elle-même.

Chaque fois que des données personnelles sont traitées dans le cadre d'évaluations, la législation de l'UE relative à la protection des données à caractère personnel s'applique, notamment la directive 95/46/CE relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, les mesures d'application nationales s'y rapportant, et le règlement (CE) n° 45/2001 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les institutions et organes communautaires et à la libre circulation de ces données.

### **3.6 Calendrier de l'évaluation**

Le plan d'évaluation doit fournir des informations sur le calendrier de l'évaluation (lequel doit être défini avec précision et en fonction de la durée d'autorisation du régime d'aides) et comporter des échéances, par exemple pour la collecte des données, la réalisation de l'évaluation et la présentation du rapport final. Le calendrier étant susceptible de varier en fonction du régime, il fera l'objet d'une concertation et d'un accord au cas par cas avec la Commission. Il est conseillé aux personnes responsables de la gestion des régimes d'aides de favoriser les échanges de vue informels sur le contenu du plan d'évaluation avant de le notifier officiellement à la Commission.

En cas de projet de prorogation d'un régime d'aides, il convient que le rapport final d'évaluation soit remis à la Commission suffisamment tôt (par exemple six mois avant la fin programmée du régime d'aides) pour permettre l'examen du projet de prorogation. Si aucune prorogation n'est envisagée, le rapport pourra être remis une fois le régime arrivé à expiration.



Graphique 2: Vue d'ensemble du processus d'évaluation dans le cas d'un régime d'aides notifié

### 3.7 Organe évaluateur: garantir indépendance et expertise

L'évaluation des effets des régimes d'aides d'État doit être objective, rigoureuse, impartiale et transparente<sup>11</sup>. Chaque évaluation doit être réalisée selon des méthodes éprouvées, par des experts disposant d'une expérience adéquate avérée ainsi que des connaissances méthodologiques nécessaires pour effectuer ce type de tâches.

Les évaluations seront effectuées par un organisme indépendant, au moins sur le plan fonctionnel, de l'autorité octroyant les aides et attestant des compétences nécessaires ainsi que d'un personnel dûment qualifié pour mener à bien ce type d'évaluations. Cette indépendance est essentielle pour garantir la qualité et la crédibilité de l'évaluation. Elle ne requiert toutefois pas forcément la création d'un nouvel organe ni l'externalisation de l'évaluation pour la confier à des évaluateurs privés. Il est donc envisageable, en fonction des instances en place dans chaque État membre, de solliciter les compétences d'organismes indépendants tels que les instituts nationaux de statistique, les banques centrales, les cours des comptes ainsi que les universités et centres de recherche publics comme privés. Le choix pourra être fait, pour chaque régime d'aides, au cas par cas.

<sup>11</sup> Se reporter par exemple aux standards d'évaluation de la Commission européenne, aux normes et standards d'évaluation de l'OCDE, aux règles d'évaluation applicables au sein du système des Nations unies et au document de la Banque mondiale intitulé «Independent Evaluation: Principles, Guidelines and Good Practice».

L'intervention précoce de l'organe évaluateur, par exemple dès la conception du régime d'aides, est importante pour la réussite de l'évaluation, dans la mesure où le régime d'aides d'État pourra être évalué selon les modalités proposées et où il sera possible de recueillir les données indispensables à l'évaluation. Il convient donc que le plan d'évaluation soit, si possible, élaboré par l'évaluateur désigné ou en étroite collaboration avec lui. Ce plan devra également comporter, même à titre purement indicatif, des informations sur les ressources humaines et financières nécessaires au bon déroulement de l'évaluation. Des informations sur l'identité et le rôle de chacun des principaux experts participant à l'évaluation de même qu'une estimation de leur degré de participation seront également nécessaires.

Le plan d'évaluation doit par ailleurs décrire précisément l'organe évaluateur ou, s'il n'est pas encore choisi, les critères qui présideront à son choix, notamment en termes d'indépendance, d'expérience et de compétences. Dans la mesure du possible, il devra également proposer des seconds choix. Si l'évaluateur n'a pas encore été choisi ou s'il l'a été mais n'a pas été activement associé à l'élaboration du plan d'évaluation, les raisons devront en être clairement précisées. Car même dans ce cas de figure, le plan d'évaluation doit être suffisamment détaillé pour permettre de juger de la validité de l'évaluation à réaliser.

### **3.8 Publicité: faciliter la participation des parties prenantes**

L'évaluation doit être rendue publique. Cela suppose que tant le plan d'évaluation que le rapport d'évaluation, une fois ce dernier approuvé, fassent l'objet d'une publicité suffisante et soient accessibles dans tous les lieux mentionnés dans le plan d'évaluation, y compris sur l'internet. La Commission peut elle aussi rendre ces documents publics<sup>12</sup>.

Si les données utilisées aux fins de l'évaluation sont à caractère personnel et/ou de nature confidentielle, leur confidentialité devra être garantie tout au long du processus d'évaluation, conformément aux dispositions des articles 8, 16 et 17 de la Charte des droits fondamentaux de l'Union. Cette confidentialité ne s'étend toutefois pas aux résultats de l'évaluation. De sorte que le contrat d'évaluation ne peut comporter aucune clause de confidentialité, à l'exception: 1°) des obligations de non-divulgence des données à caractère personnel et/ou des données confidentielles; 2°) des obligations de respect des dispositions générales régissant le droit statistique national et le secret statistique, telles que celles liées à la présentation des résultats.

Les données collectées au cours de l'évaluation doivent être rendues accessibles, non seulement à des fins de réplique des résultats mais aussi en vue d'autres études, et soumises aux mêmes conditions de restriction que celles imposées à l'organe chargé de l'évaluation initiale.

---

12 Exception faite des secrets d'affaires et autres informations confidentielles dans les cas dûment justifiés. Voir la communication de la Commission sur le secret professionnel dans les décisions en matière d'aides d'État, C(2003) 4582 (JO C 297 du 9.12.2003, p. 6). La publication de données à caractère personnel doit se faire dans le respect de la législation de l'UE relative à la protection des données à caractère personnel, et notamment de la directive 95/46/CE et de ses mesures nationales d'exécution ainsi que du règlement (CE) n° 45/2001.

L'autorité d'octroi des aides d'État pourrait veiller à ce que toutes les parties intéressées soient associées et consultées au moins une fois au cours de la mise en œuvre du plan d'évaluation. Les parties pourront par exemple être invitées à analyser les résultats de l'évaluation initiale à la lumière du rapport intermédiaire. Ces dispositions pourront, le cas échéant, figurer dans le plan d'évaluation.

#### **4 Critères de sélection des régimes d'aides d'État à évaluer**

En principe, tout régime d'aides d'État peut faire l'objet d'une évaluation. Cependant, si l'évaluation est considérée comme une bonne pratique, elle n'est toutefois pas exigée par les règles relatives aux aides d'État dans tous les cas. L'évaluation des aides d'État doit rester proportionnée et concerner les régimes susceptibles d'avoir une incidence significative sur le marché intérieur et de créer d'importantes distorsions de concurrence si leur mise en œuvre n'est pas examinée en temps utile. Dans les lignes directrices applicables concernant les aides d'État, l'accent est mis sur les régimes d'aides d'État: (1) de grande ampleur, y compris ceux relevant du règlement général d'exemption par catégorie; (2) présentant des caractéristiques nouvelles; ou (3) susceptibles de connaître dans un proche avenir des changements importants (en ce qui concerne le marché, la technologie ou la réglementation), de nature à nécessiter une nouvelle appréciation du régime. Les différentes lignes directrices mentionnent également d'autres types de régimes pour lesquels une évaluation pourrait se révéler utile.

##### **4.1 Régimes d'aides d'État de grande ampleur, y compris ceux relevant du règlement général d'exemption par catégorie**

Conformément à la communication sur la modernisation de la politique en matière d'aides d'État, la Commission pourrait demander à ce que les régimes d'aides d'État de très grande ampleur fassent l'objet d'une évaluation, étant donné que: (1) s'ils ne sont pas bien conçus, ces régimes sont susceptibles d'avoir une forte incidence sur le marché unique; (2) des gains d'efficacité considérables sont possibles du fait de leurs budgets élevés; et (3) parce qu'ils s'adressent à différentes catégories de bénéficiaires, ils peuvent fournir une quantité suffisante de données en vue d'une évaluation.

Certains régimes d'aides d'État sont susceptibles d'échapper à toute évaluation dès lors que, nonobstant leur taille, ils ne comportent aucun problème spécifique (cas ordinaires, cas où un nombre important de bénéficiaires reçoivent une aide de faible montant, cas où aucun changement important ou aucune distorsion de concurrence sévère ne risque de survenir).

Qui plus est, le nouveau règlement général d'exemption par catégorie (RGEC) définit les régimes d'aides de grande ampleur sur la base de leur budget (budget annuel moyen dépassant 150 millions d'euros) et prévoit une évaluation pour certaines catégories d'aide<sup>13</sup>.

---

13 Aides à finalité régionale (à l'exception des aides au fonctionnement à finalité régionale), aides en faveur des PME, aides en faveur de l'accès des PME au financement, aides à la recherche, au développement et à l'innovation, aides à la protection de l'environnement (à l'exception des aides sous forme de réductions de taxes environnementales au titre de la directive 2003/96/CE) et aides en faveur des infrastructures à haut débit.

Afin de ne pas retarder l'entrée en vigueur de ces régimes de grande ampleur, mais aussi pour garantir qu'ils feront l'objet d'une évaluation efficace, le nouveau RGEC prévoit une exemption de notification pendant une période maximale de six mois, susceptible d'être prorogée par la Commission après approbation du plan d'évaluation<sup>14</sup>. Ce dernier sera notifié dès que possible et au plus tard dans les 20 jours ouvrables suivant l'entrée en vigueur du régime d'aides.

Le nouveau RGEC prévoit également des cas de modification ou de remplacement des régimes d'aides de grande ampleur soumis à évaluation. Ces cas devront également être notifiés, sauf lorsque les modifications sont de nature purement formelle ou administrative ou opérées dans le cadre de mesures cofinancées par l'UE.

#### **4.2 Régimes d'aides présentant des caractéristiques nouvelles**

La définition de la notion de «nouveau» est susceptible de varier d'un instrument d'aide à un autre ou d'un État membre à un autre. En principe, elle portera sur la nature du régime d'aides ou des marchés ciblés par ce dernier, par exemple des marchés émergents qui ne sont encore qu'à un stade très précoce de leur développement. Ces régimes ont le pouvoir de façonner durablement et profondément les contours de ces industries naissantes. C'est dire si l'ampleur des avantages qu'ils confèrent mais aussi des risques de distorsion qu'ils induisent est grande. La nouveauté peut par exemple résider dans la mise en place d'un nouveau mécanisme sur les capacités énergétiques, des aides à de nouvelles technologies ou un nouveau régime de soutien aux énergies renouvelables, dans le cadre des aides en faveur de l'environnement. L'évaluation des régimes d'aides présentant des caractéristiques nouvelles permettra en tout état de cause à leurs concepteurs de tenir compte des dernières évolutions sur le marché.

#### **4.3 Régimes d'aides concernés par des changements importants prévisibles**

La possibilité de changements importants (en ce qui concerne la technologie, la réglementation ou le marché) dans un proche avenir sera évaluée au cas par cas. Ces changements importants peuvent notamment prendre la forme d'une révision anticipée de la réglementation applicable ou d'aides à des secteurs d'activité évoluant rapidement dans lesquels l'environnement commercial et les technologies disponibles se renouvellent sans cesse. Si les régimes n'ont pas été conçus pour s'adapter aux conséquences de ces changements, les fonds publics risquent de ne pas être utilisés à bon escient (par exemple ils risquent de cibler une «défaillance du marché» potentielle appelée à disparaître rapidement). De graves distorsions risquent également d'apparaître et de pénaliser davantage les nouveaux entrants sur le marché que les opérateurs historiques, voire de créer des inégalités entre les nouvelles technologies et les technologies moins récentes. La révision d'un cadre réglementaire existant (par exemple dans le secteur des communications électroniques), la forte fluctuation des prix des intrants ou des extrants (par exemple dans le cas des panneaux

---

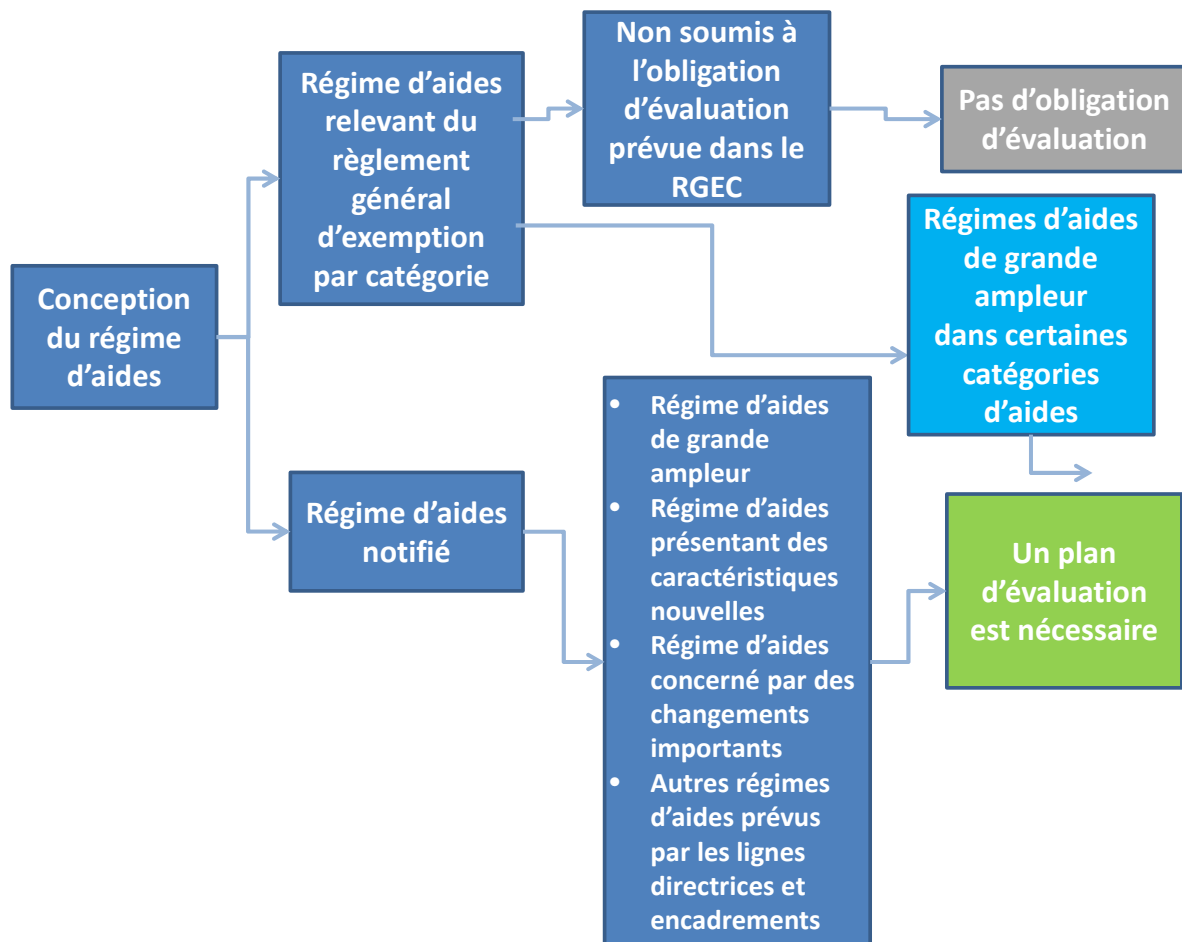
<sup>14</sup> La Commission pourrait aussi décider, à titre exceptionnel, qu'en égard aux particularités de l'espèce, l'évaluation n'est pas nécessaire.



solaires) ou le lancement d'une nouvelle technologie sur le marché (par exemple la mise en place du réseau de quatrième génération pour la téléphonie mobile à haut débit) sont autant de cas où une évaluation se justifie pleinement si l'on veut que les futurs régimes d'aides tiennent compte de l'évolution du marché.

#### 4.4 Autres régimes d'aides

Les lignes directrices applicables aux différents types d'aides d'État recensent un certain nombre de régimes d'aides pour lesquels une évaluation serait particulièrement utile.



**Graphique 3: Sélection de régimes d'aides aux fins d'une évaluation**

## **Annexe I:     Annexe technique concernant les méthodes d'identification de l'incidence causale**

Tout régime d'aides d'État est susceptible d'avoir une incidence à différents niveaux. En particulier, il est censé avoir un effet direct au niveau du bénéficiaire. Comprendre l'ampleur de cet effet est indispensable pour juger de l'efficacité et de l'efficience d'une mesure publique. Cependant, dans la mesure où les aides sont destinées à des entreprises qui interagissent sur des marchés ou dans des régions rivalisant entre elles pour attirer les activités économiques, les aides d'État ont aussi, en principe, des effets indirects. Ces derniers peuvent prendre la forme de retombées sur l'activité d'autres entreprises (retombées positives issues de la recherche et du développement mais aussi éviction des investissements consentis par des entreprises concurrentes) ou d'effets de déplacement (déplacements de l'activité économique d'une région vers une autre). Ces effets indirects sont à l'origine des avantages et inconvénients induits par les interventions des pouvoirs publics dans l'économie. C'est la raison pour laquelle l'évaluation des interventions publiques doit également se pencher sur l'ampleur de ces effets indirects.

Le recours à différents instruments est, en principe, nécessaire pour mesurer les effets directs et indirects d'une politique. De nombreuses méthodes et techniques destinées à apprécier les effets directs des politiques sur leurs bénéficiaires ont été élaborées au cours des décennies passées. Elles sont décrites en détail plus loin dans cette section. Malheureusement, il est assez rare que ces mêmes techniques permettent aussi d'évaluer les effets indirects des régimes d'aides sur les entreprises ou les régions. L'évaluation des effets indirects des régimes d'aides d'État passe en règle générale par d'autres types de preuves que celles utilisées pour évaluer les effets directs sur les bénéficiaires, et leur interprétation repose, en principe, davantage sur la théorie économique et la modélisation. Il est plus difficile de fournir des orientations précises sur ce type d'exercice, étant donné qu'il doit être conçu sur mesure pour pouvoir évaluer les effets, potentiels ou escomptés, positifs ou négatifs, d'une politique. Aussi cette évaluation doit-elle s'appuyer sur une analyse soignée et rigoureuse des effets indirects les plus vraisemblables de la politique en question. Forts de cette analyse, les évaluateurs pourront ensuite effectuer des mesures, notamment à partir des microdonnées émanant des non-bénéficiaires, en particulier dans la même région, la même filière ou le même secteur, ainsi que dans les régions voisines. Cette démarche formera le cœur de l'évaluation des effets indirects des régimes d'aides d'État. Au besoin, elle pourra être complétée par des données macroéconomiques et, plus important encore, des études de cas sélectionnées avec soin.

Si l'évaluation des effets directs constitue une première étape nécessaire et capitale, celle des effets indirects constitue également – à condition d'être rigoureuse – un élément important du processus d'évaluation des effets du régime d'aides dans leur ensemble. Et si l'absence d'investissements supplémentaires de la part des bénéficiaires de l'aide témoigne globalement de l'échec d'une politique, l'existence d'un effet positif ne suffit pas non plus à conclure que la politique en question a atteint ses objectifs. En particulier, s'il s'avère que l'incidence

directe de l'aide sur les bénéficiaires est extrêmement modeste voire inexistante, le régime d'aides sera très probablement considéré comme n'ayant pas atteint ses objectifs, sauf si de solides arguments peuvent être avancés pour prouver l'existence d'effets indirects bénéfiques sensibles. Le contraire est vrai également: si l'évaluation met au jour les effets directs positifs de l'aide, il reste à déterminer si d'éventuels effets indirects négatifs sont susceptibles de les contrebalancer, voire de l'emporter sur les effets positifs.

Par ailleurs, il n'est pas toujours aisé d'opérer une distinction claire entre effets directs et indirects. Une entreprise peut avoir investi davantage (effets directs présumés) parce que ses propres investissements ont évincé les investissements effectués par des entreprises concurrentes (interaction des effets indirects). Elle peut aussi avoir investi davantage parce qu'elle anticipe un effet de contagion ou s'attend à des investissements de la part d'autres entreprises. Il n'est pas non plus exclu que l'aide en soi ou le simple octroi d'une aide puisse avoir l'un ou l'autre effet. L'existence probable, la nature et la portée escomptée des effets indirects doivent dès lors être examinées en détail dans l'évaluation des effets directs. La théorie économique qui établit un lien entre l'aide et les effets indirects doit être clairement exposée et les informations complémentaires susceptibles de confirmer le bien-fondé de cette théorie doivent faire partie intégrante de l'évaluation<sup>15</sup>.

### *Inférence causale*

L'incidence causale de l'aide représente la différence entre le résultat obtenu grâce à l'aide et celui obtenu en l'absence d'aide. Le résultat obtenu grâce à l'aide est observé auprès des entreprises bénéficiaires de l'aide, mais le résultat obtenu en l'absence d'aide n'est mesuré qu'auprès d'entreprises non bénéficiaires. Le résultat qu'auraient obtenu les entreprises bénéficiaires en l'absence d'aide n'est donc, par définition, pas mesuré. Pour estimer l'effet de l'aide sur les entreprises qui en bénéficient, il convient donc de construire un scénario *contrefactuel*, c'est-à-dire un scénario plausible rendant compte de ce qui serait arrivé aux entreprises bénéficiaires si ces dernières n'avaient pas reçu l'aide en cause. Cette construction nécessite de trouver un *groupe de contrôle*, autrement dit un groupe d'entreprises affichant à tous égards la similarité la plus grande possible avec le groupe d'entreprises bénéficiaires – hormis le fait qu'elles n'ont pas bénéficié de l'aide en question.

La qualité du groupe de contrôle est une condition fondamentale pour la validité de l'évaluation. Les entreprises qui bénéficient d'aides présentent généralement des caractéristiques différentes de celles des entreprises non bénéficiaires. Elles opèrent, par exemple, dans des zones plus pauvres offrant un plus faible potentiel commercial; elles ont un accès plus limité au crédit, peuvent être plus ou moins efficaces, n'ont pas forcément de projet à développer, etc. Dès lors, une évaluation qui se bornerait à comparer des entreprises bénéficiaires à des entreprises non bénéficiaires risquerait de rendre compte d'abord et avant tout de cette réalité, au détriment de l'effet proprement dit de la politique mise en œuvre.

---

15 Bien que le présent document concerne essentiellement les effets directs des aides, le simple fait que ces aides puissent avoir des effets indirects pose un défi analytique à l'évaluation des effets directs. Une attention toute particulière devra par ailleurs être portée aux effets des interactions sur les marchés.

La difficulté majeure de toute évaluation consiste donc à s'assurer que cette différence systématique (appelée effet de sélection) entre les bénéficiaires et les non-bénéficiaires d'aides d'État ne faussera pas les résultats. Différentes méthodes à la fiabilité éprouvée ont été mises au point à cette fin au cours des dernières décennies. Leur choix dépend de la politique à évaluer et des données disponibles. Chacune d'elles a néanmoins ses propres limites et n'est valable que si certaines hypothèses se vérifient. Identifier et traiter expressément ces limites peut donc contribuer à accroître la crédibilité d'une étude. La présente annexe technique passe en revue les méthodes les plus pertinentes. Elle se propose d'en analyser avant tout les aspects les plus pratiques, tout en soulignant l'importance d'une bonne stratégie d'identification<sup>16</sup>.

## **A. Expériences aléatoires**

Pour obtenir une estimation correcte (autrement dit non faussée) des effets d'une politique, il est impératif de disposer d'un groupe de contrôle approprié. Le cas le plus favorable à cet égard est celui d'une absence totale d'effet de sélection du fait de la sélection aléatoire des bénéficiaires<sup>17</sup>. Il n'existe alors aucune différence systématique entre les bénéficiaires et les non-bénéficiaires, à l'exception de l'aide elle-même, et les résultats observés peuvent être attribués à la politique évaluée.

Toutefois la sélection aléatoire des bénéficiaires est parfois jugée contradictoire avec la finalité des régimes d'aides, lesquels visent à sélectionner les meilleurs bénéficiaires possibles sur la base de critères objectifs. Il n'en reste pas moins envisageable, dans certaines circonstances, de donner un caractère relativement aléatoire aux critères d'admissibilité ou aux éléments destinés à inciter les entreprises à participer. Une possibilité consiste, par exemple, à allouer un budget fixe au régime concerné. Si, à caractéristiques globalement égales, les demandes de soutien des candidats dépassent le budget proposé, il est envisageable de rendre le traitement aléatoire. Une autre solution consiste à exposer, de manière aléatoire, les bénéficiaires potentiels à différents niveaux d'information à propos du régime.

Les projets pilotes peuvent fournir d'autres occasions de procéder à une attribution aléatoire des aides. Ainsi, dans le cas de politiques innovantes, peut-il s'avérer judicieux d'évaluer, en premier lieu, un projet pilote de faible envergure, dans lequel il sera plus facile d'opérer une sélection aléatoire des bénéficiaires. Une autre piste envisageable est celle d'un déploiement progressif du régime d'aides, visant à rendre admissibles 25 % des entreprises sélectionnées aléatoirement la première année, puis 50 % la deuxième, 75 % la troisième et enfin 100 % la quatrième (ou à promouvoir le régime auprès d'un public de plus en plus large). Un déploiement progressif s'impose souvent pour des raisons administratives dans le cas des politiques entièrement nouvelles.

---

16 La présente annexe offre une présentation rapide et non technique des méthodes économétriques utilisables pour l'évaluation des politiques, dont de nombreux éléments ont été puisés dans Givord (2010). Imbens et Wooldridge (2009) ou encore Angrist et Pischke (2008) proposent eux aussi des présentations de très grande qualité.

17 Pendant des dizaines d'années, les expériences aléatoires ont constitué les seules méthodes acceptables pour évaluer les effets des médicaments et des traitements médicaux.

Une approche de ce type peut se révéler particulièrement adaptée dans le cas de régimes totalement nouveaux ou d'importantes modifications de régimes existants. Et s'il est probablement difficile de rendre l'admissibilité (directement ou indirectement) aléatoire lors de la poursuite d'un régime existant, les expériences aléatoires n'en restent pas moins en partie utiles aux fins de l'évaluation. En particulier, il est possible d'opérer une sélection aléatoire des bénéficiaires pour mettre en place des variantes plus efficaces, mieux ciblées et/ou provoquant moins d'effets de distorsion. Il est par exemple possible de remplacer aléatoirement un programme de subventions par un régime de prêts nouvellement créé.

## **B. Méthodes quasi expérimentales**

Bien qu'elles constituent le meilleur outil d'évaluation de l'effet des politiques publiques, il n'est pas toujours possible de procéder à des expériences aléatoires. Aussi d'autres méthodes ont-elles été conçues pour permettre une analyse ex post de ces effets. Toutes se proposent d'utiliser les variations exogènes de l'environnement dans lequel les entreprises opèrent et de créer des situations quasi expérimentales. C'est pourquoi on les qualifie d'expériences naturelles ou de quasi-expériences.

Il est généralement très difficile d'identifier des expériences naturelles ou des quasi-expériences dans le cadre d'évaluations ex post. Une analyse approfondie de la façon dont la politique est conçue peut être l'occasion d'étudier la présence éventuelle d'une variation exogène suffisante. S'il y a lieu, la configuration initiale peut également être adaptée afin d'introduire davantage d'éléments permettant d'identifier les effets de la politique.

### *Contrôle des différences observables*

Comme expliqué plus haut, les bénéficiaires et les non-bénéficiaires de régimes d'aides présentent normalement des différences majeures. Des différences dont il est nécessaire de tenir compte lors de la comparaison des résultats des deux groupes.

La plupart des différences entre caractéristiques sont généralement observables. La méthode la plus habituelle pour prendre en compte ces différences consiste à opérer une régression linéaire. La régression linéaire vise à contrôler l'influence des caractéristiques observables sur les résultats. Elle suppose l'existence d'une relation linéaire entre le résultat (par exemple des investissements en recherche et développement) et d'autres caractéristiques de l'entreprise étudiée, comme le secteur d'activité, l'ancienneté, la taille, etc., y compris l'octroi de l'aide. La régression linéaire peut également être envisagée comme une approximation linéaire de relations plus complexes<sup>18</sup>. Utilisables dans des cadres d'évaluation très différents, les régressions linéaires sont des techniques polyvalentes.

Les techniques d'appariement constituent d'autres méthodes envisageables. Elles consistent à appairer chaque bénéficiaire avec une entreprise apparemment très similaire, qui n'a pas

---

<sup>18</sup> De plus, il est possible de faire entrer des caractéristiques en interaction (par exemple le chiffre d'affaires avec le secteur d'activité) et d'introduire des fonctions de ces caractéristiques (par exemple des carrés de variables).

bénéficié de l'aide. Les éléments observables utilisés pour cet appariement peuvent être les caractéristiques des entreprises ou la probabilité estimée de recevoir l'aide (appariement par score de propension). L'appariement peut s'avérer utile pour le contrôle d'éléments observables dans le cadre de stratégies empiriques valables.

Quoi qu'il en soit, tant la régression linéaire simple que l'appariement sont intrinsèquement limités. Leur validité n'est assurée que dans l'hypothèse dite d'une indépendance conditionnelle. Cette condition nécessite qu'une fois l'incidence des caractéristiques observables prise en compte, le résultat soit indépendant de ces dernières. Concrètement, elle implique que chaque variable ayant une incidence à la fois sur le résultat et sur la sélection soit observable (et prise en compte dans la forme fonctionnelle appropriée). Si tel n'est pas le cas, le simple fait qu'une entreprise soit participante traduit certaines caractéristiques (non observées) de cette entreprise, qui déterminent aussi sa performance. Ni la régression linéaire ni l'appariement ne peuvent alors fournir une évaluation valable. Par exemple, le fait qu'une entreprise ait un «projet prometteur» aura une influence tant sur la probabilité de sa candidature (et de son acceptation dans le régime d'aides) que sur sa probable capacité à développer ses activités avec succès. Ne pas en tenir compte fausserait les résultats.

En matière d'appariement, la comparaison des résultats obtenus par le bénéficiaire et son «jumeau» non bénéficiaire ne permet d'éviter l'effet de sélection que si l'octroi de l'aide n'est pas lié à des caractéristiques non observées qui influencent également le résultat. En réalité, cette hypothèse se vérifie rarement. Mesurer l'ensemble des variables ayant une incidence sur le fait de demander ou d'obtenir une aide se révèle rarement possible. L'application de méthodes d'appariement nécessite en outre que les entreprises bénéficiaires d'une aide présentent des caractéristiques observables très proches de celles des entreprises non bénéficiaires. Si les entreprises appariées possèdent réellement des caractéristiques observables similaires, les motifs pour lesquels l'une bénéficie de l'aide et l'autre pas seront, par définition, impossibles à observer. La validité d'une évaluation par appariement ou d'une régression linéaire plus simple ne saurait donc reposer sur la simple existence d'un jeu de données exhaustif comportant un grand nombre de caractéristiques observées.

Au contraire, le recours à la technique de l'appariement ou à la régression linéaire ne peut se justifier que par le fait que les motifs non observés qui expliquent l'admissibilité d'une entreprise ou l'attribution d'une aide n'ont aucune influence directe ni indirecte sur les résultats (une fois les éléments observables contrôlés). Pour garantir la validité d'une évaluation fondée sur l'appariement ou sur la régression linéaire, il faut avoir la certitude que le groupe d'entreprises non bénéficiaires a été déterminé de manière exogène. Cela implique qu'une fois les caractéristiques observables contrôlées, il ne subsiste aucun facteur non observé susceptible d'expliquer l'admissibilité ou l'attribution de l'aide et d'exercer une influence directe ou indirecte sur les résultats. En règle générale, l'appariement d'entreprises admissibles dans les mêmes conditions ne permet pas de remplir ce critère. Par exemple, si toutes les entreprises sont admissibles, on peut vraisemblablement conclure que les

entreprises bénéficiaires avaient un projet à proposer et que les entreprises non bénéficiaires n'en avaient pas (dans le cas contraire, elles auraient aussi demandé et obtenu une aide). Globalement, les entreprises ayant un projet à développer ont plus de chances de prospérer et de voir leur chiffre d'affaires comme leurs effectifs augmenter, mais cette probabilité n'est pas liée à l'octroi d'aides. Il est donc impossible de dissocier les deux aspects au moyen d'un appariement fondé sur des caractéristiques observables (à moins qu'il ne s'agisse de mesurer l'existence d'un projet d'investissement comparable).

L'hypothèse de l'indépendance conditionnelle étant, dans de nombreux cas, vouée à l'échec, il peut se révéler nécessaire de recourir à des techniques autres que l'appariement ou la simple régression linéaire pour tenir compte de l'existence d'une sélection non observée dans le traitement.

La suite de la présente section décrit en détail les méthodes les plus couramment utilisées dans l'évaluation de l'incidence des politiques, comme la méthode des doubles différences, la régression par discontinuité (RD) ou encore les variables instrumentales. Ces méthodes tirent leur validité d'hypothèses différentes. Leur choix dépend généralement du contexte dans lequel s'inscrit la politique évaluée et de la disponibilité des données. Nous passerons en revue les avantages et les inconvénients de chacune d'elles. À l'exception notable des essais aléatoires contrôlés présentés ci-dessus, aucune technique ne surpasse globalement les autres. Le choix de la technique à employer doit être guidé par une analyse approfondie du contexte particulier de la mesure ainsi que des données disponibles.

Il convient de souligner ici que ce n'est pas tant la technique économétrique utilisée qui permet de mettre en évidence les effets d'une politique, mais l'exogénéité du groupe de contrôle et, partant, la qualité du scénario contrefactuel. La qualité de l'étude d'évaluation dépendra donc en grande partie de la capacité du chercheur à établir de manière convaincante l'exogénéité du groupe de contrôle. Si certains biais subsistent, il est impératif de les examiner en détail, en analysant leurs sources, leur direction et l'ampleur potentielle de leurs effets sur les résultats.

#### **a. Doubles différences**

##### *Logique et identification*

Comme expliqué plus haut, la simple comparaison entre un groupe de bénéficiaires et un groupe de non-bénéficiaires, même sélectionné avec soin, a peu de chances de produire une évaluation pertinente. En effet, on ne saurait exclure l'existence de différences non observées entre les deux groupes, susceptibles d'entraîner des écarts de résultats même en l'absence d'aide. Il ne suffit pas non plus de comparer les résultats avant et après l'attribution d'aides aux bénéficiaires, car une telle comparaison ne permet pas de distinguer les effets produits par l'aide de ceux d'autres facteurs susceptibles d'influer sur les résultats des deux groupes, comme par exemple la conjoncture économique globale, l'évolution du cadre réglementaire ou encore l'augmentation des coûts du travail.

Conjuguer les deux approches permet néanmoins d'évaluer l'effet causal de l'aide: c'est la méthode dite des doubles différences. Celle-ci consiste à évaluer sur le long terme les écarts de résultats observés entre les entreprises. Les différences préexistantes étant attribuées à des facteurs autres que les aides d'État, seule l'évolution des écarts de résultats (les doubles différences) sera ainsi imputée aux aides d'État. En d'autres termes, la méthode vise à comparer l'écart de performance entre les bénéficiaires et le groupe de contrôle *avant* et *après* l'octroi des aides, puis à imputer l'évolution observée de cet écart aux aides d'État. La méthode est efficace si tant les bénéficiaires que le groupe de contrôle sont affectés, sur la durée, par les autres facteurs et que ces facteurs influent sur leur performance de la même manière. Dans ces conditions, il est possible de conclure que les aides constituent l'unique facteur pertinent susceptible d'expliquer l'évolution observée dans la performance des bénéficiaires par rapport au groupe de contrôle.

La méthode repose sur l'hypothèse fondamentale suivante: les écarts entre les bénéficiaires et le groupe de contrôle demeurent stables dans le temps et les deux groupes sont touchés de la même manière par des chocs communs (écarts par rapport à la moyenne) au cours de la période prise en compte. Reste que cette hypothèse peut se révéler inexacte dans la pratique. Si le groupe des bénéficiaires est constitué des entreprises les plus vulnérables, ces dernières sont susceptibles d'être davantage touchées par un ralentissement de l'économie ou la morosité du climat des affaires. Aussi le groupe de contrôle devra-t-il lui aussi comprendre des entreprises vulnérables. On le voit, la validité de la méthode dépend essentiellement du choix du groupe de contrôle. L'identification des effets d'une politique ne réside pas dans l'utilisation des doubles différences (simple mise en œuvre technique), mais dans le choix du groupe de contrôle adéquat.

Il convient d'accorder un soin tout particulier à la constitution du groupe de contrôle si les non-bénéficiaires décident eux-mêmes de ne pas demander des aides. On peut estimer que le fait de demander ou non des aides d'État dépend du bénéfice escompté de ces aides. Aussi y a-t-il tout lieu de croire que les résultats escomptés par les entreprises qui ne demandent aucune aide (sur le plan de l'emploi, de la productivité, du chiffre d'affaires, etc.) diffèrent des résultats escomptés par les bénéficiaires. À titre d'exemple, si toutes les entreprises qui demandent des aides en obtiennent, les seules entreprises admissibles qui n'en demandent pas sont celles qui n'ont pas de projet à soutenir (dans l'hypothèse où les frais liés à une demande d'aide sont peu élevés). Ces entreprises sont non seulement vraisemblablement moins performantes en termes absolus, mais également, sur la durée, en termes comparatifs, dans la mesure où les autres entreprises mettront en œuvre leurs projets et prospéreront. Il est logique que l'évolution de l'emploi, de la productivité et du chiffre d'affaires ne demeure pas parallèle. Or la double différenciation ne résout pas, en général, le problème.

Dans ces conditions, il faut que les entreprises du groupe de contrôle qui n'ont pas bénéficié d'aides aient été sélectionnées pour des motifs dépourvus d'effets sur les résultats mesurés. Ces entreprises ne peuvent pas s'être auto-sélectionnées et avoir décidé elles-mêmes, volontairement, de ne pas participer au régime. La configuration la plus convaincante est celle d'une non-participation liée à une non-admissibilité consécutive à une expérience naturelle.



Dans ce cas, il est peu probable que la non-admissibilité soit due à des facteurs non observés exerçant eux aussi une influence sur les résultats. Les groupes de contrôle pourraient ainsi être constitués, par exemple, d'entreprises situées dans des régions qui ne sont plus admissibles au bénéfice du régime (à condition que cette admissibilité ne soit pas liée à leurs performances mais à un événement exogène).

### *Mise en œuvre*

Sur le plan technique, la méthode des doubles différences peut être mise en œuvre soit dans le cadre d'un modèle de régression linéaire soit par appariement. Dans le premier cas, le groupe de contrôle est choisi indépendamment des caractéristiques observables et est ainsi globalement comparable à l'ensemble du groupe des bénéficiaires. Les différences observables sont alors prises en compte dans une régression linéaire classique. Dans le second cas, le groupe de contrôle est constitué d'entreprises comparables individuellement à chacune des entreprises bénéficiaires au sein de l'échantillon, sur la base des facteurs observables. Le résultat de chaque entreprise est comparé au résultat de la ou des entreprises qui lui sont le plus comparables et les résultats sont ensuite agrégés. Si les deux méthodes proposent des moyens distincts de prendre en compte les différences observables, elles ne divergent toutefois pas fondamentalement l'une de l'autre en ce qui concerne l'identification de l'effet causal de la politique.

Selon les circonstances, il peut être intéressant de comparer les variations de résultats entre les bénéficiaires et le groupe de contrôle avant l'octroi de l'aide. Si les résultats diffèrent systématiquement avant même l'octroi de l'aide, il est probable que le groupe de contrôle et le groupe des bénéficiaires divergent pour des motifs n'ayant aucun lien avec l'aide et que la méthode ne permette dès lors pas d'évaluer valablement l'effet causal de cette aide. Bien que ne permettant pas de tester rigoureusement la validité de l'hypothèse (le test n'existant pas), cette démarche offre à tout le moins un contrôle de bon sens non dénué d'utilité.

Des méthodes complémentaires ainsi que des tests de robustesse peuvent être appliqués lorsque plusieurs groupes de contrôle potentiels a priori valables coexistent. Le test de robustesse le plus naturel consiste à mettre en œuvre plusieurs estimateurs de doubles différences et à comparer les résultats. Il est également possible d'utiliser ces différents groupes de contrôle pour accroître la fiabilité de l'estimation. Prenons l'exemple d'un régime d'aides destiné aux PME dans une région donnée. Deux groupes de contrôle sont envisageables: l'un constitué d'entreprises qui ne sont pas des PME, l'autre constitué de PME situées dans une région voisine. Aucune de ces entreprises n'a volontairement décidé de ne pas demander d'aide. Elles ne remplissent simplement pas les critères d'admissibilité. Cependant, les deux groupes de contrôle sont imparfaits: les entreprises de plus grande taille de la région sont susceptibles de réagir différemment à la conjoncture économique générale et les PME de la région voisine risquent d'être soumises à des chocs régionaux spécifiques. Plutôt que de choisir entre ces deux estimateurs par doubles différences, il est possible de les combiner et de mettre ainsi en œuvre un estimateur de triples différences: en partant de l'estimation classique des doubles différences entre des PME et des entreprises autres que des PME dans la région concernée, il est possible de déduire la même double différence de la

région voisine afin d'annuler la variation de résultat entre les PME et les entreprises autres que des PME dans la région bénéficiant d'aides<sup>19</sup>. Il est également possible de tenter de constituer un groupe de contrôle synthétique composé de PME de plusieurs régions voisines et d'entreprises autres que des PME implantées dans la région concernée, afin de reproduire plus fidèlement le profil de résultats des bénéficiaires avant l'octroi des aides (pour plus de détails, voir Abadie, Diamond et Hainmuller, 2010).

### *Inférence*

Outre la conception et le choix rigoureux du groupe de contrôle, se pose également tout particulièrement la question de l'inférence. Dans le contexte qui nous occupe, la notion d'inférence se rapporte au fait de savoir si les effets estimés sont réellement significatifs. La signification statistique diffère de la signification économique. Cette dernière se rapporte à l'ampleur des effets estimés par comparaison avec d'autres paramètres pertinents sur le plan de la théorie économique. Comme expliqué plus haut, la signification économique revêt une importance cruciale. Néanmoins, ce débat n'a en principe un intérêt que si, d'un point de vue statistique, les effets font l'objet d'une estimation suffisamment précise, autrement dit que l'on peut exclure l'absence d'effets.

Il y a tout lieu de croire que dans le cas d'hypothèses standard (comme celle d'une homoscedasticité ou d'une absence d'autocorrélation), une inférence simple risque de surestimer la signification statistique des effets<sup>20</sup>.

Le premier problème tient au regroupement des données. Si le groupe de contrôle et le groupe des bénéficiaires présentent chacun une forte homogénéité (tout en étant distinct l'un de l'autre), toutes les entreprises, au sein de chaque groupe, risquent d'être affectées par des écarts identiques par rapport à la moyenne (chocs). En termes statistiques, cela signifie que le terme d'erreur a une composante commune. Si la variance de cette composante commune est grande par rapport à la variance de résultat observée pour chaque entreprise, l'inférence sera faussée. Le problème peut se révéler particulièrement épineux et s'apparenter à un problème d'identification avec deux périodes et deux groupes, car il devient impossible de distinguer l'effet des chocs affectant l'ensemble des entreprises au sein de chaque groupe de l'effet de la politique. Le problème s'atténue lorsque les groupes présentent une moindre homogénéité. Mais il n'en reste pas moins nécessaire de s'interroger sur l'existence de chocs communs à l'intérieur des sous-ensembles – homogènes – constitués au sein de chaque groupe. Si la demande est locale, il conviendra en principe de compenser la structure agrégée du terme d'erreur au niveau des localités. Il en ira de même au niveau des industries ou des secteurs.

---

19 Prenons l'exemple d'un régime régional d'aides aux PME destiné à favoriser les créations d'emploi. Supposons qu'à la fin du régime, il s'avère que les PME de la région ont créé 20 % d'emplois en plus que les grandes entreprises implantées dans la région. Si, dans une région voisine comparable (mais non bénéficiaire du régime d'aides), les PME ont également créé davantage d'emplois que les grandes entreprises (par exemple 15 %), l'incidence du régime d'aides pourra être estimée à 5 % environ.

20 Ce problème a été largement mis en évidence dans le cas de la méthode des doubles différences, mais il peut également se poser avec les autres méthodes étudiées dans le présent document.

Un deuxième problème apparaît en cas d'utilisation de données de panel. Les termes d'erreur de la plupart des données d'entreprises, comme l'emploi, la productivité et l'investissement, sont normalement autocorrélés, ce qui signifie que les écarts par rapport à la moyenne observés sur une période sont de nature à se prolonger au cours de la période suivante. Ne pas en tenir compte reviendrait à surestimer la précision de l'estimation des effets et à rejeter trop souvent l'hypothèse d'une absence d'effet de la politique. Comme l'ont montré Bertrand, Duflo, et Mullainathan (2004), ce problème peut s'avérer important.

## **b. Variables instrumentales**

### *Logique et identification*

Les variables instrumentales constituent une méthode classique pour traiter les problèmes d'endogénéité des variables explicatives. Le fait de bénéficier d'aides pouvant être considéré, dans le cadre d'une régression linéaire, comme une variable explicative endogène de la performance d'une entreprise, il est naturel d'utiliser des variables instrumentales afin d'évaluer l'effet de ces aides.

Une variable est dite endogène lorsqu'elle est corrélée à un élément non observé qui détermine également le résultat. Imaginons par exemple que l'on tente de mesurer les effets d'une subvention publique sur l'emploi au sein des entreprises en opérant une régression de l'emploi par rapport à la participation au programme et à d'autres éléments observables. Imaginons par ailleurs que le programme d'aides cible en particulier les entreprises peu efficaces confrontées à des conditions défavorables sur le marché local. L'évaluateur n'est pas en mesure d'observer les conditions du marché, de sorte que celles-ci ne peuvent pas être contrôlées directement. Néanmoins, si l'on ignore cette variable, les effets de la subvention risquent d'être sous-estimés par l'évaluateur en raison de l'endogénéité de la participation au programme. Le fait qu'une entreprise soit soumise à des conditions du marché défavorables ou au contraire favorables a une incidence à la fois sur sa participation au programme et sur le recrutement qu'elle opère. Autrement dit, la participation au programme est corrélée au terme d'erreur qui explique l'emploi. L'incidence des conditions du marché sur la participation au programme fait qu'il est impossible d'attribuer la totalité de la corrélation entre la participation au programme et l'emploi à l'effet causal de la subvention.

Il existe néanmoins d'autres facteurs susceptibles d'expliquer la participation au programme mais non l'emploi. Par exemple, comme dans Criscuolo et al. (2012), l'implantation géographique peut déterminer le montant total alloué au programme dans la région. La liste des régions couvertes par le programme peut également évoluer dans le temps. La variation dans le temps du budget alloué à une région donnée pour des raisons externes (par ex. la chute du PIB moyen par habitant au sein de l'UE) peut en outre avoir des effets sur la participation au programme, mais non sur les conditions du marché rencontrées par les entreprises au niveau local. L'évolution de l'emploi liée aux changements exogènes affectant la couverture du programme n'a pas de rapport avec les conditions du marché locales. En ciblant cette «partie» de la variable de participation, il est possible d'isoler l'incidence réelle de la

participation au programme sur l'emploi au sein de l'entreprise, sans interférence de la part des conditions du marché locales. Telle est la raison d'être des variables instrumentales.

Aux fins de l'évaluation des aides d'État, une variable instrumentale est une variable qui peut expliquer le fait de recevoir ces aides, mais qui n'a pas d'incidence directe sur les autres déterminants non observés du résultat à mesurer. Les variables instrumentales permettent dès lors de se concentrer sur la participation au régime d'aides sans que les effets de la sélection n'interfèrent. À titre indicatif, on peut envisager la logique qui sous-tend les variables instrumentales comme suit<sup>21</sup>: on opère dans un premier temps une régression de la participation au programme sur l'ensemble des variables exogènes, y compris les variables instrumentales. Puis la variable de participation (autrement dit la variable indiquant si l'aide a été reçue) est remplacée par la participation telle qu'escomptée lors de la première étape: cette participation escomptée n'est pas corrélée à l'élément non observé qui détermine également le résultat.

### *Problèmes liés aux instruments dits faibles*

Pour simple et classique qu'elle soit, la définition de la variable instrumentale envisagée comme une variable susceptible d'expliquer la réception des aides mais dépourvue d'incidence directe sur les autres déterminants non observés du résultat à mesurer, recèle un certain nombre de difficultés pratiques. En effet, s'il existe des tests qui permettent de vérifier la cohérence des instruments lorsque ces derniers sont plus nombreux que ce qui est strictement nécessaire pour identifier un modèle, il n'existe pas de test pour s'assurer de la validité desdits instruments. En général, les études utilisant des variables instrumentales ont pour but principal d'expliquer pourquoi chaque instrument pris individuellement peut être réputé non corrélé aux déterminants non observés de la performance des entreprises, qu'il s'agisse de l'emploi, de la productivité, du chiffre d'affaires, de l'investissement, etc. Bien que nécessaires pour mesurer la validité de l'évaluation, ces explications – fondées à la fois sur des arguments économiques et des éléments factuels – ne suffisent pas, notamment lorsque plusieurs instruments sont utilisés.

Le débat sur la qualité des variables instrumentales doit également examiner la question des instruments dits faibles, autrement dit des instruments faiblement corrélés à la variable de résultat. Lorsque les variables instrumentales sont faiblement corrélées à la variable endogène, les estimations risquent d'être imprécises. On peut alors être tenté d'ajouter davantage de variables instrumentales, dans la mesure où il est notoire que le fait de se doter d'un nombre suffisant de variables permet de récupérer une part suffisante de la variable initiale pour obtenir des résultats statistiquement significatifs. Dans le même temps, l'estimateur des doubles moindres carrés tend naturellement vers l'estimateur des moindres carrés ordinaires, lequel est biaisé<sup>22</sup>. Ce risque de biais doit être expressément traité dans toutes les évaluations

---

21 Dans la pratique, les moindres carrés en deux étapes (ou doubles moindres carrés) sont mis en œuvre en une étape pour des raisons évidentes d'inférence.

22 La discussion de Bound, Jaeger et Baker's (1995) sur les biais statistiques dans Angrist et Krueger (1991) offre un débat très intéressant sur les biais créés par les instruments faibles. En outre, les estimateurs de

fondées sur la méthode des variables instrumentales. Il convient en particulier d'examiner la question de la crédibilité non seulement de leur exogénéité individuelle mais également de leur exogénéité conjointe.

Un cas particulier se présente dans l'hypothèse d'une autocorrélation de la variable endogène. À supposer que la source d'endogénéité soit exclusivement contemporaine, il est possible d'utiliser des valeurs passées comme variables instrumentales. Pour autant, il conviendra de s'interroger sur la validité de cette approche. Par exemple, si des variables explicatives sont autocorrélées, le résultat mesuré risque de l'être aussi. Dans ce cas, les variables retardées sont également endogènes. De manière générale, plus l'autocorrélation des variables explicatives est forte, plus les hypothèses en matière d'exogénéité risquent de se révéler fausses. Si elle est faible, on peut avoir recours à plusieurs décalages (voire à des valeurs futures) mais on risque de tomber dans l'écueil décrit plus haut (utilisation de nombreux instruments faibles). Dans l'ensemble, recourir à des valeurs passées peut constituer une stratégie valable, mais il convient d'en user avec précaution.

En règle générale, pour éviter les problèmes décrits plus haut, il est vivement recommandé de n'utiliser qu'un faible nombre d'instruments convaincants. Reste que dans ce cas aussi, il y a lieu de démontrer que les instruments sont de bons prédicteurs de la variable explicative endogène<sup>23</sup>.

#### *Variations des estimations en deux étapes: le modèle de sélection de Heckman (1979)*

Lorsque la variable endogène est une variable de traitement (variable muette de participation), la première régression des doubles moindres carrés peut être envisagée comme un modèle de probabilité linéaire de la probabilité d'être traité. Ce modèle de probabilité linéaire est une approximation linéaire. Cependant, il arrive dans certains cas que la probabilité d'être traité, bien que limitée aux entreprises admissibles, soit faible. L'approximation linéaire risque alors d'être trop grossière pour reproduire efficacement la probabilité d'être traité et cibler les extrémités de la distribution, qui sont justement les plus dignes d'intérêt. Ce problème peut être résolu de différentes manières, qui reposent toutes sur le remplacement du modèle de probabilité linéaire de la probabilité d'être traité par une fonction non linéaire<sup>24</sup>.

Une approche classique consiste à traiter le problème de l'évaluation dans le cadre d'un modèle de sélection (Heckman, 1979). Cette approche traite l'effet de sélection comme un problème d'omission de variables dans la régression linéaire du résultat sur les

---

variables instrumentales sont biaisés à distance finie, de sorte qu'il peut subsister d'importants biais non-asymptotiques même avec des ensembles de données suffisamment vastes pour garantir une signification statistique apparente.

23 Cette démonstration peut consister à calculer la statistique de Fisher de la régression de première étape. Plus la statistique de Fisher est élevée, moins les instruments risquent d'être faibles. Stock, Wright et Yogo (2002) proposent un test formel: pour un instrument, il est par exemple nécessaire que la statistique de Fisher de la régression de première étape soit supérieure à 10.

24 La présente section propose une description très sommaire des modèles de sélection. Pour une description plus détaillée, voir l'article fondamental de Heckman (1979) et, plus généralement, Wooldridge (2002), chapitre 17.

caractéristiques observables et la participation<sup>25</sup>. Cette méthode comporte plusieurs variantes, consistant par exemple à estimer l'ensemble du modèle par la méthode du maximum de vraisemblance ou à instrumenter l'octroi de l'aide par la valeur prédite de l'équation de sélection<sup>26</sup>.

Quoi qu'il en soit, il est absolument indispensable de se pencher sur la question de l'identification et en particulier du choix des variables car utiliser les mêmes variables lors des deux étapes de l'estimation n'est pas satisfaisant, même si les résultats offrent une précision suffisante<sup>27</sup>.

La seule approche fiable consiste à estimer un modèle de sélection à l'aide d'une variable dite d'exclusion. Une variable exclue est une variable qui explique la sélection des bénéficiaires de l'aide, mais non le résultat. Il ne suffit pas de retirer une variable de l'équation principale pour l'ajouter à la liste des variables explicatives de l'équation de sélection. Au contraire, cette variable d'exclusion doit expliquer la sélection mais ne doit pas avoir d'incidence sur le résultat que l'on tente d'expliquer. Une telle variable est par nature très proche d'une variable instrumentale valide. Son choix ne doit pas obéir à un souci de commodité mais doit être guidé par la théorie économique, la structure institutionnelle et/ou l'expérience.

### **c. Méthode de la régression par discontinuité**

La régression par discontinuité (RD) est le dernier né des outils d'évaluation<sup>28</sup>. Largement plébiscitée par les milieux scientifiques ces dix dernières années, notamment pour sa simplicité, cette méthode consiste à exploiter l'existence d'une variable exerçant une incidence discontinue sur la probabilité d'être touché par une politique. Dans le cadre des

---

25 La variable omise représente la différence dans l'espérance conditionnelle du résultat pour l'échantillon sélectionné (ici les bénéficiaires de l'aide). Dans certaines hypothèses concernant le processus de sélection des bénéficiaires de l'aide (par ex. un modèle de type probit ou logit), cette différence peut être formellement établie (rapport inverse de Mills); elle constitue une fonction des paramètres de sélection. Dès lors, il est possible d'identifier l'effet de la politique en ajoutant la variable omise à la régression. Les paramètres de sélection ne sont pas connus mais des paramètres convergents peuvent être récupérés au moyen d'une estimation de première étape du processus de sélection. Cette démarche aboutit à une procédure d'estimation parfois dénommée méthode Heckit. Celle-ci implique tout d'abord de récupérer les paramètres dignes d'intérêt aux fins de la sélection des bénéficiaires de l'aide, par exemple au moyen d'une spécification de type logit ou probit. Il est ensuite possible d'obtenir un estimateur convergent de l'effet de la politique en ajoutant l'estimation du rapport inverse de Mills à la régression linéaire. Les logiciels statistiques comportent généralement une fonction permettant d'effectuer cette estimation par la méthode de Heckman.

26 Pour une présentation complète de toutes ces méthodes, voir par exemple Wooldridge (2002), chapitre 17.

27 Lorsque l'équation de sélection n'est pas linéaire, le rapport inverse de Mills n'est pas colinéaire avec les autres variables explicatives, même si la première équation n'inclut qu'un sous-ensemble de ces variables explicatives. En théorie, le modèle est alors déjà identifié. Dans ce cas, il arrive très souvent que le ratio inverse de Mills ne présente pas une variation suffisante, ce qui aboutit à des estimations très imprécises. Reste que l'estimation pourra tout de même déboucher sur des résultats significatifs, notamment si les échantillons sont importants. Pour autant, lorsque toutes les variables de l'équation de sélection se trouvent également dans l'équation principale, le modèle n'est identifié qu'en raison des non-linéarités d'une forme paramétrique particulière.

28 Pour une description formelle et complète de la régression par discontinuité, voir Imbens et Lemieux (2008).

régimes d'aides d'État, plusieurs types de discontinuités peuvent présenter un intérêt. Le premier type regroupe les frontières géographiques: l'admissibilité au régime est liée à des frontières administratives précises (localités, régions NUTS, etc.). Le deuxième englobe les conditions imposées aux entreprises bénéficiant du régime, notamment en termes d'ancienneté et de taille.

Imaginons que des projets présentés par des entreprises soient notés sur un total de 100 points. Pour bénéficier du régime, les entreprises doivent obtenir au moins 70 points. Le projet d'une entreprise qui a obtenu 71 points est à peine meilleur que celui d'une entreprise qui a obtenu 69 points. Pourtant, cette différence marginale va avoir des conséquences majeures, dans la mesure où l'une des deux entreprises va bénéficier d'aides d'État alors que l'autre n'en bénéficiera pas. La comparaison entre les résultats des deux entreprises donne dès lors une indication très précise de l'effet causal des aides d'État.

En principe, il faut que la probabilité de bénéficier du régime d'aides soit discontinue et que toutes les autres variables soient continues<sup>29</sup>. La mise en œuvre technique de la méthode de régression par discontinuité peut être très semblable à celle des variables instrumentales, avec pour instrument une variable muette de franchissement de seuil. Il existe néanmoins deux différences majeures. Tout d'abord, l'approche de la régression par discontinuité repose sur des hypothèses plus faibles. En particulier, elle n'implique pas, a priori, l'indépendance de l'instrument. Dans le cas de la sélection par notation, par exemple, il se peut que les entreprises présentant de bons projets aient été plus nombreuses à demander une aide que les entreprises présentant de mauvais projets. La seule exigence réside dans l'absence de discontinuité de la probabilité de demander une aide dans le voisinage du seuil. La deuxième différence tient à ce que les estimations réalisées reposent uniquement sur des entreprises situées très près des deux côtés du seuil, ce qui a une contrepartie: les estimations selon la méthode de régression par discontinuité sont beaucoup plus localisées que les estimations fondées sur les variables instrumentales. Si l'on observe également des effets divergents sur les entreprises situées plus à distance du seuil, il faut en conclure que la méthode n'offre pas une estimation correcte de l'effet du régime d'aides sur l'ensemble des bénéficiaires.

Le caractère extrêmement localisé de ces estimations peut s'avérer problématique si l'on anticipe de fortes divergences à distance du seuil. Par ailleurs, certaines entreprises situées de l'autre côté de la frontière peuvent également être fortement touchées par le régime d'aides, notamment en cas d'effets de déplacement importants. Dès lors, l'application de la méthode de la régression par discontinuité près de la frontière géographique n'est pas une bonne option sur le plan empirique. Enfin, la force de la régression par discontinuité est de se concentrer sur

---

29 Sur le plan formel, on distingue deux types de régression par discontinuité: une approche précise et une approche plus imprécise. Dans l'approche précise, qui est implicitement décrite ici, seules les entreprises situées au-dessus d'un certain seuil sont traitées. Dans l'autre approche, la discontinuité est moins drastique: il existe une discontinuité de la probabilité d'être traitée mais celle-ci ne passe pas de 0 à 1. Dans l'absolu, pour ce qui est des régimes d'aides d'État fondés sur des critères d'admissibilité, la régression n'est précise que si l'on considère que le traitement est l'admissibilité. Si le traitement est la réception des aides, la régression est imprécise. Au contraire, lorsque l'allocation repose sur une notation, les seules entreprises prises en compte sont celles qui ont demandé une aide, et la régression est précise.

une zone très étroite autour de la discontinuité. Si la zone est trop large, l'incidence des autres caractéristiques ne pourra pas être considérée comme constante. En règle générale, le contrôle des paramètres observables, qui repose sur une forme fonctionnelle particulière, ne permet pas de résoudre ce problème.

L'examen graphique des données permet de conforter la fiabilité de l'hypothèse sous-jacente à la méthode de la régression par discontinuité. Pour ce faire, il importe en particulier de contrôler trois éléments, et en premier lieu la présence effective d'une discontinuité au seuil d'octroi des aides. Ensuite, il convient de s'assurer que les résultats à mesurer présentent une discontinuité au même moment mais qu'ils ne présentent pas d'autre discontinuité du même type à un autre moment. Enfin, il y a lieu de vérifier que les autres paramètres corrélés au résultat (notamment la propension à introduire une demande d'aide) ne présentent pas de discontinuité.

Pour finir, il est envisageable de créer sciemment des discontinuités afin de permettre l'évaluation d'un régime. Il est notamment possible d'accélérer une politique afin de créer des discontinuités et de faciliter l'identification de ses effets.

### **C. Estimation structurelle**

Dans certains cas, il est possible de franchir une étape supplémentaire et de confronter les données à un modèle théorique, par exemple celui des investissements consentis par les entreprises, afin d'obtenir les paramètres essentiels. Cette approche est qualitativement différente des précédentes. L'estimation structurelle applique un modèle théorique de comportement des entreprises défini dans sa totalité. Elle permet ainsi d'obtenir les paramètres essentiels qui déterminent le comportement d'une entreprise. L'évaluation s'effectue donc au plus près des déterminants du comportement individuel des entreprises et permet de simuler l'efficacité d'autres outils. Reste que l'estimation structurelle nécessite généralement davantage de ressources, de données et d'hypothèses.

S'il est impossible de fournir des orientations précises sur l'estimation structurelle dans la mesure où l'identification, l'estimation et l'inférence doivent être déduites au cas par cas, les indications générales données plus haut demeurent toutefois d'application. Tout d'abord, le modèle théorique doit refléter les principaux faits stylisés du marché. Ensuite, la question des caractéristiques non observées et de la sélection doit être expressément et correctement traitée.

### **D. Remarques méthodologiques complémentaires**

#### *Hétérogénéité des effets du traitement*

Les sections qui précèdent sont consacrées à l'estimation de l'effet moyen du traitement sur les populations traitées. Le terme même indique que l'effet des aides varie d'un bénéficiaire à l'autre. Cette hétérogénéité peut avoir des causes et des conséquences multiples. La première conséquence possible est que l'effet moyen risque d'être statistiquement insignifiant si les aides se révèlent extrêmement efficaces pour certaines entreprises et beaucoup moins pour d'autres. Cette absence d'effet statistiquement significatif ne veut pas dire pour autant que les



aides octroyées n'ont d'effet sur aucune entreprise. S'agissant des politiques publiques, la performance moyenne d'un régime constitue un premier indicateur très intéressant. Cependant, il est également important de s'attacher à comprendre les causes de cette hétérogénéité afin d'améliorer la conception de ces régimes et de cibler précisément les entreprises pour lesquelles les aides sont les plus efficaces et occasionnent le moins de distorsions.

Il est dès lors conseillé, dans la mesure du possible, d'estimer l'effet des aides sur différents types d'entreprises, par exemple sur les entreprises de petite taille par rapport aux entreprises de grande taille, sur les entreprises jeunes par rapport aux entreprises plus anciennes, sur les entreprises innovantes, les entreprises ayant un accès limité au crédit, etc.<sup>30</sup>.

#### *Distorsions à l'égard des entreprises non bénéficiaires*

L'évaluation – directe ou indirecte – de l'incidence du régime d'aides sur les non-participants fournit des informations précieuses pour l'évaluation des aides d'État. En effet, ces dernières peuvent créer des distorsions sur les marchés via les effets qu'elles produisent sur les non-bénéficiaires, soit en favorisant un transfert de connaissances en provenance des bénéficiaires soit en réduisant la compétitivité relative des non-bénéficiaires par rapport aux bénéficiaires.

De plus, les effets produits sur les entreprises ou les sites non bénéficiaires peuvent également influencer sur la validité de l'évaluation. Par exemple, les effets des aides à finalité régionale peuvent se matérialiser en partie par des mouvements à la frontière: des entreprises historiquement situées du «mauvais côté» d'une frontière peuvent ainsi être amenées à déménager leurs locaux de l'autre côté de cette frontière. Dès lors, une régression par discontinuité à la frontière risquerait de ne prendre en compte que cet effet de déplacement et de surestimer ainsi le réel effet agrégé de la politique. En pareil cas, il convient de recourir à une autre stratégie empirique (par exemple en vérifiant la robustesse de l'évaluation sur des régions plus vastes).

---

30 Une autre approche pourrait consister à procéder à l'estimation systématique des différents effets du traitement sur les entreprises, en différents points de la distribution conditionnelle. De plus en plus d'auteurs se penchent sur l'estimation de ces effets de traitement sur les quantiles, à commencer par Abadie, Angrist et Imbens (2002). Il existe un outil très utile pour comprendre la nature intrinsèquement hétérogène du traitement mais son utilité diminue sur le plan de la politique proprement dite, sauf à pouvoir cibler directement différentes entreprises en fonction de leur position au sein de la distribution conditionnelle.

## E. Données

Toute évaluation d'impact nécessite d'avoir accès à des données microéconomiques appropriées, caractérisées par une bonne cohérence entre les données se rapportant aux bénéficiaires et celles se rapportant aux non-bénéficiaires. Toutes les données doivent donc provenir de la même source, à l'exception bien sûr des informations relatives aux aides elles-mêmes. Les données doivent être le plus affinées possible, même si dans certains cas, il peut s'avérer nécessaire de procéder ultérieurement à certaines agrégations.

Il convient également de disposer de données rendant compte des indicateurs de résultat tant pour le groupe de traitement que pour le groupe de contrôle – et notamment de données indiquant la période de mesure du résultat – ainsi que du maximum de données possible sur les facteurs susceptibles d'influencer le résultat et la décision des entités de participer ou non au programme d'aides. Ces données, utilisées afin d'analyser les différences entre les groupes de traitement et les groupes de contrôle, indiqueront, par exemple, à l'échelle des entreprises, la taille, la localisation, les effectifs, les facteurs de production, etc.

L'origine la plus logique des données est administrative (bilans fiscaux, enquêtes nationales, etc.). Ces sources sont en effet riches d'informations sur la localisation et le domaine d'activité des entreprises, voire de leurs différents sites d'implantation. En principe, elles permettent de retracer l'investissement et le chiffre d'affaires par activité et de calculer les ratios financiers. Les grandes enquêtes nationales ou européennes, comme les enquêtes communautaires sur l'innovation, présentent aussi un intérêt, dans la mesure où elles couvrent un échantillon largement représentatif de par sa taille et fournissent des informations complémentaires très utiles sur des sujets précis. Les séries de données couplées employeurs-salariés constituent également une source d'informations privilégiée car elles permettent généralement de relier les caractéristiques de la main-d'œuvre à chaque site d'implantation, ce qui peut se révéler capital lorsque la dimension géographique de la main-d'œuvre intervient.

Outre les indicateurs de résultats et les caractéristiques des bénéficiaires, il est également important de disposer de données sur les aides et sur leur processus d'octroi. En règle générale, ces informations proviennent directement des autorités d'octroi des aides en la matière et font état à la fois du montant et de la date d'octroi des aides aux bénéficiaires. Cependant, il peut aussi être utile d'obtenir des informations plus générales sur le processus d'attribution des aides, ainsi que sur les demandes rejetées, en particulier si l'octroi des aides repose sur un mécanisme de notation<sup>31</sup>.

L'accès à ce type de données confidentielles étant généralement réglementé, il est impératif de veiller à ce que l'équipe scientifique chargée de l'évaluation puisse les consulter en temps opportun. Les délais d'accès aux sources administratives devront eux aussi être pris en compte dans la conception du plan d'évaluation.

---

31 Les données sur les demandes rejetées sont particulièrement précieuses pour les études fondées sur la méthode de la régression par discontinuité.

## F. Exemples

**Exemple 1 (aides à finalité régionale):** Criscuolo et al. (2012)<sup>32</sup> ont procédé à une évaluation du régime intitulé «Regional Selective Assistance (RSA)», un régime d'aides régionales sélectif instauré au Royaume-Uni entre 1986 et 2004. Durant cette période, des subventions à caractère discrétionnaire ont été attribuées au titre du RSA à des entreprises situées dans des zones défavorisées. Le RSA constituait alors le principal régime d'aides en faveur des entreprises au Royaume-Uni. Les aides octroyées au titre de ce régime relevaient du champ d'application des lignes directrices concernant les aides d'État à finalité régionale, et en particulier des cartes des régions admissibles (cartes des aides à finalité régionale), lesquelles ont évolué au fil du temps. En règle générale, le critère d'admissibilité retenu pour ce type d'aides est la position relative de la région en termes de PIB par habitant ou de chômage. Le statut d'une région peut dès lors évoluer, notamment en raison de son développement ou parce que le PIB moyen par habitant au sein de l'UE a évolué (par exemple à la suite de l'adhésion de nouveaux États membres, comme en 1995). Les indicateurs utilisés pour déterminer l'admissibilité des régions fluctuent eux aussi. Dans ces conditions, l'évolution de l'admissibilité des entreprises ne dépend pas uniquement de la situation des entreprises proprement dites mais aussi d'événements extérieurs au Royaume-Uni ou de modifications apportées aux règles administratives. En se concentrant sur cet aspect de l'évolution de l'admissibilité et en évaluant ses répercussions sur l'évolution des investissements, de l'emploi et de la productivité, Criscuolo et al. (2012) sont parvenus à identifier de manière très convaincante les effets du RSA<sup>33</sup>.

**Exemple 2 (soutien aux entreprises):** Martini et Bondonio (2012)<sup>34</sup> se sont penchés sur deux cas de soutien aux entreprises en Italie: une subvention à l'investissement proposée à l'échelle nationale (loi 488) d'une part, divers régimes d'aides en faveur des PME dans la région du Piémont d'autre part. La première évaluation est particulièrement intéressante dans la mesure où elle confronte les entreprises qui ont vu leur demande acceptée (autrement dit les bénéficiaires des aides) à des entreprises comparables dont la demande a été rejetée pour cause de tarissement du budget alloué à la mesure d'aide. L'inclusion des demandeurs écartés a permis d'éviter le biais de sélection qui se produit trop souvent lorsque l'évaluation se limite à comparer les demandeurs avec des non-demandeurs. Le groupe des demandeurs écartés a passé avec succès le premier contrôle de qualité, ce qui signifie qu'ils avaient des projets d'investissement crédibles. Aussi partageaient-ils avec les bénéficiaires la même ambition d'investir dans un projet viable. C'est en raison de contraintes budgétaires (rationnement) qu'ils n'ont pas pu bénéficier de la mesure d'aide. L'écart de performance entre les

32 Criscuolo, C., R. Martin, H. Overman et J. Van Reenen, 2012. «The causal effects of an industrial policy», CEPR Discussion Papers 8818, C.E.P.R. Discussion Papers.

33 Techniquement, Criscuolo et al. (2012) ont appliqué la méthode des variables instrumentales telle qu'elle est présentée dans cette annexe.

34 Rapport pour la DG REGIO, A. Martini, D. Bondonio: «Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy» (2012).

demandeurs (simplement) acceptés et les demandeurs rejetés (de justesse) a fourni une estimation fiable de l'effet de l'aide.

**Exemple 3 (garanties de prêts):** Lelarge, Sraer et Thesmar (2010) ont évalué les effets du programme de garantie de prêts proposé en France sous le nom de SOFARIS. Ce programme assure les prêteurs, au moyen de garanties, contre les risques de défaut des emprunteurs. Les emprunteurs versent une prime de risque, mais cette prime est subventionnée. Lelarge, Sraer et Thesmar (2010) ont décrit de manière très explicite la nature des effets de sélection. Premièrement, les entreprises qui présentent des projets plus rentables sont plus disposées à accepter de s'acquitter des frais liés à la garantie. Deuxièmement, il est probable que les gestionnaires du programme sélectionnent des projets socialement souhaitables, qui sans cela n'auraient pas accès à des financements privés. De manière générale, les entreprises s'autosélectionnent pour participer au programme, sachant qu'une sélection a également lieu lors de la phase d'octroi. Les résultats d'évaluations naïves, fondées par exemple sur des régressions linéaires classiques ou sur des comparaisons avec l'entreprise la plus comparable, peuvent s'en trouver affectés<sup>35</sup>. Le contexte factuel et institutionnel du programme constitue toutefois une source d'identification des effets de la politique. Le programme a été mis en œuvre à la fin des années 1980. Initialement circonscrit aux entreprises opérant dans les secteurs manufacturier et des services aux entreprises, il a vu sa dotation publique s'accroître en 1995 et a été étendu à de nouveaux secteurs d'activités (bâtiment, commerce de gros et de détail, transports, hôtellerie et restauration, services à la personne). Lelarge, Sraer et Thesmar (2010) ont comparé les entreprises nouvellement admissibles aux entreprises anciennement admissibles afin de mesurer les effets du programme sur un éventail d'indicateurs comme l'endettement, l'emploi, la croissance du capital, les dépenses financières et la probabilité de faire faillite. Bien que vraisemblablement différentes, les entreprises des deux groupes n'en subissent pas moins des chocs macroéconomiques analogues, de sorte que leurs différences n'évoluent a priori pas dans le temps, exception faite des effets escomptés du programme proprement dit<sup>36</sup>.

---

35 Dans ce cas, les techniques d'appariement, en l'occurrence l'appariement avec le plus proche voisin, n'offrent pas de meilleure solution aux problèmes de sélection que les moindres carrés ordinaires. Comme expliqué plus haut dans la présente annexe, les techniques d'appariement ne permettent généralement pas de résoudre la question des effets de sélection en l'absence d'expériences naturelles.

36 Concrètement, les auteurs ont appliqué un modèle de sélection par la méthode d'Heckman, avec une variable d'exclusion à l'échelle des entreprises et une approche classique fondée sur les variables instrumentales à l'échelle des secteurs d'activités. Pour en savoir plus sur ces méthodes, lire les pages précédentes de la présente annexe.

**Exemple 4 (crédits à la créativité):** Bakhshi et al.<sup>37</sup> ont eu recours à une expérience par essai aléatoire contrôlé afin d'évaluer les effets d'un régime de soutien aux entreprises innovantes. L'étude pilote, commencée à Manchester en 2009, s'articulait autour de «chèques» ou «crédits à la créativité», alloués de manière aléatoire à de petites ou moyennes entreprises désireuses d'investir dans des projets créatifs comme le développement de sites Internet, la production vidéo ou le lancement de campagnes de marketing créatif. L'objectif était de s'assurer de leurs effets sur l'innovation. En augmentant d'au moins 84 % la probabilité de voir des entreprises lancer des projets innovants avec des partenaires créatifs avec lesquels elles n'avaient jamais travaillé auparavant, ces crédits ont profondément modifié les rapports des PME avec les entreprises innovantes. Les chercheurs ont constaté que les entreprises qui bénéficiaient de crédits récompensant leur créativité connaissaient à court terme un fort regain d'innovation et qu'elles enregistraient une hausse de leur chiffre d'affaires dans les six mois suivant l'achèvement de leur projet. Mais ces effets positifs s'estompaient ensuite et il ne subsistait plus, au bout de 12 mois, de différence statistiquement significative entre le groupe des bénéficiaires de crédits à la créativité et le groupe des non-bénéficiaires. Les auteurs soutiennent que ces résultats n'auraient jamais pu être mis en évidence par les méthodes d'évaluation classiques utilisées par les pouvoirs publics et plaident pour un recours plus large aux essais aléatoires contrôlés dans l'évaluation des politiques de soutien aux entreprises.

**Exemple 5 (soutien à la RDI):** Einiö (2013) a étudié les effets des subventions à la RDI allouées par le Tekes, l'Agence finlandaise pour la technologie et l'innovation, sur les investissements en matière de recherche et développement, ainsi que sur l'emploi et la productivité, entre 2000 et 2006. L'étude s'est attachée à exploiter les variations régionales dans le budget du Tekes en matière de soutien à la RDI induites par l'augmentation de l'enveloppe financière allouée à certaines parties du nord et de l'est de la Finlande au titre du FEDER (zones relevant de l'objectif 1). Ces zones avaient été initialement déterminées lors des négociations d'adhésion de la Finlande en 1995 sur la base d'une densité de population inférieure ou égale à 8 personnes par kilomètre carré. Le budget de soutien à la RDI y étant relativement plus élevé, la probabilité de bénéficier de ce soutien était dès lors plus importante dans les zones relevant de l'objectif 1 que dans d'autres régions du pays. Il en a résulté une variation régionale dans le traitement, caractérisée par un nombre sensiblement supérieur d'entreprises soutenues dans la région relevant de l'objectif 1. L'allocation régionale étant fondée sur la règle de la densité de population telle qu'établie en 1993 (et non, par exemple, sur les niveaux attendus d'investissement en RDI ou les performances économiques des entreprises), l'étude s'est penchée sur la densité de population de 1993, ce qui a permis de traiter effectivement la question de la sélection régionale. Concrètement, les effets de traitement ont été estimés selon la méthode des variables instrumentales, sachant que tout indicateur relatif aux régions relevant de l'objectif 1 faisait également figure d'instrument

---

37 Rapport commandé par le Nesta, «Creative Credits, a randomised controlled industrial policy experiment», Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley et N. Rathbone, juin 2013, disponible sur [http://www.nesta.org.uk/sites/default/files/creative\\_credits.pdf](http://www.nesta.org.uk/sites/default/files/creative_credits.pdf).

d'accès au programme. Cette approche visait à identifier l'incidence du soutien sur les entreprises ayant participé au régime d'aides du fait de l'augmentation de l'enveloppe financière allouée aux zones relevant de l'objectif 1. La validité de la construction a été confirmée par le fait que les entreprises appelées à bénéficier du régime et celles composant le groupe de contrôle affichaient, avant le programme, des tendances similaires. Einiö (2013) a mis en évidence une incidence positive sur les investissements en matière de R&D, l'emploi et le chiffre d'affaires des participants ayant bénéficié d'une aide à la R&D grâce au financement agrégé supplémentaire alloué à leur région. Malgré l'absence d'incidence instantanée sur la productivité, l'étude a également fait apparaître des gains de productivité à long terme.

## Annexe II: Liste d'indicateurs de résultats envisageables

La liste proposée ci-dessous n'est donnée qu'à titre d'exemple. Les indicateurs de résultats devront être établis en fonction des objectifs du régime d'aides et de l'évaluation.

### Incidence directe des aides au niveau des bénéficiaires

	DIMENSION DU RÉSULTAT	INDICATEURS DE RÉSULTATS
<b>Aides à finalité régionale</b>	Effets positifs	Investissements privés complétant un soutien public  Augmentation de l'emploi dans les entreprises bénéficiant d'un soutien
<b>Aides à la recherche, au développement et à l'innovation (RDI)</b>	Activités supplémentaires dans le domaine de la RDI	Investissements privés complétant un soutien public  Dépenses supplémentaires consenties dans ce domaine par les entreprises bénéficiant d'un soutien  Nombre de chercheurs embauchés dans les entreprises bénéficiant d'un soutien  Nombre de nouveaux brevets déposés  Nombre d'entreprises bénéficiant d'un soutien pour lancer des produits nouveaux sur le marché
<b>Aides à la protection de l'environnement</b>	Effets positifs sur l'environnement	Réduction des émissions de CO2 des entreprises bénéficiaires  Capacités supplémentaires de production d'énergies renouvelables  Réduction du nombre de déchets enfouis ou incinérés  Nombre de sites contaminés assainis
	Adoption anticipée de normes environnementales	Pourcentage d'entreprises satisfaisant aux nouvelles normes environnementales au moins X mois/années avant leur entrée en vigueur [l'exigence a été fixée à 12 mois minimum, et des intensités d'aide plus élevées ont été autorisées en cas d'adoption antérieure à 3 ans]

<b>Aides à l'énergie (et en faveur des infrastructures énergétiques)</b>	Diminution de la consommation d'énergie	<p>Nombre de ménages dont le classement en matière de consommation énergétique s'est amélioré</p> <p>Diminution de la consommation annuelle d'énergie primaire des bâtiments publics</p> <p>Nombre d'utilisateurs d'énergie supplémentaires connectés à des réseaux intelligents</p>
	Soutien aux énergies renouvelables	Part de la production d'énergie tirée de SER
<b>Financement des risques</b>	Effets positifs	<p>Rendements obtenus au sein du fonds</p> <p>Mobilisation des investissements privés</p> <p>Nombre d'entreprises bénéficiant de capital-risque</p>
	Perdants	Mauvaise performance moyenne des entreprises bénéficiaires, en raison d'une piètre gestion commerciale/d'une participation privée insuffisante
	Degré de diversification insuffisant	Fonds trop petits ou trop limités sur le plan régional, assortis de perspectives de rendement réduites, présentant peu d'attrait aux yeux des investisseurs privés
<b>Aides en faveur du haut débit</b>	Augmentation de la couverture du haut débit	<p>Ménages supplémentaires bénéficiant d'un accès à haut débit d'au moins 30 Mbps</p> <p>Ménages supplémentaires bénéficiant d'une implantation ou d'un accès à haut débit d'au moins 100 Mbps</p>
	Efficiences	<p>Coûts d'investissement/aide pour la connexion d'un ménage (foyers connectés)</p> <p>Nombre de ménages souscrivant de nouveaux services</p>
<b>Sauvetage et restructuration</b>	Effets positifs	<p>Maintien de l'emploi et de l'activité à l'échelle des entreprises et au niveau régional</p> <p>Évolution des parts de marché et de la productivité des entreprises aidées</p>



<b>Transport aérien</b>	Effets positifs	<p>Nombre de transporteurs aériens utilisant l'aéroport</p> <p>Investissements privés complétant un soutien public</p> <p>Accroissement de la productivité régionale et/ou de la valeur ajoutée brute (VAB)</p>
	Effets négatifs	<p>Duplication des infrastructures ou des liaisons aériennes déficitaires</p> <p>Détérioration du trafic dans les infrastructures existantes (présence d'autres aéroports ou d'autres moyens de transport dans la zone desservie, etc.)</p>

## Incidence indirecte du régime d'aides

	DIMENSION DU RÉSULTAT	INDICATEURS DE RÉSULTATS
Effets positifs potentiels	Avantages macroéconomiques	Augmentation de l'emploi Accroissement de la productivité et/ou de la valeur ajoutée brute (VAB)
	Diversification de l'économie régionale	Nombre de secteurs répertoriés sous les différents codes NACE
	Coopération renforcée entre le secteur public et le secteur privé	Nombre d'entreprises coopérant avec des organismes de recherche
	Externalité positive / effets de contagion	Nombre de bénéficiaires indirects (par ex. nombre de tiers ayant accès à l'infrastructure) Évolution de l'emploi et de l'activité dans d'autres entreprises ou d'autres régions (transport aérien) Nombre d'habitants bénéficiant de meilleurs moyens de transport dans la zone desservie
Effets négatifs potentiels sur la concurrence et les échanges	Biais sectoriel	Aide majoritairement orientée vers une branche au sein d'un régime plurisectoriel
	Tendance à favoriser les entreprises déficitaires ou à faible productivité (pour prévenir toute sortie)	Proportion d'entreprises à forte productivité par rapport aux entreprises à faible productivité
	Tendance à favoriser les opérateurs historiques	Proportion d'entreprises anciennes par rapport aux entreprises plus jeunes
	Renforcement d'un pouvoir de marché	Évolution du pouvoir de marché détenu par un acteur dominant
	Effet lié au site d'implantation	Déplacement d'une région pauvre vers une région plus développée
	Sécurité de l'approvisionnement	Enfermement dans des sources d'énergie à haute teneur en carbone Évaluation de l'existence ou de la persistance de risques de coupures de courant Verrouillage des marchés nationaux de l'électricité

	Infrastructure énergétique	<p>Verrouillage des marchés nationaux de l'électricité</p> <p>Renforcement du pouvoir de marché d'un opérateur historique</p>
	Sauvetage et restructuration	<p>Évolution de l'emploi et de l'activité dans d'autres entreprises ou d'autres régions</p> <p>Évolution des parts de marché et de la productivité des entreprises aidées</p>
	Transport aérien	<p>Duplication des infrastructures ou des liaisons aériennes déficitaires</p> <p>Détérioration du trafic dans les infrastructures existantes (présence d'autres aéroports ou d'autres moyens de transport dans la zone desservie, etc.)</p>

### Annexe III: Glossaire

<b>Valeur de référence</b>	Valeur de l'indicateur avant la mise en œuvre de l'intervention concernée.
<b>Groupe de contrôle</b>	Toute analyse contrefactuelle suppose de trouver la(les) entreprise(s) ou le groupe de contrôle se prêtant le mieux à la comparaison, c'est-à-dire un groupe d'entreprises affichant la similarité la plus grande possible avec le groupe d'entreprises bénéficiaires – hormis le fait qu'elles n'ont pas reçu l'aide en question.
<b>Scénario contrefactuel</b>	Afin d'évaluer l'effet de l'aide sur ses bénéficiaires, il est nécessaire de construire un «scénario contrefactuel», autrement dit d'échafauder un scénario plausible rendant compte de ce qui serait arrivé aux bénéficiaires si ces derniers n'avaient pas reçu l'aide en question.
<b>Évaluation</b>	Collecte et analyse systématiques d'informations sur les programmes et les projets, leurs objectifs et leur mise en œuvre. L'évaluation en retire des connaissances sur leur incidence afin de pouvoir émettre des jugements. Les évaluations sont utilisées pour améliorer l'efficacité et permettre des prises de décision éclairées quant à la programmation en cours et future.
<b>Incidence</b>	Changement pouvant être imputé de manière crédible à une intervention. Synonyme d'«effet» d'une intervention ou de sa «contribution au changement».
<b>Indicateur</b>	Variable fournissant des informations quantitatives ou qualitatives sur un phénomène donné. Elle inclut normalement une valeur et une unité de mesure.
<b>Méthode</b>	Les méthodes sont des familles de techniques et d'outils d'évaluation remplissant des objectifs distincts. Elles consistent généralement en des procédures et des protocoles assurant la systématisation et la cohérence des modalités de réalisation des évaluations. Les méthodes peuvent être centrées sur la collecte ou l'analyse d'informations et de données, être quantitatives ou qualitatives, viser à décrire, expliquer, prédire ou influencer les mesures à mettre en place. Le choix des méthodes dépend de la nature de l'intervention, des questions d'évaluation posées et du mode d'enquête (causal, exploratoire, normatif, etc.).
<b>Résultat</b>	Dimension spécifique du bien-être des personnes motivant l'adoption de mesures, c'est-à-dire dimension appelée à être modifiée par les interventions conçues et mises en œuvre par une politique. Exemples: mobilité dans une zone donnée, compétences dans un secteur d'activité particulier.
<b>Indicateur de résultat</b>	Indicateur décrivant un aspect spécifique d'un résultat, une caractéristique pouvant être mesurée. Exemples: temps requis pour se déplacer du point W au point Y à une vitesse moyenne (aspect de la mobilité), résultats d'examens menés dans un domaine donné (aspect de compétence), part des entreprises qui se sont vu refuser un crédit à un certain taux d'intérêt (aspect du rationnement des financements par les banques).

## Annexe IV: Références

Abadie, A., A. Diamond et J. Hainmueller (2007), «Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Programme», *Journal of the American Statistical Association*, juin 2010, vol. 105, n° 490.

Abadie, A., J. Angrist et G. W. Imbens (2002), «Instrumental Variables Estimates of the Effect of Subsidised Training on the Quantiles of Trainee Earnings», *Econometrica*, 70(1), 91–117.

Angrist, J. D., et J. Pischke (2010), «The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics», *Journal of Economic Perspectives*, 24(2): 3-30.

Angrist, J. et A. Krueger (1991), «Does Compulsory School Attendance Affect Schooling and Earnings», *Quarterly Journal of Economics*, 106.

Angrist, J. et J. Pischke (2008), «Mostly Harmless Econometrics: An Empiricist's Companion», Princeton University Press.

Bakhshi, H., J. Edwards, S. Roper, J. Scully, D. Shaw, L. Morley et N. Rathbone (2013), «Creative credits, a randomized controlled industrial policy experiment», rapport commandé par le Nesta, disponible sur [http://www.nesta.org.uk/sites/default/files/creative\\_credits.pdf](http://www.nesta.org.uk/sites/default/files/creative_credits.pdf).

Banque mondiale (2003), «Independent Evaluation: Principles, Guidelines and Good Practice». Disponible sur:  
<http://siteresources.worldbank.org/INTDGF/Resources/Evaluation&LearningNote.pdf>.

Bertrand, M., E. Duflo et S. Mullainathan (2004), «How much should we trust differences-in-differences estimates ? », *The Quarterly Journal of Economics*, 119, 249–275.

Bound, J., D. Jaeger et R. Baker (1995), «Problems with Instrumental Variable Estimation When the Correlation Between the Instruments and the Endogenous is weak», *Journal of the American Statistical Association*, 90(430), 443–450.

C. Lelarge, D. Sraer et D. Thesmar (2010), 'Entrepreneurship and Credit Constraints: Evidence from a French Loan Guarantee Programme,' NBER Chapters, in: *International Differences in Entrepreneurship*, pages 243-273, National Bureau of Economic Research, Inc.

Criscuolo, C., R. Martin, H. Overman et J. Van Reenen (2012), «The causal effects of an industrial policy», CEPR Discussion Papers 8818, C.E.P.R. Discussion Papers.

Duflo, E., et M. Kremer (2005), «Use of Randomisation in the Evaluation of Development Effectiveness», in *Evaluating Development Effectiveness*, éd. par O. Feinstein, G. K. Ingram,

et G. K. Pitman. New Brunswick, New Jersey and London, U.K.: Transaction Publishers, vol. 7, pp. 205{232}.

Duflo, E., R. Glennerster & M. Kremer (2007), «Using Randomisation in Development Economics Research: A Toolkit», CEPR Discussion Papers 6059, C.E.P.R. Discussion Papers.

Einiö, Elias (2013), «R&D Subsidies and Company Performance: Evidence from Geographic Variation in Government Funding Based on the ERDF Population-Density Rule», *The Review of Economics and Statistics* (à venir).

Garicano, L. C. Lelarge et J. Van Reenen, (2012), «Firm Size Distortions and the Productivity Distribution: Evidence from France», CEP Discussion Papers dp1128, Centre for Economic Performance, LSE.

Givord, P. (2010), «Méthodes économétriques pour l'évaluation de politiques publiques», WPD3E n° G2010-08.

Givord, P., Rathelot, R. et P. Sillard (2013), «Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines programme», *Regional Science and Urban Economics*, volume 43, numéro 1, janvier 2013, pages 151-163

Groupe des Nations Unies pour l'évaluation (2005), «Règles d'évaluation applicables au sein du système des Nations Unies», disponibles sur: [http://www.uneval.org/papersandpubs/documentdetail.jsp?doc\\_id=22](http://www.uneval.org/papersandpubs/documentdetail.jsp?doc_id=22).

Heckman, J. J. (1979), «Sample Selection Bias as a Specification Error», *Econometrica* 47, 153–161.

Imbens, G. et J. Wooldridge (2009), *Recent Developments in the Econometrics of Programme Evaluation*, *Journal of Economic Literature*, 47:1, 5-86

Imbens, G. W., et T. Lemieux (2008), «Regression discontinuity designs: A guide to practice», *Journal of Econometrics*, 142(2), 615–635.

Keane, M. P. (2010), «A Structural Perspective on the Experimentalist School», *Journal of Economic Perspectives*, 24(2): 47-58.

Martini, A. et D. Bondonio (2012), «Counterfactual impact evaluation of cohesion policy: impact and cost effectiveness of investment subsidies in Italy», rapport remis à la Commission européenne, DG Regio.

Nederlandse Rijksoverheid (2012), «Durf te meten», Eindrapport Expertwerkgroep Effectmeting, disponible sur <http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/2012/11/23/durf-te-meten-eindrapport-expertwerkgroep-effectmeting.html>

Nevo, A. et M. D. Whinston, (2010), «Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference», Journal of Economic Perspectives, 24(2): 69-82.

Normes et standards d'évaluation de l'OCDE. Disponibles sur:  
<http://www.oecd.org/fr/cad/evaluation/49898837.pdf>

Sims, C. A. (2010), «But Economics Is Not an Experimental Science», Journal of Economic Perspectives, 24(2): 59-68.

Standards d'évaluation de la Commission européenne. Disponibles sur:  
[http://ec.europa.eu/dgs/secretariat\\_general/evaluation/docs/standards\\_c\\_2002\\_5267\\_final\\_fr.pdf](http://ec.europa.eu/dgs/secretariat_general/evaluation/docs/standards_c_2002_5267_final_fr.pdf).

Stock, J., J. Wright et M. Yogo (2002): «A Survey of Weak Instruments and Weak Identification in Generalised Method of Moments», Journal of Business and Economic Statistics, 20(4), 518–29.

Wooldridge, J. (2002), «Econometric Analysis of Cross Section and Panel Data», Cambridge : MIT Press