

Non classifié

DSTI/IND/PME(2003)4/FINAL



Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

28-May-2004

Français - Or. Anglais

**DIRECTION DE LA SCIENCE, DE LA TECHNOLOGIE ET DE L'INDUSTRIE
COMITE DE L'INDUSTRIE ET DE L'ENVIRONNEMENT DE L'ENTREPRISE**

**DSTI/IND/PME(2003)4/FINAL
Non classifié**

Groupe de travail sur les petites et moyennes entreprises et l'entrepreneuriat

EVALUATION DES POLITIQUES ET PROGRAMMES A L'EGARD DES PME

Processus de Bologne de l'OCDE

2e Conférence ministérielle de l'OCDE sur les PME, sur le thème "Promouvoir l'entrepreneuriat et les PME innovantes dans une économie mondialisée -- Vers une mondialisation plus responsable et fédératrice", organisée conjointement par l'OCDE et le ministère turc de l'Industrie et du Commerce

Cette Conférence aura lieu à ISTANBUL, Turquie, du 3 au 5 juin 2004.

JT00165117

**Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format**

Français - Or. Anglais

AVANT PROPOS

Lors de la première Conférence de l'OCDE des Ministres en charge des PME accueillie par le gouvernement italien à Bologne (Italie) en juin 2000, les ministres de près de 50 pays Membres et non membres de l'Organisation ont adopté la Charte de Bologne sur les politiques à l'égard des PME. Pour eux, la Conférence de Bologne a marqué le début d'un dialogue sur les politiques impliquant à la fois des pays Membres de l'OCDE et des économies non membres, et devait donner lieu à un suivi continu dans le cadre de la mise en œuvre de la Charte de Bologne. Ce dialogue et ce suivi sont aujourd'hui désignés par l'appellation « Processus de Bologne de l'OCDE ». La deuxième Conférence de l'OCDE des Ministres en charge des PME accueillie par le ministère turc de l'Industrie et du Commerce (Istanbul, du 3 au 5 juin 2004), est une occasion d'évaluer l'impact sur les PME de la nouvelle donne que crée la mondialisation.

Le présent document est l'un des dix rapports de référence établis pour la Conférence ministérielle d'Istanbul, le thème de chacun étant lié à un atelier spécifique de cette Conférence. Plusieurs versions précédentes de ce rapport ont été examinées par le Groupe de travail sur les petites et moyennes entreprises et l'entrepreneuriat de l'OCDE dont les observations ont été intégrées dans la version finale. Les économies non membres participant au Processus de Bologne de l'OCDE ont également eu l'occasion de formuler des commentaires. La version finale du rapport énonce les principaux messages et les principales recommandations d'action résultant des travaux préparatoires menés depuis la Conférence de Bologne au sein du Groupe de travail sur les PME et l'entrepreneuriat de l'OCDE. En raison du caractère très disparate du développement économique, des dispositifs institutionnels et du contexte politique que l'on rencontre dans les économies participant au Processus de Bologne – désormais au nombre de 80 – les composantes de telle ou telle politique ou de tel ou tel programme ne sont pas toujours adaptées à l'ensemble des protagonistes. Néanmoins, dans leurs grandes lignes, les messages et recommandations d'action ci-inclus représentent un ensemble qui traduit dans les faits les thèmes de la Charte de Bologne, et dont les responsables politiques sont invités à s'inspirer pour œuvrer en faveur des PME innovantes dans une économie mondiale. Les ministres examineront cet ensemble ainsi que d'autres recommandations lors de leurs délibérations à la Conférence d'Istanbul.

Le présent rapport a été établi par le professeur David Storey, Directeur du Small Business Centre, University of Warwick, Royaume-Uni, en étroite collaboration avec le Secrétariat de l'OCDE (Unité PME).

Le présent rapport est diffusé sous la responsabilité du Secrétaire général de l'OCDE. Les vues exprimées sont celles de ses auteurs et ne reflètent pas nécessairement celles de l'Organisation ou de ses pays Membres.

Les rapports de référence préparés pour la 2^{ème} Conférence Ministérielle de l'OCDE sur les PME sont disponibles in extenso sur les sites Internet suivants :

Site Internet de l'unité PME : <http://www.oecd.org/sti/smes>

Site Internet de la Conférence : <http://www.oecd-istanbul.sme2004.org>

Copyright OCDE, 2004.

Toute demande d'autorisation de reproduction ou de traduction totale ou partielle de cette publication doit être adressée : Chef des services des publications, OCDE, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

TABLE DES MATIÈRES

AVANT PROPOS	2
NOTE DE SYNTHÈSE	4
I. INTRODUCTION	7
CONTEXTE	9
II. ÉVALUATION DES POLITIQUES A L'ÉGARD DES PME	11
Le contexte de l'évaluation	11
Le cadre d'évaluation	17
La méthode d'évaluation en six étapes.....	18
La méthode en six étapes : modifications et inconvénients	25
Exemples d'évaluation relevant des bonnes pratiques	29
Enseignements dégagés.....	34
III. CONCLUSIONS.....	39
ANNEXE 1	41

Évaluation des politiques et programmes à l'égard des PME

NOTE DE SYNTHÈSE

Ce document présente un examen approfondi des méthodes employées pour évaluer les composantes de la politique à l'égard des PME. Il s'inspire essentiellement de l'expérience des pays Membres de l'OCDE mais, le cas échéant, fait aussi référence à des économies non membres. Les conclusions principales en sont présentées et résumées dans l'encadré ci-après.

En premier lieu, cette évaluation ne peut être conduite correctement avant que les objectifs et cibles de la politique à l'égard des PME ne soient clairement spécifiés. Le document propose d'utiliser le cadre COCE pour formuler les politiques. La première lettre « C » y représente à la fois la clarté et la cohérence de l'action publique. « Clarté », dans ce contexte, signifie que la politique en question est pleinement comprise par ceux qui l'appliquent et par ceux qui sont censés en bénéficier. Le concept de « cohérence », quant à lui, est un peu plus complexe. Il a trait au besoin de veiller à ce que tous les secteurs de l'Administration s'emploient à travailler collectivement en faveur des PME. Plus précisément, les politiques à l'égard des PME ont peu de chance d'être « cohérentes » si elles sont appliquées, sans coordination, par de multiples administrations ou si quelques administrations travaillent à la réglementation des PME et d'autres à « l'assistance », sans mettre en place un dialogue approprié.

La lettre « O » signifie « Objectifs ». Pour garantir la clarté des politiques, il convient d'énoncer précisément leurs objectifs ; concrètement, ceux-ci peuvent consister à augmenter le nombre d'entreprises, le taux de création des entreprises ou leur valeur ajoutée. Il peut aussi s'agir de chercher à accroître le nombre d'entrepreneurs parmi des groupes définis en termes d'appartenance ethnique, de sexe, de désavantage ou de localisation.

La deuxième lettre « C » correspond à « cibles ». Celles-ci doivent être spécifiées de manière à ce qu'une évaluation ultérieure puisse être réalisée pour définir si elles sont atteintes ou pas. A titre d'exemple de transformation d'un « objectif » en « cible », prenons l'objectif consistant à « augmenter le taux de création des entreprises ». Dans ce cas, la cible pourrait être « d'augmenter la création d'entreprises de 10 % dans les trois prochaines années ». Ce n'est qu'ainsi que l'on pourra définir si la cible est atteinte ou pas.

La lettre « E » signifie « Évaluation », qui est le thème central du présent document. Celui-ci conclut au premier chef qu'il est impossible de conduire une évaluation en l'absence de cibles, car une évaluation ne peut s'effectuer que dans un cadre où les effets escomptés des politiques sont clairement spécifiés.

Le document conclut en deuxième lieu que l'évaluation doit être au cœur du processus d'élaboration des politiques. Elle ne doit pas être engagée uniquement à titre d'exercice comptable rétrospectif pour déterminer si l'argent public a été judicieusement dépensé, bien que cette fonction soit utile. L'évaluation ne doit pas être « gardée pour la fin », mais éclairer les politiques en vigueur de sorte que les objectifs et cibles puissent être modifiés en fonction de l'efficacité des mesures. De ce fait, les réflexions sur les méthodes d'évaluation devraient être intégrées à la formulation des politiques au stade de l'élaboration des nouvelles idées. Elles pourraient même influencer les choix opérés par les pouvoirs publics sur la meilleure façon d'apporter un soutien aux PME.

Veiller à ce que l'évaluation reçoive une priorité plus élevée signifie que tous les programmes, et pas quelques-uns seulement, sont évalués. Le document ne précise toutefois pas si tous les programmes doivent être évalués de manière aussi approfondie. L'argument fondamental en faveur d'un degré uniforme d'évaluation est qu'il permet d'établir des comparaisons valables entre les différents programmes. L'argument inverse avance qu'il est déraisonnable de prévoir des budgets d'évaluation similaires pour des programmes d'envergures très diverses.

Quoi qu'il en soit, le document souligne qu'il est désormais possible, sur le plan technique, de mesurer précisément l'impact des programmes. Il montre, dans le cadre de son approche « Six étapes vers une méthode d'évaluation idéale » que, si la méthode la plus perfectionnée est parfois appliquée, c'est là l'exception plutôt que la règle. Le problème vient de ce que les méthodes moins approfondies apportent généralement des réponses fallacieuses, le plus souvent parce qu'elles surestiment l'effet des politiques. La troisième conclusion principale du document est que, lorsque cela se justifie, il convient d'employer les méthodes les plus perfectionnées (étapes 5 ou 6). La principale imperfection de cette « approche en six étapes », c'est-à-dire le fait qu'elle soit, à juste titre, utilisée pour évaluer l'impact de programmes et de politiques particuliers, plutôt que celui de la politique globale à l'égard des PME, est également pris en considération.

Principales recommandations

- **Inscrire la politique visant les PME dans le cadre analytique « COTE » (initiales en gras dans l'énumération ci-après) :**
 - La démonstration du bien-fondé de l'intervention publique et de l'intention est un préalable à des politiques visant les PME claires et homogènes. Les différentes instances publiques qui interagissent avec les PME pour faciliter leur développement doivent s'assurer que leurs efforts sont homogènes et coordonnés.
 - Les objectifs des politiques visant les PME doivent être clairement stipulés. On peut citer par exemple la création de nouvelles entreprises, la croissance des entreprises existantes ou la promotion de l'esprit d'entreprise auprès de catégories ciblées de la population.
 - Les cibles (t de « *targets* ») doivent être spécifiées de manière mesurable de façon à faciliter l'évaluation de la réalisation des objectifs.
 - L'évaluation des politiques, qui doit reposer sur des objectifs clairs, doit être le test le plus important de son efficacité.
- **Développer une « culture de l'évaluation » en plaçant l'évaluation des programmes au centre du processus d'action.** L'évaluation ne doit pas simplement être entreprise à titre d'exercice de comptabilité a posteriori, mais servir à informer la politique en vigueur, dont les objectifs des cibles peuvent être modifiés à la lumière des résultats de l'évaluation.
- **Intégrer la méthodologie et le budget d'évaluation des programmes dans le processus législatif.** Ne pas reporter ces points après la mise en œuvre du programme.
- **Garantir que tous les programmes, et non pas seulement quelques-uns, sont soumis à évaluation.**
- **Les évaluations principales doivent être confiées à des organes indépendants mais informés.** Néanmoins, l'implication des équipes des programmes, qu'il s'agisse des décideurs ou des prestataires, est essentielle.
- **Évaluer tous les programmes à l'aide des techniques les plus élaborées tout en prenant en compte la nécessité d'une rentabilité économique.** En pratique, des démarches plus approximatives s'imposeront pour l'évaluation de programmes de moindre ampleur.

Au moment d'engager ces évaluations, il convient de ne pas perdre de vue les éventuelles incompatibilités. D'abord, dans l'idéal, ceux qui conduisent les évaluations doivent être indépendants des responsables des programmes, comme dans tout audit. Mais si l'évaluation a pour fonction essentielle de concourir au perfectionnement constant de la politique, les évaluateurs ont besoin de la collaboration et de la participation actives de ceux qui l'ont élaborée et de ceux qui l'appliquent. Cela risque d'être plus difficile si l'évaluateur est vu comme un intrus. L'un dans l'autre, le document conclut, mais de justesse, que l'indépendance de l'évaluation revêt une importance cruciale.

Enfin, le document conclut que l'OCDE elle-même pourrait jouer un plus grand rôle dans la diffusion des « bonnes pratiques » en matière de techniques d'évaluation des politiques à l'égard des PME.

Évaluation des politiques et programmes à l'égard des PME

I. INTRODUCTION

Le thème « transversal » du présent document est l'évaluation des politiques et programmes en faveur des PME. Les Ministres sont priés d'engager une réflexion sur le mode et le type d'évaluation des politiques à l'égard des PME et de l'entrepreneuriat qui sont censés constituer une pratique optimale. Il est également associé à un deuxième thème – assurer la cohérence des politiques à l'égard des PME. Cette cohérence est obtenue quand toutes les mesures publiques, vues dans leur globalité, atteignent les objectifs généraux de la politique définie. Les diverses composantes de l'action publique doivent se renforcer mutuellement et non se contredire. Un test déterminant consiste à vérifier s'il est possible de réallouer les fonds de manière à apporter des avantages supplémentaires/supérieurs à l'économie. Si cette réallocation n'est pas possible, la politique pourra être jugée cohérente [OCDE 2003]¹.

Dans la pratique, l'importance que les responsables publics et les professionnels attachent à l'évaluation de l'action gouvernementale varie selon les pays, voire selon les programmes à l'intérieur d'un même pays. Cet intérêt transparait dans le degré de perfectionnement des techniques d'évaluation employées. Il est de ce fait difficile, lorsque les méthodes utilisées sont différentes, de comparer l'impact de mesures analogues dans différents pays. L'OCDE a donc un rôle essentiel à jouer, qui consiste à harmoniser les procédures d'évaluation et de permettre ainsi de comparer plus aisément l'efficacité des politiques. Il existe néanmoins des dénominateurs communs à tous les pays, à savoir la nécessité d'assigner des objectifs précis aux programmes, de prévoir des budgets réservés aux évaluations, d'aborder ces évaluations sous un angle analytique et de veiller à ce qu'elles soient de grande qualité sur le plan technique.

Le document aboutit aux conclusions essentielles suivantes :

- Les politiques et programmes doivent être assortis d'objectifs et de cibles qu'il est en principe possible de mesurer.
- La méthode d'évaluation et le budget qui lui est destiné doivent être définis dans le cadre du processus législatif et non pas reportés jusqu'à la réalisation du programme.
- Tous les programmes, et pas seulement certains d'entre eux, doivent être évalués.
- Les programmes de grande envergure doivent faire appel aux méthodes d'évaluation les plus perfectionnées qui soient, et la technicité de l'évaluation diminuer proportionnellement à l'ampleur du programme.

¹

Dans leur examen sur la cohérence des politiques dans le cadre du développement mondial, OCDE (2003), les auteurs disent que « il ne sert à rien que les pays donateurs consentent une aide au développement pour améliorer les capacités d'un pays dans le domaine des échanges s'ils maintiennent par ailleurs des obstacles commerciaux qui excluent du marché les produits en provenance des pays en développement. »

Le document soulève par ailleurs les questions suivantes :

- Qui doit être chargé de réaliser les évaluations ? Faut-il confier cette tâche aux responsables de l'exécution du programme, à des fonctionnaires indépendants ou seulement à des organismes extérieurs ?
- L'OCDE pourrait-elle faire fonction de banque d'expertise en évaluation, à laquelle les États Membres et non membres pourraient faire appel selon leurs besoins ?

CONTEXTE

L'OCDE a déjà entrepris des travaux de grande envergure dans le domaine de l'évaluation des politiques. Dans une publication parue en 1997, elle énonçait cinq recommandations en matière d'évaluation :

- La méthode d'évaluation doit être conçue en même temps que la politique ou le programme, de sorte que l'on puisse obtenir les données nécessaires et s'assurer de l'acceptation des critères et procédures d'évaluation par les parties concernées.
- Il faut veiller à ce que les conclusions de l'évaluation soient adoptées au plus haut niveau décisionnel en encourageant le débat public à leur propos et en faisant en sorte que les responsables publics y réagissent.
- L'évaluation doit être orientée vers les utilisateurs, répondre aux besoins d'information des différents clients et participants au programme.
- Les évaluations doivent faire appel à une combinaison de méthodes pour répondre aux différents besoins d'information.
- Les évaluations sont encore souvent peu concluantes en raison d'une estimation inadéquate de l'impact des programmes.

A l'origine, ces recommandations ont été formulées afin de résoudre les problèmes rencontrés dans le cadre de l'évaluation de l'action gouvernementale dans les domaines de l'innovation et de la technologie. Le Groupe de travail sur les PME [OCDE (2000)] a élargi ce cadre aux aspects intéressant les PME et l'entrepreneuriat. Il a dégagé les sept critères suivants pour apprécier les mesures prises dans ce secteur :

- Justification
- Additionnalité
- Pertinence
- Supériorité
- Efficience systémique
- Efficience propre
- Efficience d'adaptation

Ces critères se définissent comme suit :

- *La justification* : le bien-fondé du programme ;
- *L'additionnalité* : la valeur ajoutée nette du programme ;
- *La pertinence* : la mesure dans laquelle le programme s'attaque à une défaillance clairement identifiée du marché ;
- *La supériorité* : l'efficacité du programme par rapport à d'autres programmes qui permettraient d'atteindre les mêmes objectifs ;
- *L'efficacité systémique* : la mesure dans laquelle le programme entretient des interactions positives ou négatives avec d'autres actions gouvernementales ;
- *L'efficacité propre* : la rentabilité du programme dans la réalisation de ses objectifs spécifiques ;
- *L'efficacité d'adaptation* : la mesure dans laquelle les évaluations entraînent une réorientation du programme.

La taxinomie présentée ci-dessus montre bien que l'évaluation des programmes ne doit pas être considérée comme la dernière étape du processus. Elle a plutôt pour objectif d'apporter des aménagements et des modifications aux programmes, qui sont ensuite corrigés à la lumière de l'expérience. Ces modifications peuvent se traduire par un changement de cibles et d'objectifs, par une orientation sur d'autres groupes de clients, par l'application de méthodes de mise en œuvre différentes ou par la mobilisation de moyens à d'autres niveaux. Il s'agit avant tout d'encourager les responsables politiques et les professionnels à ne pas voir dans les évaluations un moyen d'imputer les responsabilités et de mettre un terme aux programmes inefficaces - bien que cela puisse parfois s'avérer nécessaire. L'évaluation devrait au contraire avoir pour finalité première d'améliorer la prestation des services.

Pour qu'une évaluation soit probante, il est impératif de définir au préalable des finalités et des objectifs précis pour les politiques. Les objectifs étant souvent énoncés en termes « visionnaires » pour obtenir un soutien politique, ils sont souvent obscurs et appellent à une interprétation de la part de l'évaluateur².

L'évaluation ne pourra jouer un rôle essentiel que si les professionnels et les responsables politiques y adhèrent. Des travaux antérieurs de l'OCDE ont mis l'accent sur l'importance de l'évaluation pour définir les principes qui doivent orienter les politiques. Ils ont aussi souligné que l'efficacité et l'efficacité des politiques à l'égard des PME doivent faire l'objet d'une évaluation périodique.

² A Lisbonne, le Conseil européen a défini ses objectifs dans la mesure où, à l'horizon 2010, l'Union entend « devenir l'économie de la connaissance la plus compétitive et la plus dynamique du monde, capable d'une croissance économique durable accompagnée d'une amélioration quantitative et qualitative de l'emploi et d'une plus grande cohésion sociale ». Union européenne, Livre vert, « *L'esprit d'entreprise en Europe* » 2003.

Le Royaume-Uni avait auparavant annoncé un objectif tout aussi ambitieux, qui était de « faire du Royaume-Uni l'endroit le plus favorable au monde à la création et à l'expansion des entreprises », UK Small Business Service, 2000

II. ÉVALUATION DES POLITIQUES A L'ÉGARD DES PME

Le rapport comportera cinq parties :

- le contexte de l'évaluation ;
- le cadre d'évaluation ;
- la démarche en six étapes ;
- les imperfections de la démarche en six étapes ;
- l'évaluation des bonnes pratiques ;
- les enseignements dégagés de l'expérience.

Le contexte de l'évaluation

Dans leur examen de l'évaluation des politiques dans les domaines de l'innovation et des techniques, Papaconstantinou et Polt (1997) offrent une définition très utile de l'évaluation :

« L'évaluation est un processus par lequel on cherche à déterminer aussi systématiquement et objectivement que possible la pertinence, l'efficacité et l'impact d'une activité au regard de ses objectifs, notamment par l'analyse de sa mise en œuvre et de sa gestion administrative. »

Plusieurs mots ou expressions de cette définition méritent d'être soulignés. Le premier est le mot « processus ». L'un des thèmes essentiels du présent rapport est que l'évaluation n'est pas une activité « ponctuelle », réalisée à l'issue d'un programme donné, mais qu'elle fait partie intégrante d'un processus visant à améliorer une politique ou la prestation d'un service.

Pour illustrer cette notion, la figure 1 montre de manière très schématique de quelle manière une politique est mise en application. Elle suppose un processus linéaire, qui commence par un « énoncé de projet » et aboutit à la formulation d'une loi. Les deuxième et troisième étapes, dites « cibles » et « mesures », se produisent quand la loi est « interprétée » par les fonctionnaires afin de pouvoir l'appliquer aux circonstances pratiques auxquelles ils sont confrontés. On citera pour exemple de ces objectifs et cibles ceux du *Small Business Service* du Royaume-Uni en 2000-2001. Les objectifs sont énumérés à l'encadré 1, et les cibles à l'encadré 2.

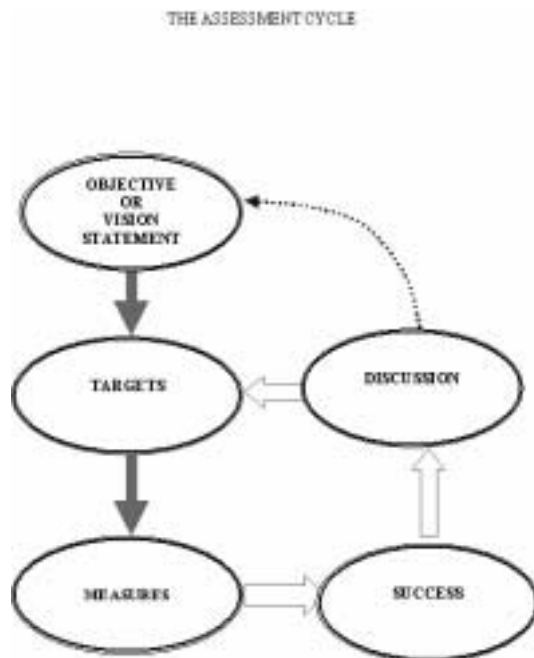
Encadré 1. Les objectifs stratégiques du Small Business Service du Royaume-Uni 2000-2001

- Être le porte-parole efficace des petites entreprises au sein du gouvernement - veiller à ce que les pouvoirs publics soient conscients des besoins des entreprises
- Œuvrer en faveur d'un cadre réglementaire qui allège au maximum les charges des entreprises
- Établir et assurer le fonctionnement d'un service d'aide aux entreprises de niveau international pour rehausser la compétitivité et la rentabilité des petites entreprises
- Promouvoir l'esprit d'entreprise à tous les niveaux de la société, notamment auprès des groupes sous-représentés et défavorisés.

Figure 1. Le cycle d'évaluation : l'évaluation en soi



FIGURE 2. LE CYCLE D'ÉVALUATION : L'ÉVALUATION DANS LE CADRE DU PROCESSUS D'ACTION PUBLIQUE



Encadré 2. Les résultats et les cibles du Small Business Service du Royaume-Uni 2000-2001

- Offrir une assistance à 5 000 entreprises avant mars 2002.
- Taux de satisfaction de la clientèle de 80 % pour l'ensemble des services.
- Accroître la notoriété spontanée et assistée des services de la marque *Business Link* parmi les petites entreprises à 20 % et 80 % respectivement avant mars 2002.
- Cinq mille demandes par jour sur le site Internet de *Business Link*.
- Deux mille appels par jour au centre d'appel national.
- Faire en sorte que tous les prestataires de services *Business Link* respectent les nouvelles normes en avril 2002 au plus tard.
- Mettre en place *Business Link University* avant septembre 2001.
- Établir de nouveaux flux de financements pour assurer le maillon manquant du financement des entreprises nouvellement créées et des petites entreprises en expansion.
- Lancer avant octobre 2001 un programme destiné à aider les petites entreprises à comprendre l'investissement extérieur et à s'y préparer.
- Créer de nouveaux fonds de capital-risque régionaux et y effectuer les premiers investissements avant mars 2002.
- Lancer un nouveau fonds d'incubation d'entreprises avant septembre 2001 et annoncer les premiers prêts en décembre 2001.
- Atteindre l'objectif d'un million d'entreprises britanniques offrant des services en ligne avant mars 2002.
- Mille programmes TCS avant la fin de 2001.
- Quinze mille placements STEP d'ici septembre 2001.
- Élaborer un plan d'action pour mettre en application les recommandations du *Council for Excellence in Management and Leadership* avant octobre 2001.

Revenons maintenant à la figure 1. Celle-ci suppose que la politique mise en œuvre fera l'objet, à l'issue d'un certain délai, d'un « examen » ou d'une « évaluation ». On peut voir dans cette opération la quatrième étape. Elle est souvent réalisée soit quand il est proposé d'amender la législation, soit quand il s'agit d'amplifier sa portée, de la réduire ou de l'abolir. Ou bien, comme le signale le document de la CEAP (2002), même si l'évaluation n'a pas d'influence sur les projets en cours, elle est porteuse d'enseignements pour les projets ultérieurs.

Le fait que l'évaluation constitue la quatrième étape du processus linéaire (et la dernière de la figure 1) justifie qu'on la qualifie de « dernière étape du processus ». Il est alors difficile de raccorder cette notion à la définition de l'évaluation énoncée plus haut, qui la décrit comme un « processus ». Dans les cas où l'évaluation n'a été réalisée que pour donner des indications quant à l'ampleur du budget alloué à un programme, il est difficile d'y voir un « processus ».

Pour que ce processus fonctionne bien, deux changements s'imposent. Tous deux font néanmoins intervenir une cinquième étape, illustrée sous la forme d'une case supplémentaire, dite « Analyse », à la figure 2. Le premier type d'analyse reconnaît qu'une évaluation, lorsqu'elle a lieu, a toutes les chances de conclure que le programme a été plus fructueux sur certains plans que sur d'autres. Il arrive aussi fréquemment que les programmes fonctionnent bien dans des domaines que leurs concepteurs n'ont pas envisagés, ou qu'ils jugeaient peu importants au moment où le programme a été élaboré [voir l'encadré 3]. L'exercice d'évaluation a donné l'occasion d'examiner les objectifs du programme. Il a fallu que les

responsables décident si ce dernier devait mettre davantage l'accent sur les domaines où il obtenait manifestement des résultats probants et accorder moins d'importance à ceux où son incidence était moindre. L'autre solution consistait à réaffirmer l'importance des objectifs originaux et à remodeler le programme en conséquence.

Encadré 3. Évaluation du Programme « Shell Technology Enterprise » (STEP)

STEP est un programme de huit semaines qui permet aux étudiants de deuxième année du premier cycle d'études universitaires de réaliser, au cours des vacances d'été, un projet supervisé et choisi dans une PME. Des prix sont accordés aux projets d'excellence.

A l'origine, le programme STEP avait pour objectifs d'infléchir le point de vue des diplômés de l'enseignement supérieur à l'égard du travail dans les PME et de les inciter à créer leur propre société. Il escomptait par ailleurs qu'après avoir engagé pour la première fois un diplômé, les petites entreprises seraient plus disposées à renouveler l'expérience ultérieurement.

L'évaluation a conclu qu'aucun des objectifs initiaux ne s'était véritablement concrétisé. De fait, les étudiants ayant participé au programme étaient plus enclins à quitter l'université pour travailler dans une grande entreprise. Le STEP a toutefois eu des retombées favorables en ce qu'il a augmenté les chances des diplômés (des moins qualifiés et de ceux qui venaient de régions fortement touchées par le chômage en particulier) de trouver rapidement un emploi.

À la lumière de ces constatations, le programme a été élargi, mais il n'a plus pour ambition première de faire évoluer les mentalités.

Source : Fraser, Storey et Westhead (2003)

L'exemple ci-dessus montre qu'il est possible de faire de l'évaluation un *processus* en réinjectant ses conclusions dans le débat une fois qu'elle est achevée. Un deuxième apport au *processus*, peut-être plus efficace encore, consiste à intégrer l'expertise de ceux qui doivent la réaliser aux objectifs et à l'énoncé de projet au moment où la législation est élaborée. C'est ce que la figure 2 illustre sous forme de pointillés. L'évaluateur serait chargé de vérifier que la loi a été formulée de manière à ce qu'une évaluation satisfaisante puisse être réalisée. Sa présence garantirait que les objectifs et les cibles de la législation sont suffisamment précis pour permettre aux futurs évaluateurs d'effectuer leur travail. Il veillerait également à la mise en place d'un dispositif de collecte des données afin de pouvoir conduire cette opération. Enfin, il convient d'allouer un budget à l'évaluation, qui prenne en considération l'ampleur et la complexité du programme, et de formuler une déclaration-cadre sur la façon d'y procéder.

L'intervention de l'évaluateur dès les phases liminaires de la réforme législative permet de s'attaquer à trois problèmes de fond. Le premier, qui sera analysé plus en détail ultérieurement, est que les objectifs des programmes sont souvent exprimés en termes trop flous pour permettre aux évaluateurs de mesurer les effets du programme. L'intervention de l'évaluateur au moment où la loi est formulée oblige les législateurs à se poser deux questions importantes :

- « Comment savoir si la loi/le programme a fait ses preuves ? »
- « Comment savoir si la loi/le programme a déçu les attentes ? »

Le deuxième problème auquel se heurtent les évaluateurs est celui de l'absence de données valables. Dans certains cas, on ne dispose même pas de renseignements sur les entreprises bénéficiaires du programme, mais c'est relativement rare. Bien plus souvent, l'évaluateur est gêné par le manque d'informations sur les entreprises qui n'ont pas participé au programme et pourraient ainsi servir de « groupe de contrôle » pour estimer l'effet des politiques. Enfin, comme on le verra plus loin, l'évaluation tirerait également profit de données sur les entreprises qui ont demandé à participer au programme mais ont essuyé un refus, ou sur celles qui n'étaient pas informées du programme mais remplissaient les conditions requises pour y participer, ou encore sur celles qui en étaient informées mais ont choisi de ne pas soumettre de demande. Quand l'évaluateur entre en jeu dès le début, la collecte des données peut commencer dès la mise en œuvre du programme, voire avant, ce qui permet de réaliser une évaluation de meilleure qualité.

Le troisième problème est que, l'évaluation n'étant pas envisagée avant le début du programme, voire longtemps après, les responsables du programme ont le sentiment qu'elle leur est imposée. Un débat s'ouvre alors quant à la nature, à l'ampleur et au budget de l'opération. Si les personnes chargées de l'exécution des programmes se trouvent soudainement contraintes de procéder à une évaluation, il est concevable qu'elles en soient irritées. Il s'agit inévitablement d'un exercice de longue haleine, surtout pour le personnel « en première ligne », pour lequel le coût d'opportunité lié à la prestation de services aux entreprises peut paraître très élevé. Dans les cas où l'exercice d'évaluation est financé sur le budget propre du programme, la réaction risque d'être encore plus vive. Pour le plus grand bénéfice de l'ensemble des parties, il convient de chercher à résoudre tous ces problèmes au moment où la législation est élaborée, de manière à ce que chacun soit informé de ce qu'une évaluation aura lieu, des principes sur lesquels elle sera fondée et de la façon dont elle sera financée. C'est pourquoi il est essentiel que l'évaluation soit envisagée dans le cadre d'un processus, mais à condition que toutes les parties soient informées de son intégration à tous les niveaux.

La deuxième expression clé de la définition de l'évaluation que donnent Papaconstantinou et Polt est « *aussi systématiquement et objectivement que possible* ». L'évaluation ayant habituellement lieu « en bout de chaîne », il existera sans doute des intérêts bien ancrés après plusieurs années d'exécution d'un programme, notamment parmi les bénéficiaires directs des programmes, comme les entreprises à qui les fonds sont alloués, mais aussi parmi les responsables de leur mise en œuvre et de leur administration. Toutes choses égales par ailleurs, ces groupes se prononceront vraisemblablement en faveur de la poursuite ou de l'élargissement des opérations. Or, la tâche de l'évaluateur consiste à mesurer « *systématiquement et objectivement* » les mérites du programme, et il se heurtera probablement à ceux qui y sont attachés par principe. Seules des techniques objectives, dont bon nombre sont examinées plus loin, lui permettront de prouver son indépendance aux exécutants du programme.

Quoi qu'il en soit, l'objectivité est certes un élément important, mais l'évaluateur accomplira sa mission au mieux s'il bénéficie du soutien des responsables du programme. Il est censé être compétent en matière d'évaluation, tandis que les exécutants des programmes détiennent, de par leurs fonctions, des informations de première main. Ils pourront souvent fournir une explication simple, mais inattendue, aux « associations » établies par les évaluateurs. Sans leur savoir et leur appui, ces derniers risquent donc d'aboutir à des déductions erronées. La distinction est par conséquent subtile entre la nécessité pour l'évaluateur de faire preuve d'objectivité et d'indépendance et celle d'obtenir le soutien et les informations privilégiées des exécutants des programmes. Il sera nettement plus facile de mobiliser cette assistance si les responsables du programme sont clairement informés pendant la phase de conception qu'une évaluation aura lieu.

La troisième expression importante de la définition de Papaconstantinou et Polt est « *la pertinence, l'efficacité et l'impact d'une activité au regard de ses objectifs* ». L'hypothèse implicite est ici que la stratégie a des objectifs clairs, qui sont énoncés en termes suffisamment précis pour servir à

l'évaluateur. Dans la pratique, cela se vérifie rarement. Les objectifs sont souvent énoncés en termes génériques (on parle par exemple de « développer l'esprit d'entreprise d'une économie »), comme on l'a signalé au paragraphe 19, ce qui est d'une utilité limitée pour l'évaluateur. Ainsi, dans le cas simple de l'objectif consistant à « développer l'esprit d'entreprise d'une économie », l'évaluateur pourrait envisager toute une panoplie de mesures pour estimer l'impact de la stratégie retenue. Il pourrait par exemple se fonder sur le nombre de nouvelles entreprises créées, sur le nombre de sociétés créées par des personnes jeunes ou défavorisées, ou par des personnes qui n'ont jamais travaillé dans les affaires auparavant, sur l'augmentation du nombre total d'entreprises ; il pourrait vérifier si ces entreprises se développent, si elles s'internationalisent et, bien entendu, si elles survivent. Tous ces éléments permettraient de mesurer « l'esprit d'entreprise » d'un pays, mais celui-ci obtiendra en toute vraisemblance de meilleurs résultats sur certains points que sur d'autres. Quand toutes sortes de mesures permettent d'évaluer une politique, et que celle-ci obtient de bons résultats dans certains domaines et des résultats moins probants dans d'autres, les partisans de la poursuite d'un programme seront inévitablement enclins à choisir les mesures ou les cibles où celui-ci s'est avéré performant. A l'inverse, ceux qui souhaitent y mettre un terme ou confier son exécution à d'autres retiendront les mesures qui leur conviennent. Richard Harrison et Claire Leitch (1996) ont remarquablement résumé ce phénomène dans un essai intitulé « *Whatever you hit call the target: An alternative approach to small business policy* », qui souligne à quel point il importe de définir clairement les cibles qui correspondent le plus précisément aux objectifs stratégiques³.

Or, dans la réalité, les mesures permettant d'évaluer l'efficacité d'un programme ne sont souvent pas spécifiées. L'évaluateur doit donc déduire intuitivement les objectifs originaux du législateur au moment où il a élaboré le programme. Il lui faut souvent, bien des années après qu'une législation est entrée en vigueur, définir plusieurs mesures qui cadrent avec son objectif initial. Cela le conduit alors inévitablement à retenir plusieurs mesures de performance, probablement en coopération avec les responsables de la mise en œuvre du programme.

En principe, cette analyse est vraiment utile et témoigne bien du rôle de l'évaluation en tant que « *processus* ». Dans la pratique, malheureusement, elle amène les responsables du programme à préconiser les mesures sur lesquelles ils ont enregistré d'assez bons résultats pour évaluer celui-ci, ce qui donne fréquemment lieu à un débat quelque peu déplacé dont ni les évaluateurs ni les responsables du programme ne ressortent grandis.

La conclusion essentielle de cet examen du contexte de l'évaluation est que celle-ci doit faire partie intégrante du *processus* stratégique. L'évaluation ne doit pas être « gardée pour la fin ». Elle doit au contraire être un élément fondamental de la formulation initiale de la stratégie. Dès que la stratégie est mise en œuvre, tous les organismes et individus responsables de son exécution doivent être informés qu'une évaluation aura lieu. Une fois engagée, celle-ci doit servir de base à un dialogue avec les responsables politiques dans le but d'améliorer la stratégie. Ses résultats pourront par la suite servir de fondement à un débat de fond sur les modalités idéales de l'interaction entre les pouvoirs publics et les PME.

³ Bien évidemment, cela ne signifie pas que les cibles ne peuvent être modifiées. De fait, les évaluations ont souvent pour effet utile d'amener de telles modifications. (Voir l'exemple de l'encadré 3). Il est en revanche déconseillé d'en sélectionner plusieurs, ce qui permet aux responsables politiques de prétendre que les programmes sont performants quelle que soit la cible qu'ils se trouvent atteindre à un moment donné. Il est donc essentiel de définir clairement les critères permettant de juger de l'échec d'une stratégie.

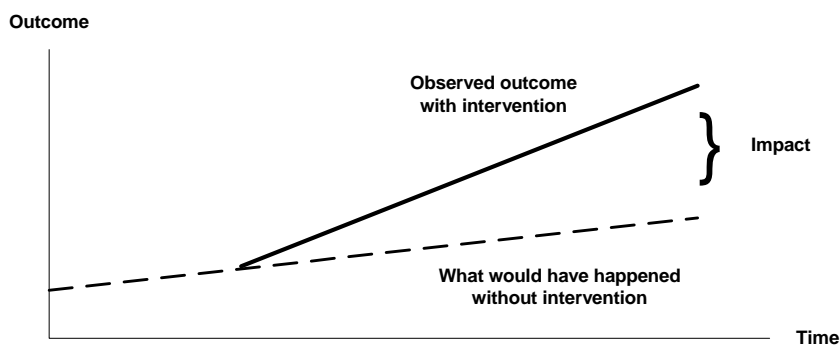
Le cadre d'évaluation

Le Groupe de travail sur les PME a dégagé sept critères qui permettent d'évaluer les stratégies : justification, additionnalité, pertinence, supériorité, efficacité systémique, efficacité propre et efficacité d'adaptation.

La question de l'additionnalité est un élément primordial de l'évaluation. Il s'agit de l'impact réel du dispositif ou programme. Si celui-ci n'est certes pas facile à mesurer, on peut néanmoins penser qu'il se traduit par une augmentation de la production, de l'emploi, du chiffre d'affaires ou de l'activité à l'exportation que l'on peut imputer à l'existence du programme. En d'autres termes, il s'agit de l'activité qui n'aurait pas existé sans le programme, mais qui est attribuable aux entreprises qui y participent. C'est ce qu'illustre la figure 3 ci-dessous, tirée d'Oldsman (2002).

La figure montre que, pour tout résultat donné, l'impact d'une stratégie peut être mesuré par la différence entre le résultat observé par suite de l'intervention et ce qui se serait produit sans cette intervention. On voit que les deux résultats commencent à diverger après la mise en œuvre de la stratégie.

Figure 3. Incidence d'une intervention



Source : Oldsman (2002)

Le concept est simple, mais définir l'impact du programme « *aussi systématiquement et objectivement que possible* » peut poser de multiples problèmes car, pour diverses raisons, celui-ci n'est pas facile à observer. D'abord, les changements susceptibles d'intervenir dans les entreprises du fait de leur participation au programme ne sont pas toujours clairement définissables ; autrement dit, la mesure des résultats est imprécise. Certains programmes sont censés augmenter les chances de survie des entreprises ; d'autres doivent amener une progression des ventes, des bénéfices ou de l'emploi ; d'autres encore sont supposés stimuler l'innovation ou les ventes sur les marchés étrangers. Certains devraient faire tout cela à la fois et, dans d'autres cas, on ne sait précisément sur quels terrains les entreprises sont censées marquer une amélioration. Il est donc impératif de définir des mesures appropriées des résultats avant de procéder à l'évaluation.

Le deuxième problème est que la participation au programme précédera les améliorations. Si l'on prend l'exemple de la figure 3, le point de divergence des lignes n'interviendra pas immédiatement après l'exécution du programme. Certains programmes auront peut-être des retombées plusieurs années avant d'autres. Ainsi, un programme qui prévoit de subventionner la participation des PME à un Salon

professionnel international afin de les inciter à s'internationaliser⁴ aura probablement des retombées rapides sur les ventes. Par contre, un programme destiné à financer les activités de recherche et de développement des PME ne devrait guère avoir d'effet avant plusieurs années.

Le troisième problème est la multitude de facteurs qui influencent les résultats d'une PME, en dehors de sa participation au programme : la compétence de son propriétaire ; son secteur et son lieu d'activité ; le climat macroéconomique ; et le hasard. En principe, c'est seulement quand ces facteurs « exogènes » sont pleinement pris en compte que l'impact d'un programme peut être estimé. La section qui suit montre de quelles manières différentes ces problèmes ont été abordés.

La méthode d'évaluation en six étapes

Bon nombre de méthodes ont été adoptées pour tenter d'évaluer l'impact des programmes, mais toutes peuvent s'inscrire dans le cadre de la méthode d'évaluation idéale en six étapes [Storey (2000)], présentée au tableau 1 ci-dessous.

Tableau 1. Six étapes vers une méthode d'évaluation idéale de l'impact des politiques à l'égard des PME

Suivi	
Étape I	* Mise en place des programmes
Étape II	* Opinion des bénéficiaires
Étape III	* Avis des bénéficiaires sur le plus apporté par le dispositif d'aide
Évaluation	
Étape IV	Comparaison des performances des entreprises « bénéficiaires » avec celles d'entreprises « de référence »
Étape V	Comparaison avec des entreprises « témoins »
Étape VI	Prise en compte des biais induits par la sélection

Les six étapes décrites correspondent à une graduation dans la complexité des procédures d'évaluation, l'étape I équivalant à la procédure la moins sophistiquée et l'étape VI à la plus perfectionnée. Cette complexité vise ici à traduire le degré de certitude avec lequel le responsable politique peut attribuer les changements intervenus dans les entreprises assistées dans le cadre du programme - souvent réunies sous l'appellation de « groupe bénéficiaire » - à leur participation au programme en question. Autrement dit, on peut y voir dans quelle mesure le responsable politique est certain que tous les autres facteurs sont maintenus constants.

Le tableau 1 établit une distinction entre suivi et évaluation. Les étapes I à III ne sont pas considérées comme relevant d'un exercice d'évaluation et sont plutôt assimilées à des activités de suivi. La différence entre suivi et évaluation tient à ce que le suivi repose exclusivement sur les avis des bénéficiaires, alors que l'évaluation vise, d'une manière ou d'une autre, à mettre en regard les avis exprimés ou les décisions prises par les bénéficiaires d'une mesure et ceux émanant d'entreprises n'en

⁴ « Accès aux marchés internationaux », exposé de la délégation grecque, Groupe de travail sur les petites et moyennes entreprises de l'OCDE, Athènes, 26-28 avril 1998 [DSTI/IND/PME(98)6].

ayant pas bénéficié afin de présenter le « contrefactuel » illustré par la ligne en pointillés à la figure 3. On estime que le décalage entre les changements réels et le contrefactuel représente l'impact de la politique, à savoir son « additionnalité ».

La suite du chapitre décrit brièvement les différentes étapes afin de justifier l'adoption d'une méthode d'évaluation plus perfectionnée. Le tableau 2 montre qu'il est utile de connaître les caractéristiques des entreprises qui participent aux programmes, mais que ces informations ne permettent pas au responsable politique d'évaluer si le programme a rehaussé la performance de ces entreprises.

Tableau 2. ÉTAPE I : Mise en place des programmes

<p><u>QUESTIONS :</u></p> <ul style="list-style-type: none"> * Nombre d'entreprises participantes * Secteurs d'activité * Lieu d'activité * Taille * Sommes engagées <p><u>PROBLÈMES :</u></p> <ul style="list-style-type: none"> * Ne donne pratiquement aucune indication quant à l'efficacité de la stratégie * Ne donne pratiquement aucune indication en ce qui concerne la satisfaction des objectifs.

Une méthode simple et très répandue d'aborder cette question consiste à s'adresser aux entreprises qui ont participé au programme, comme indiqué au tableau 3⁵. Cette démarche peut fournir des indications utiles quant à l'exécution du programme, mais présente des imperfections notables quand il s'agit d'évaluer son impact global. Certaines des personnes interrogées, si leur entreprise prospère depuis qu'elle participe au programme, peuvent voir dans ces progrès la marque de leur propre réussite et se montrer de ce fait très réticentes à attribuer un quelconque changement à leur participation au programme. A l'inverse, d'autres peuvent juger qu'il est de leur intérêt que le programme se poursuive et exagèrent donc les avantages qu'elles en retirent. Un troisième groupe sera tout simplement incapable de définir cette incidence compte tenu de toutes les autres mutations intervenues dans l'entreprise. Enfin, ces enquêtes porteront en toute probabilité sur les PME qui ont survécu plutôt que sur celles qui ont cessé leur activité. Toutes choses égales par ailleurs, les entreprises demeurées en activité auront vraisemblablement une opinion plus positive, ce qui aura pour effet d'introduire un biais favorable dans l'estimation du programme.

Tout cela signifie que si l'on se contente de demander l'opinion des bénéficiaires, on court le risque d'introduire une distorsion incompatible à laquelle il est difficile de remédier. De nombreuses « évaluations » relevant de l'étape II ont néanmoins été effectuées, sans doute dans le but de satisfaire à l'obligation de conduire cette opération, et ce au plus faible coût possible. Les problèmes associés à cette démarche sont toutefois si importants que ce type d'évaluation devrait être découragé, sauf s'il s'agit d'études de cas réalisées dans le but d'obtenir des renseignements.

⁵ Se reporter au document « Évaluation thématique des impacts des fonds structurels sur les petites et moyennes entreprises », présenté à l'OCDE, novembre/décembre 1998.

Tableau 3. Étape II : Opinion des bénéficiaires

<p>QUESTIONS :</p> <ul style="list-style-type: none">* Participants au programme : l'ont-ils apprécié ?* Entreprises : Ont-elles rencontré des problèmes lors de sa mise en œuvre ?<ul style="list-style-type: none">: Les formalités ont-elles été trop longues ?: ou trop lourdes ? <p>PROBLÈMES :</p> <ul style="list-style-type: none">* Le fait que le programme ait été apprécié ne permet pas de conclure à son efficacité.* Cette démarche peut seulement apporter des indications quant à l'exécution de la stratégie<ul style="list-style-type: none">- mais ce n'est pas là la question essentielle

L'étape III fait appel à une méthode un peu plus évoluée. Elle interroge également les bénéficiaires du programme, mais aborde explicitement la question du contrefactuel. Elle leur demande ce qui ce serait produit si leur entreprise n'avait pas participé au programme. De ce fait, les personnes interrogées sont tenues de prendre le contrefactuel en considération dans leurs réponses.

Tableau 4. Étape III : Avis des bénéficiaires sur le plus apporté par le dispositif d'aide

<p>QUESTIONS :</p> <ul style="list-style-type: none">* Les entreprises pensent-elles que le dispositif a fourni une « additionnalité » ?* Auraient-elles obtenu ces résultats de toute façon ?* Le dispositif a-t-il un effet de « substitution » ? <p>PROBLÈMES :</p> <ul style="list-style-type: none">* Les bénéficiaires donnent les réponses qu'ils croient attendues* Aucune vérification n'est possible* Cette méthode ne donne qu'un aperçu des entreprises survivantes

Cette démarche impose toutefois une gymnastique cérébrale considérable aux personnes interrogées. Certaines seront éventuellement à même de procéder à ce calcul complexe, mais d'autres n'en seront pas capables. Malheureusement, le chercheur ne peut déterminer avec précision quelles personnes appartiennent à chaque catégorie. Ces problèmes se multiplient quand la personne interrogée ne travaillait pas dans l'entreprise au moment où le programme a débuté.

Tous ces éléments font que l'étape III est aussi très imparfaite et montre bien qu'un contrefactuel explicite est nécessaire. Celui-ci caractérise les étapes IV à VI, définies sous le nom d'évaluation.

L'évaluation la plus simple consiste à comparer la performance des entreprises « bénéficiaires » (celles qui participent au programme) à celle d'un autre groupe de sociétés non participantes. Cette comparaison repose sur le principe que la différence entre les entreprises « bénéficiaires » et « non bénéficiaires » est la participation au programme, et que cela transparaît dans les écarts de performance observés. Les entreprises non bénéficiaires prises en considération à l'étape IV font généralement aussi partie de l'économie ou de la localité.

Tableau 5. Étape IV : Comparaison entre la performance des entreprises « bénéficiaires » et celle des entreprises « de référence »

<p><u>APPROCHE</u> :</p> <ul style="list-style-type: none"> * Croissance de l'emploi/des ventes dans les entreprises bénéficiaires par rapport aux entreprises « de référence » * Survie des entreprises bénéficiaires en comparaison aux entreprises « de référence » <p><u>PROBLÈME</u> : * Les entreprises bénéficiaires ne sont pas des entreprises ordinaires</p>
--

Or, il est évident que les sociétés participant au programme peuvent différer systématiquement des autres entreprises de l'économie. Elles peuvent s'en distinguer sur le plan de la taille, de la localisation géographique, du secteur, de l'âge ou d'autres éléments observables, mais qui tous peuvent influencer systématiquement sur leur performance. Ces facteurs, parce qu'ils sont en principe identifiables et mesurables, sont dits « observables ». Les écarts de performance entre les deux groupes peuvent donc dériver de la participation au programme, mais aussi de ce que les facteurs « observables » ne sont pas maintenus constants pendant l'étape IV. La non prise en compte de ces facteurs signifierait, si l'on se réfère à la figure 3, que la ligne en pointillés n'a pas été correctement tracée. L'impact du programme serait en fait partiellement dû aux caractéristiques des entreprises ; autrement dit, son incidence véritable est soit surestimée, soit sous-estimée.

L'étape V s'emploie à remédier à ce problème. Elle retient généralement des facteurs observables de première importance, comme le régime de propriété, la taille, le secteur et la localisation géographique, et vérifie, par un processus d'appariement, que les entreprises bénéficiaires et non bénéficiaires ne diffèrent pas les unes des autres sur ces points. Une fois cette opération terminée, les écarts de performance entre les deux groupes sont imputés à la participation au programme.

Tableau 6. Étape V : Comparaison avec les entreprises « témoins »

<u>APPROCHE</u> :	
* Comparer les entreprises bénéficiaires et les entreprises témoins en termes :	
- d'âge	
- de secteur	
- de régime de propriété	
- de localisation géographique	
* Comparer la performance des deux groupes sur une même période	
<u>PROBLÈMES</u> :	* Il peut s'avérer extrêmement difficile de procéder à un appariement parfait en fonction de quatre critères.
	* Biais de sélection de l'échantillon
	- Ce sont les entreprises les plus « motivées » qui demandent à participer au programme
	- Cette méthode attribue les écarts de performance au programme et pas à la motivation

L'étape V s'efforce véritablement de résoudre le problème du contrefactuel, mais ne saurait pour autant être considérée comme une « pratique optimale ». Les résultats demeurent quelque peu ambigus dans la mesure où les écarts de performance constatés peuvent aussi être dus à la présence de facteurs « non observables ». Le plus manifeste en serait le dynamisme ou la motivation du propriétaire de la PME, ce qui introduirait différents biais de sélection dans les techniques d'évaluation.

On peut citer deux exemples de biais de sélection : l'autosélection et la sélection par un comité. Le biais d'autosélection peut par exemple se produire dans le cas de programmes destinés à assister les entreprises en pleine essor : les sociétés désireuses de se développer seront probablement plus enclines à demander une aide que celles dont les ambitions sont modérées. Le problème surgit quand cette caractéristique n'apparaît que partiellement dans les facteurs « observables », puisque ces derniers sous-estiment les résultats que l'entreprise aurait enregistrés si le programme n'avait pas existé. La ligne « contrefactuelle » en pointillés, si l'on se réfère à la figure 3, est tracée trop bas, et l'impact du programme est par conséquent surestimé.

Autrement dit, les entreprises demandeuses, étant donné la plus forte motivation de leurs propriétaires, auraient probablement affiché une meilleure performance que les entreprises « témoins », même en l'absence de programme. Il est donc abusif d'attribuer tous les écarts de performance à la participation au programme, mais il est difficile de déterminer quelle part exacte lui est imputable.

Le deuxième exemple est celui de la sélection par un comité. Celle-ci concerne les programmes où seule une partie des entreprises candidates voient leur demande satisfaite. Dans ce cas, la décision relève d'un comité ou d'un groupe équivalent, les « meilleures » entreprises ou demandes obtenant le financement. Si le comité est efficace, il écartera les entreprises les moins prometteuses, de sorte que les autres seront clairement sélectionnées de manière non aléatoire parmi les candidates. Comme dans le cas

précédent, les entreprises choisies auraient vraisemblablement affiché une meilleure performance que les autres, même si le programme n'avait pas existé. Les écarts de performance observés entre les sociétés participantes et les sociétés témoins ne peuvent donc être attribués uniquement à la participation au programme. 6

Tableau 7. Étape VI : Comparaison avec des entreprises « témoins » de façon à corriger les biais induits par la sélection

<p><u>APPROCHE</u></p> <p>* Utilisation de techniques statistiques : estimateur de Heckman en deux étapes</p> <p>aux fins d'essais et d'ajustement</p> <p>* Recours à des groupes aléatoires</p> <p><u>PROBLÈMES</u> * Les responsables politiques (et certains universitaires) sont réservés à l'égard des « ajustements » statistiques</p> <p>* L'utilisation de groupes aléatoires pourrait signifier que les deniers publics sont octroyés à des entreprises ou à des personnes dont on sait qu'elles n'en tireront pas profit.</p>

Plusieurs techniques statistiques permettent désormais de résoudre le problème du biais de sélection. Curieusement, si ces techniques sont aujourd'hui couramment employées pour évaluer le marché du travail et les programmes d'aide sociale, elles le sont beaucoup moins pour estimer l'impact du soutien aux entreprises. Smith (2002) en a effectué une analyse utile, qui est présentée à l'encadré 4 ci-dessous. Il montre qu'il est désormais possible, sur le plan technique, de résoudre le problème de la sélection et de réaliser des évaluations conformément à la méthode de l'étape VI.

⁶ Il se peut, bien entendu, que le comité soit incapable de faire la différence entre les entreprises prometteuses et les autres. Dans ce cas, il risque de ne pas jouer un rôle utile. Il pourrait aussi remplir une autre fonction et veiller à ce qu'un pourcentage donné de certaines catégories d'entreprises reçoivent un financement. Il pourrait s'agir d'entreprises implantées dans certaines régions, ou dont les propriétaires sont désavantagés à certains égards. Si ce rôle est clairement défini, la méthode appliquée à l'étape V conviendra. Par contre, si cette tâche n'est pas explicitement confiée au comité, l'application de cette méthode pourrait entraîner une sous-estimation des effets du programme.

Encadré 4. Est-il possible d'obtenir des groupes de comparaison non aléatoires véritablement comparables par les méthodes statistiques ?

Il existe au moins cinq techniques statistiques, pas nécessairement inconciliables, pour déterminer l'effet véritable d'un programme sur une quelconque variable de résultat quand les usagers du programme se distinguent des non-usagers sur d'autres plans que celui de l'utilisation du programme. D'abord, on peut simplement procéder à un contrôle statistique des variables observées qui influencent le résultat économique et pourraient être corrélées à l'utilisation du programme en les intégrant à l'équation d'estimation utilisée pour prédire la variable de résultat. Cette méthode permet notamment de réduire le biais d'estimation de l'impact du programme quand on dispose de données sur le plus grand nombre possible de variables de ce type. Elle ne peut par contre corriger les biais éventuellement dus à des variables non observées qui sont corrélées à la fois aux résultats économiques et à l'utilisation du programme. Elle suppose par ailleurs que la forme fonctionnelle en vertu de laquelle les variables observées influencent les résultats économiques est connue.

Une deuxième approche, qui est une variante de la première, est dite estimation « de la différence entre les différences », ou estimation « de la différence entre les différences entre les différences » (ou estimation DD et DDD) (Meyer, 1995). Dans le cadre de la méthode DD, on compare la différence, avant et après le programme, entre les différences distinguant les bénéficiaires des non-bénéficiaires du programme ou des mesures. Avec la méthode DDD, si l'on a des raisons de penser que ces mesures auront plus d'effets sur certains types d'usagers que sur d'autres, on peut comparer la différence entre les groupes supposés à fort impact et à faible impact dans le groupe des bénéficiaires et dans celui des non-bénéficiaires, avant et après l'application des mesures. Une approche DD revient à supposer qu'il est possible de neutraliser des facteurs qui influencent par ailleurs les résultats économiques en tenant compte des effets de la période temporelle, que l'entité appartienne au groupe des bénéficiaires ou à celui des non-bénéficiaires, ou bien à un groupe à fort impact ou à faible impact. Cette méthode a pour inconvénient que de nombreuses autres variables, observées et non observées, peuvent aussi influencer les résultats économiques et sont corrélées à l'utilisation du programme. On peut conjuguer la deuxième approche à la première en ajoutant quelques-unes de ces autres variables observées à l'équation d'estimation.

Une troisième méthode consiste à appairer les usagers du programme à des non-usagers similaires sur le plan des caractéristiques observées. Des études récentes ont montré que cette opération devrait s'efforcer de trouver des usagers et des non-usagers aussi semblables que possible en termes de « score de propension », à savoir la probabilité estimée en fonction de variables observées qu'une entité donnée fera appel au programme (Smith et Todd 2001 ; Heckman, Ichimura, et Todd, 1999). Ce mécanisme fonctionne bien si l'on dispose de données sur toutes les variables qui permettent de prédire l'utilisation du programme et sont en outre corrélées aux résultats économiques. Il ne fonctionnera pas correctement si les variables non observées qui prédisent l'utilisation du programme et sont corrélées avec les résultats économiques sont trop nombreuses. Par ailleurs, dans bon nombre de cas, il sera impossible d'appairer convenablement des usagers à des non-usagers ; les estimations de la méthode d'appariement ne mesurent alors valablement que les effets moyens du programme sur les catégories d'usagers qu'il a été possible d'appairer correctement à des non-usagers.

Une quatrième approche consiste à modéliser explicitement une sélection dans le programme, ainsi que la façon dont elle est corrélée aux variables non observées qui influencent les résultats économiques (Murnane, Newstead et Olsen, 1985). Il convient pour cela d'estimer trois équations : une équation qui détermine les résultats économiques pour les usagers du programme ; une deuxième qui détermine ces résultats pour les non-usagers ; une troisième qui définit si une entité donnée utilise le programme. L'estimation de la troisième équation permet d'ajouter un terme de « correction du biais de sélection » à chacune des deux premières équations, ce qui corrige théoriquement la distorsion due aux variables non observées qui influencent les résultats économiques et sont corrélées à l'utilisation du programme. Cette méthode suppose que l'on a défini avec exactitude les variables et la forme fonctionnelle qui doivent être incorporées aux trois équations. Elle suppose en outre une distribution statistique particulière des facteurs non observables (le « terme d'erreur ») introduits dans les trois équations.

La cinquième méthode fait intervenir une « variable instrumentale » qui prédit l'utilisation du programme et n'est pas corrélée à des variables non observables qui influencent les résultats économiques (Angrist and Krueger, 2001). Dans ce cas, on examine seulement l'évolution des résultats économiques imputable aux changements intervenus dans l'utilisation du programme qui sont associés, sur le plan statistique, à des modifications de la variable instrumentale. On suppose que les effets de ces modifications programmatiques induites par l'instrument sur les résultats économiques montrent les retombées véritables du programme, ces modifications n'étant pas corrélées à des variables non observables qui prédisent les résultats économiques du fait qu'elles sont produites par une variable instrumentale non corrélée à de telles variables. Le problème consiste à trouver des instruments de cette nature. La variable instrumentale doit être à même d'expliquer l'utilisation du programme, autrement la méthode d'estimation écarterait trop d'informations. Cela dit, sa corrélation avec les variables non observables qui influencent les résultats économiques doit être faible (dans l'idéal, nulle). Vérifier les hypothèses relatives à la corrélation d'un instrument proposé avec des variables non observables est un exercice délicat. Il est difficile de trouver de bons instruments, et ceux-ci ne conviennent pas toujours à tous.

On peut combiner ces cinq méthodes sous différentes formes. Les chercheurs peuvent par exemple créer des groupes de données appariées, intégrer à l'estimation des contrôles statistiques pour différentes variables observées et utiliser des variables instrumentales pour la variable du programme.

Source : Smith (2002)

La méthode en six étapes : modifications et inconvénients

La délégation japonaise à l'OCDE a soumis des observations sur la méthode en six étapes. Ses opinions sont reprises textuellement à l'encadré 5. En résumé, la délégation avance que les méthodes les plus perfectionnées (étape VI) conviendraient mieux dans certaines circonstances que dans d'autres. Elle fait valoir que dans certains programmes, le lien causal entre la participation au programme et la performance des entreprises apparaît plus clairement que dans d'autres.

Encadré 5. Un thème transversal : l'évaluation des politiques et programmes d'aide aux PME – Les observations formulées par la Délégation du Japon à l'attention du Groupe de travail sur les PME et l'entrepreneuriat de l'OCDE

Les « Six étapes vers une méthode d'évaluation idéale de l'impact des politiques à l'égard des PME » offrent un cadre d'excellente qualité pour évaluer l'additionnalité des programmes.

Néanmoins, étant donné la diversité des objectifs et des dispositifs des politiques publiques à l'égard des PME, le document devrait se pencher sur le fait que les méthodes de suivi ou d'évaluation appropriées peuvent varier selon les programmes. La démarche adéquate devrait être définie pour chaque programme en fonction de la relation causale entre la production, les résultats et le coût de l'évaluation, ainsi que de l'échelle budgétaire du programme.

Ainsi, il est difficile et onéreux de mesurer l'impact des programmes de prestation de services d'information ou de formation par la méthode de l'étape IV car la relation causale entre la production (information, formation) et les résultats (ventes, emploi, exportations, etc.) est plutôt vague et influencée par d'autres facteurs extérieurs. Il semblerait que les étapes II ou III conviennent à ce type de programmes, et qu'il serait superflu de recourir à des dispositifs plus perfectionnés.

Il est en revanche assez facile de mesurer l'incidence de programmes de subventions ou de garantie de prêts, par exemple, car ils sont directement liés aux opérations d'investissement. Dans ces cas, les méthodes des étapes V ou VI sembleraient adaptées.

Le tableau ci-dessous illustre les étapes cibles appropriées et devrait être soumis à l'analyse des experts.

Programme	Suivi			Évaluation			Relation causale	Coût de l'évaluation
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
							production -résultat	
Prestation de services d'information	<input type="radio"/>	<input type="radio"/>					faible	élevé
Séminaires, formation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				faible	faible
Consultation, encadrement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>				faible	faible
Subventions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	forte	faible
Garantie de prêt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	forte	faible
Programme d'équité	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	forte	faible
Incitations fiscales	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>		forte	élevé

Ces observations soulèvent la question cruciale de savoir si le même degré de raffinement doit s'appliquer à tous les programmes, indépendamment de leur envergure et de leur complexité.

L'argument en faveur d'une évaluation de l'étape VI dans tous les cas est que, si certains programmes sont évalués de manière moins approfondie que d'autres, il sera difficile de procéder à des comparaisons globales de leur efficacité. Supposons que la politique d'aide aux PME ne comporte que deux programmes : le programme A offre des subventions, le programme B dispense une formation. Si la politique est appliquée avec efficacité, l'effet marginal par unité monétaire sera identique dans les deux cas. Or, il serait impossible d'énoncer un tel jugement si le programme A était évalué selon une méthode et le programme B selon une autre.

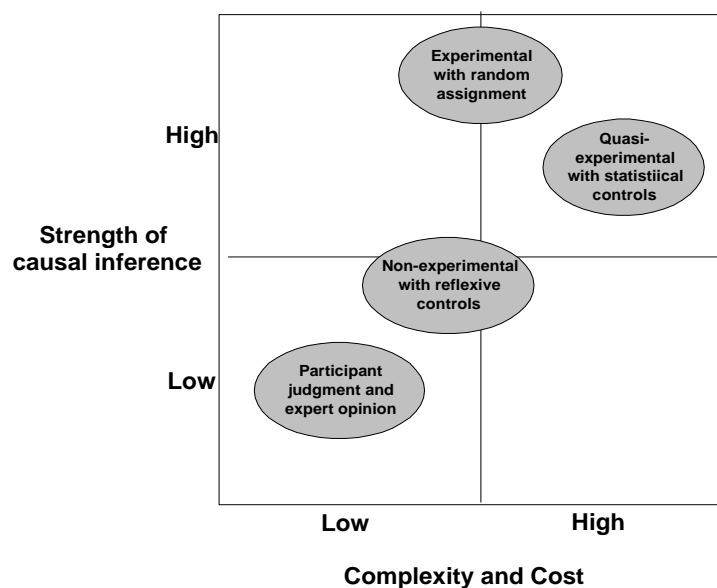
Ce point revêt une très grande importance car la méthode en six étapes montre que ce sont les procédures les plus complexes qui ont le plus de chances de définir précisément le « contrefactuel ». Dans la plupart des cas, une définition imprécise du contrefactuel entraîne une surévaluation de l'impact du programme. Il apparaît clairement que l'effet des programmes évalués selon les méthodes les moins sophistiquées (premières étapes) est généralement jugé supérieur à celui des programmes évalués au moyen de méthodes plus perfectionnées (dernières étapes). Malheureusement, les études qui font appel à ces dernières sont aussi les plus coûteuses. Un entretien avec un participant au programme est facile à réaliser, peu onéreux, et ne requiert pratiquement pas d'analyse. Le contraste avec les approches perfectionnées est flagrant, puisque celles-ci doivent prévoir l'évaluation dès le début du programme, assurer un suivi des participants et des non-participants, et faire appel à une expertise statistique généralement rare et coûteuse pour conduire l'analyse. Les responsables de l'évaluation sont donc confrontés à un problème d'ordre quasi-éthique : les études les plus perfectionnées risquent de conclure à un impact plus faible du programme, mais elles sont en outre plus chères et probablement plus longues à réaliser. Ce dilemme est analysé en détail à la section (vi).

Le problème primordial est donc que si certains programmes étaient évalués moins rigoureusement que d'autres, l'allocation des financements entre les programmes publics d'aide aux PME serait problématique. Il est cependant difficile de justifier que le même degré de perfectionnement et, partant, les mêmes budgets d'évaluation, s'appliquent à tous les programmes, quelle que soit leur envergure. La figure 3 ci-dessous montre ainsi que plus le budget est élevé, plus l'évaluation des retombées du programme est précise.

On voit apparaître ici un arbitrage important dans l'évaluation. Dans l'idéal, tous les programmes relevant d'une même politique d'aide aux PME devraient être évalués selon les mêmes méthodes, de manière à ce que l'on puisse comparer valablement leurs effets marginaux. Cela permettrait éventuellement de transférer les ressources du programme A au programme B et d'améliorer l'impact global de la politique à l'égard des PME.

Or, il est difficile de justifier un même degré de complexité analytique pour les programmes de petite et de grande envergure. Cela dit, si les petits programmes sont moins rigoureusement évalués, ils sembleront peut-être avoir un impact plus important, ce qui justifiera l'augmentation de leur budget.

Figure 3. Arbitrages dans la conception de l'évaluation



Enfin, il importe également d'insister sur le fait que la finalité de la méthode « en six étapes » est d'évaluer l'impact des programmes et des politiques particuliers. Étant donné qu'elle procède en établissant une comparaison entre les entreprises qui participent aux programmes ou qui bénéficient des politiques, et les autres, elle ne peut être utilisée pour évaluer les politiques qui s'appliquent simultanément à *l'ensemble* des entreprises d'une économie donnée. Elle ne peut par conséquent pas évaluer convenablement les incitations fiscales dont bénéficient toutes les PME. De même, elle risque de sous-estimer l'impact global des programmes d'aide aux entreprises, lorsqu'il existe entre eux une forte interdépendance. Par essence, elle évalue les effets marginaux de chaque programme ; or, la somme de ces effets peut être inférieure à l'effet global. Toutefois, ces considérations d'ordre somme toute assez technique ne doivent pas occulter le fait que les contribuables doivent être convaincus que l'impact de chaque programme a été évalué en utilisant les méthodes les plus perfectionnées sur le plan technique.

Exemples d'évaluation relevant des bonnes pratiques

On trouvera ici des exemples d'évaluations des politiques d'aide aux PME dans les pays de l'OCDE conduites selon des méthodes au moins équivalentes à celles de l'étape V. Ces études examinent la performance des entreprises « bénéficiaires » et la comparent à celle d'entreprises similaires n'ayant pas reçu d'aide. Selon les « contrôles » utilisés, cette démarche donne une mesure de l'additionnalité. Par exemple, si la politique évaluée consiste à accorder des prêts et des subventions à un groupe de PME, il faudrait, au minimum, vérifier que les entreprises bénéficiaires et les non-bénéficiaires ne diffèrent pas clairement en termes de taille, de secteur, d'âge, de localisation ou de croissance antérieure.

Cela dit, comme on l'a noté au chapitre précédent, les progrès statistiques de ces dernières années ont permis aux analystes de remédier à un problème potentiel dont la méthode de l'étape V fait abstraction. Il s'agit du fait que, outre les facteurs « observables » (âge, secteur, régime de propriété et localisation géographique), il existe parfois des facteurs « non observables » qui influencent la disposition des entreprises à demander des subventions. Un autre problème que la méthode de l'étape VI résout est celui du biais de sélection par un comité. Ce biais se produit du fait que beaucoup de programmes ne disposent pas de ressources suffisantes pour servir tous les candidats. Seuls certains d'entre eux obtiendront satisfaction, la décision de leur octroyer ou non le bénéfice de l'aide étant prise par les bailleurs de fonds compte tenu des chances de réussite qu'ils attribuent à chacun. Si les décideurs - le comité - ne sont pas à même de faire la différence entre les entreprises prometteuses et celles qui le sont moins, il existe un risque majeur de biais dans la mesure où les entreprises qui reçoivent un soutien sont celles qui auraient obtenu de meilleurs résultats que celles à qui cette chance est refusée, même si ces aides n'avaient pas existé.

L'étude présentée à l'encadré 6 est tirée d'une analyse d'Oldsman (2002), et examine l'incidence des services de formation et de conseil fournis aux PME par les Industrial Resource Centers aux États-Unis. Elle établit, au moyen d'une méthode d'appariement, que la participation au programme a eu un impact favorable sur la productivité, mais ne détermine pas aussi clairement si cette démarche résout de manière explicite le problème de la sélection.

Encadré 6. Le programme « Industrial Resource Center » (IRC)⁷.

Ce programme a été mis en place en 1988 pour aider les petites et moyennes entreprises manufacturières à perfectionner leurs pratiques commerciales et à moderniser leur capacité de production afin de stimuler la croissance économique en Pennsylvanie. Les IRC visent à accomplir leur mission au travers d'une palette complète d'activités qui allient des services de conseil et de formation. Le gouvernement de l'État a consacré quelque 84 millions de dollars au programme depuis son démarrage.

Une évaluation approfondie du programme a été conduite en 1999. Elle comportait une estimation de l'impact des services sur les sociétés participantes en termes d'accroissement de la productivité et de la production. Afin de neutraliser les effets d'un biais de sélection éventuel, l'analyse fait appel à une méthode comportant deux étapes⁸. La première consiste à estimer la probabilité pour que les entreprises deviennent des clients de l'IRC en fonction de leurs caractéristiques. La deuxième fait intervenir une estimation de l'impact du programme IRC sur les sociétés après neutralisation des facteurs qui influencent le développement de la productivité et de la production ainsi que des éventuels biais de sélection. Le modèle estimé se fonde sur la fonction de production Cobb-Douglas modifiée qui comprend des facteurs particuliers aux usines.

L'analyse s'appuie sur des données de panel portant sur chacune des usines de fabrication - le Longitudinal Research Dataset (LRD) - tenues à jour par le Center for Economic Studies du U.S. Bureau of the Census. Cet ensemble fournit des informations détaillées sur les livraisons, l'emploi, le coût des facteurs, le secteur, et d'autres identificateurs juridiques et administratifs à l'échelon des usines. Il est établi à partir du Census of Manufactures, réalisé tous les cinq ans, et de l'enquête annuelle sur les manufactures. En Pennsylvanie, le LRD a permis d'obtenir des renseignements sur les sociétés clientes et non clientes. Le lien entre les entreprises figurant dans la base de données administrative de l'IRC et le LRD a été établi au moyen d'une procédure d'appariement mise au point par le Center for Economic Studies. Cette procédure a recensé, à partir de son nombre permanent d'usines (PPN), 2 839 établissements clients exclusifs de l'IRC sur les années couvertes par le recensement. Le groupe de référence comportait un nombre équivalent de sociétés.

L'analyse a démontré que le programme avait un effet substantiel sur les clients de l'IRC. Dans le cas de la cohorte antérieure à 1992, l'écart concernant le taux de croissance de la production et de la productivité directement attribué aux services de l'IRC est estimé à 1,8 % et à 3,6 % par an sur une période de dix ans.

⁷ Nexus Associates, Inc. (1999).

⁸ On trouvera d'autres informations sur les méthodes utilisées dans le cadre de cette analyse dans les publications suivantes : Maddala (1994), Jarmin (1997) et Heckman (1974 et 1979).

Encadré 7. Programme « Small Business Innovation Research » (SBIR)⁹.

Le programme SBIR a été mis sur pied par le Congrès des États-Unis en 1982. La loi d'autorisation prescrivait à tous les organismes fédéraux consacrant chaque année plus de 100 millions dollars à la recherche extérieure de réserver un pourcentage fixe de ces financements aux petites entreprises. Ce pourcentage a augmenté au fil du temps pour atteindre 2,5 %. Entre 1983 et 1995, les petites entreprises ont reçu plus de 6 milliards de dollars à ce titre.

Une évaluation du SBIR a été effectuée en 1996, qui a essentiellement porté sur l'effet des financements en termes de progression des ventes et de l'emploi. L'analyse comparait notamment la performance des entreprises bénéficiaires de subventions à celle de sociétés similaires non-participantes, en exerçant un contrôle statistique sur les facteurs suivants : âge, localisation géographique, financement par capital-risque antérieur, activité globale de capital-risque dans la région et dans le secteur. L'échantillon comportait 541 entreprises ayant reçu des fonds dans le cadre de la phase II du SBIR au cours des trois premiers cycles programmatiques, et 594 entreprises témoins.

Les données ont été réunies à partir de la documentation du programme et des informations contenues dans les annuaires et bases de données publics.

L'analyse a démontré que les bénéficiaires du SBIR ont affiché une plus forte croissance que les entreprises témoins en termes de ventes et d'emplois, mais que ces effets étaient circonscrits à des domaines qui attiraient d'importants capitaux-risque.

Le programme américain de très grande ampleur SBIR, qui accorde des subventions aux PME pour conduire des activités de recherche et de développement, est examiné à l'encadré 7. Cette analyse est tirée d'Oldsman, mais s'inspire de l'étude originale de Lerner (1996). L'étude conclut que la participation au programme a des retombées fructueuses, mais elle fait appel à une méthode de l'étape V. Il est difficile de mesurer l'incidence de l'effet de sélection, le programme comportant à la fois une autosélection et une sélection par un comité. Même si, comme on l'a précédemment noté, l'exclusion du facteur de sélection a généralement pour effet de valoriser le programme, il se peut que ce ne soit pas le cas ici car le mandat du comité pourrait exiger une répartition géographique équitable des subventions.

Une deuxième difficulté concernant le SBIR consiste à définir dans quelle mesure il fournit une véritable additionnalité aux budgets de R-D des petites entreprises, et dans quelle mesure il « évince » les capitaux privés. Le fait que les États-Unis aient confié au National Research Council le soin d'examiner ces problèmes témoigne de l'importance qu'ils attachent à cette question¹⁰.

L'encadré 8 présente une étude britannique qui vise à évaluer l'impact des services de conseil du dispositif « Business Links » financé sur fonds publics. L'application d'une méthode de type VI lui permet de dégager quelques éléments montrant que le programme a un impact sur la productivité des entreprises. Cela dit, elle ne conclut pas à un effet substantiel de l'assistance fournie sur le chiffre d'affaires ou sur l'emploi.

⁹ Lerner, Josh (1996).

¹⁰ Charles Wessner, « US support for Innovation and Small Business: Measuring SBIR Outcome », Document présenté à la 19^e session du Groupe de travail de l'OCDE sur les petites et moyennes entreprises et l'entrepreneuriat, 9-10 décembre 2002.

Encadré 8. Mesurer les effets sur la productivité du dispositif Business Links¹¹

Business Links est le système d'accès par un « guichet unique » aux services de conseil aux (petites) entreprises bénéficiant du soutien des pouvoirs publics qui a été mis en place au Royaume-Uni.

L'évaluation avait pour objectif de déterminer si les entreprises ayant bénéficié de services dispensés via Business Links avaient obtenu de meilleurs résultats que les entreprises n'ayant bénéficié d'aucune aide. Sachant que Business Links avait notamment pour cible les entreprises employant entre 10 et 200 salariés et présentant un potentiel de croissance, l'évaluation visait également à apprécier l'efficacité de ce ciblage. L'étude a porté sur 137 entreprises ayant bénéficié de services dispensés par le réseau Business Links et sur 196 entreprises témoins, et elle concernait la période allant de 1996 à 2000. Elle a recensé les facteurs ayant une influence sur le fait qu'une entreprise sollicite ou non une assistance auprès du réseau Business Links tout en prenant en compte l'incidence des services de conseil dispensés via Business Links sur les résultats obtenus.

- Il est apparu que rien n'attestait qu'en 1996, Business Links s'adressait aux entreprises connaissant une croissance rapide.
- Des éléments, toutefois sujets à caution, attestant que l'aide dispensée via Business Links en 1996 avait des retombées positives sur l'accroissement de la productivité, ont été recueillis.
- Un effet positif, mais statistiquement non significatif, de l'aide dispensée via Business Links sur le chiffre d'affaires et l'emploi a été décelé.
- L'analyse a permis de dégager un certain nombre d'autres facteurs contribuant à l'augmentation de la productivité, du chiffre d'affaires et de l'emploi, qui vont des conditions sur le marché à la stratégie de l'entreprise, en passant par les caractéristiques propres du dirigeant et de l'entreprise elle-même. La mise en évidence de ces facteurs souligne la complexité du processus d'expansion des entreprises et les difficultés qu'elle soulève pour les spécialistes chargés de le modéliser et de le stimuler.

Dans d'autres cas, la bonne pratique consiste à évaluer l'effet non pas d'un programme unique, mais d'un groupe de programmes. L'encadré 9 prend l'exemple de la Nouvelle-Zélande et note que diverses méthodes sont employées pour étudier les différentes composantes de la politique appliquée.

¹¹ Exposé de George Bramley, DTI, au Groupe de travail sur les PME de l'OCDE à sa 17^e session, 2001

Encadré 9. L'évaluation du *Cadre de croissance et d'innovation* et des programmes de développement des entreprises en Nouvelle-Zélande comporte trois éléments essentiels :

- *Le rapport sur les indicateurs de croissance et d'innovation (Growth and Innovation Indicators Report)* – ce rapport annuel réunit les informations de référence portant sur un groupe d'indicateurs censés suivre l'évolution de la croissance globale et des résultats en matière d'innovation.
- *Le bilan annuel des programmes de développement des entreprises (The Annual Stocktake of Enterprise Development Programmes)* – ce rapport annuel regroupe les renseignements sur l'efficacité et l'efficacités (par rapport aux objectifs du programme) des programmes de développement des entreprises. Il s'agit d'une évaluation globale qui vise à recenser les données et les enseignements nouveaux surgis pendant l'année. Il présente aussi, le cas échéant, quelques commentaires sur les conclusions qui se dessinent quant à l'impact des programmes.
- *Examens approfondis de chaque programme* – ce sont des examens exhaustifs des mesures sur lesquelles se fondent les programmes, des retombées des programmes, et de l'efficacité et de l'efficacités de leur exécution. En 2003, les examens du Programme régional de partenariat et celui de l'*Investment Ready Scheme* ont été menés à terme.¹²

Enfin, l'encadré 10 présente l'évaluation du programme canadien SBLA, qui a pour objet de fournir des fonds publics à l'appui des financements reposant sur l'actif accordés par les institutions financières. L'enjeu pour l'évaluateur consiste à démontrer que sans le SBLA, le prêt n'aurait pas été consenti.

Il lui faut pour cela s'en remettre en grande part aux avis des parties concernées (banques, agents publics et bénéficiaires des prêts), qui ne sauraient être considérées comme désintéressées. Quoi qu'il en soit, il n'existe pas de groupe « non bénéficiaire » manifeste avec lequel établir une comparaison. Il est donc judicieux d'appliquer différentes méthodes quand elles tendent vers des résultats globalement analogues.

¹²

Davis, Nick, « Best practice policies for SMEs: Country presentations on SME programme evaluations, New Zealand: evaluation of Investment Ready Scheme », DSTI/IND/PME (2003), 11 mai 2003. « Regional Partnership Programme Review », Septembre 2003.

Encadré 10. La loi sur les prêts aux petites entreprises (SBLA) - Canada¹³

Le SBLA est entré en vigueur en 1961 au Canada pour aider les petites entreprises à surmonter les difficultés auxquelles elles sont confrontées pour obtenir des prêts reposant sur leurs actifs. Le programme vise à utiliser les deniers publics pour mobiliser des fonds privés auprès des institutions financières. Le gouvernement cherche à limiter autant que possible les effets d'aubaine et d'éviction.

Comme ce programme était en place depuis longtemps, bien qu'il ait connu plusieurs réorientations au fil des ans, une évaluation en a été réalisée en 1996. Il est difficile de prouver que les petites entreprises n'auraient pas obtenu de prêts sans le SBLA si l'on s'en tient à une seule méthode. L'évaluation a donc fait appel à tout un éventail de techniques, notamment des études de cas, des études économétriques et des consultations avec les parties prenantes.

Elle a conclu que jusqu'à 86 % des prêts consentis dans le cadre du SBLA étaient des prêts additionnels. Dans la plupart des cas, cela tenait au fait que, sans le programme, ils n'auraient pas été accordés ou bien que, s'ils l'avaient été, les conditions en auraient été plus rigoureuses pour les PME, ou la procédure plus lente.

Enfin, l'encadré 11 puise dans les travaux de Lars Barger-Sjogren (2003) pour citer un exemple d'une méthode de type VI appliquée en Suède à l'évaluation d'un vaste programme de formation de la main d'œuvre sur le lieu de travail.

Encadré 11. Formation en cours d'emploi en Suède

Cette étude pointue a défini la compétitivité comme étant la productivité, la marge bénéficiaire brute et la structure du capital. Elle a examiné plus de 2 000 entreprises, bénéficiaires ou pas d'un programme. Elle a également tenu compte du problème du biais de sélection.

La mesure normalisée de la productivité employée a montré qu'avant le programme, en 1995, le groupe bénéficiaire accusait un retard de près de 30 % par rapport au groupe de contrôle. A l'issue du programme, en 2002, l'écart s'était néanmoins sensiblement resserré.

Enseignements dégagés

Le présent chapitre interprète les observations ci-dessus, en s'efforçant de les rattacher aux préoccupations des responsables des politiques à l'égard des PME.

- Le financement des programmes d'aide publique aux PME est assuré par le contribuable. Les responsables de l'élaboration et de l'exécution de ces programmes doivent donc veiller à ce qu'ils aient un bon rapport coûts-résultats. Les sommes engagées ne sont pas insignifiantes, et le montant total des budgets publics consacrés aux PME est rarement quantifié, principalement en raison du nombre considérable de programmes publics de cette nature.

¹³

« Évaluation de la loi sur les prêts aux petites entreprises », document présenté à la session du Groupe de travail sur les petites et moyennes entreprises qui s'est tenue à Helsinki les 20 et 21 mai 1999.

L'étude récente résumée à l'annexe 1 fait exception à cette règle. Elle indique que le montant annuel total des sommes engagées par le gouvernement britannique pour fournir des services aux petites entreprises s'élève à quelque GBP 8 milliards (dernière année pour laquelle les chiffres sont disponibles). A titre comparatif, le budget alloué aux services de police s'élevait à GBP 7 milliards. Cela montre à quel point il importe de connaître le montant total du budget et de veiller au bon rapport coût-performance des programmes. L'évaluation devrait jouer un rôle essentiel à cet égard en informant les responsables politiques et les contribuables.

- L'exemple ci-dessus montre d'une part que les financements publics sont substantiels, et d'autre part que la responsabilité de ces fonds est largement répartie entre différentes administrations publiques. Cela pose un problème stratégique majeur car si les administrations disposent parfois de « lignes budgétaires » pour les petites entreprises, leurs priorités globales peuvent résider ailleurs. De ce fait, même si les dépenses publiques totales consacrées aux petites entreprises sont substantielles, elles risquent fort de manquer de cohérence et d'une orientation claire. C'est ce qui ressort de la recommandation de la déléguée finlandaise Annukka Lehtonen, inspirée des résultats de « l'Atelier sur les micropolitiques visant à promouvoir l'entrepreneuriat » qui s'est tenu à Helsinki en avril 2002, avec le projet public d'entrepreneuriat. La citation ci-dessous reprend la notion fondamentale de « cohérence des politiques ».

Encadré 12. Annukka Lehtonen : de la « cohérence des politiques »

La réussite des politiques en faveur des entreprises passe par une véritable coopération horizontale dans le cadre de laquelle différents ministres et ministères travaillent en collaboration. En d'autres termes, des fonctionnaires désignés de plusieurs ministères sont chargés de veiller à ce que des lois visant à promouvoir l'entrepreneuriat soient rédigées, et à ce que l'impact des diverses mesures sur l'entrepreneuriat soit pris en compte.

Bien entendu, cette coopération horizontale exige, outre la coopération administrative ordinaire, que les fonctionnaires responsables soient issus de différents ministères et mandatés par le gouvernement pour jouer un rôle « d'entrepreneur politique ». Ils doivent adopter une démarche volontariste et assumer au sein de leur propre ministère la responsabilité de l'élaboration de mesures destinées à promouvoir l'activité industrielle et commerciale. Autrement dit, les ministères demeurent responsables de l'élaboration des mesures, mais les fonctionnaires désignés sont tenus de vérifier le déroulement effectif des travaux. Il s'agit de décloisonner les différentes branches administratives, traditionnellement étroitement surveillées dans les ministères. En ce qui concerne la prise de décision, il nous faut également instaurer un groupe de travail composé de ministres sectoriels ou de ministres représentant les parties gouvernementales.

- Compte tenu de l'objectif de cohérence des politiques, le présent rapport a rappelé avec insistance les recommandations antérieures de l'OCDE selon lesquelles l'évaluation doit faire partie intégrante du processus stratégique. Elle n'est pas un « plus » facultatif. Au contraire, des éléments d'évaluation doivent être intégrés à toutes les phases de la formulation des politiques. Cela signifie qu'il faut, lors de l'élaboration des lois, établir les objectifs de

manière à ce qu'une évaluation puisse être réalisée. Il conviendrait en outre de prévoir une méthode et un budget pour l'évaluation, et même d'entamer la collecte de données.

- Le rapport indique par ailleurs clairement que des progrès considérables ont été enregistrés en matière statistique ces dernières années, et qu'il est possible de procéder à des évaluations nettement plus précises de l'impact des programmes aujourd'hui qu'il y a quelques années seulement. Nul ne peut plus prétendre que les méthodes d'évaluation sont très imparfaites. Tout simplement, les techniques existent et sont abondamment utilisées dans bien d'autres domaines des programmes publics, comme ceux ayant trait au marché du travail. Il est curieux que ces techniques aient mis tant de temps à se développer en ce qui concerne l'évaluation des politiques à l'égard des PME, mais il est désormais d'usage de les mettre au point pour différents types de programmes d'aide aux PME.

Des questions importantes restent toutefois ouvertes à la discussion et au débat, qui sont brièvement évoquées ci-dessous.

Si les techniques statistiques disponibles permettent d'évaluer bien plus précisément les politiques qu'auparavant, les responsables politiques et les fonctionnaires sont quelque peu réticents à l'idée de les voir se généraliser. Cela s'explique en partie par le fait que tous ne disposent pas des compétences mathématiques et quantitatives nécessaires à leur compréhension, et en partie par la méfiance répandue à l'égard des opérations visant à quantifier l'incidence des politiques, au motif qu'elles conduisent à négliger les aspects les moins « matériels » de ces politiques. Ainsi, quels que soient les progrès des techniques quantitatives en termes de précision, certains groupes seront toujours hostiles à leur adoption généralisée.

D'autres groupes y sont peut-être moins radicalement opposés, mais restent sceptiques quant à la réalisation d'évaluations plus perfectionnées. Ces questions sont examinées ci-dessous.

i) La relation entre la qualité technique des méthodes d'évaluation et le programme estimé

Il est clair que, de façon générale, l'impact d'un programme a tendance à être d'autant plus faible que l'évaluation qui le met en évidence est sophistiquée. Il apparaît aussi nettement que les évaluations complexes sont plus coûteuses, essentiellement parce que davantage d'éléments sont maintenus constants au lieu d'être attribués au programme. Cela pourrait facilement engendrer des conflits d'intérêt car si les responsables du programme sont également chargés d'établir le budget de l'évaluation, des pressions risquent de s'exercer pour que les programmes privilégiés soient évalués selon des méthodes moins sophistiquées. A priori, une telle marge d'appréciation semble abusive et porte à recommander une séparation des responsabilités entre les évaluateurs et les responsables des programmes.

ii) Fréquence et ampleur de l'évaluation, compte tenu de l'envergure du programme

Il se trouve que dans les pays de l'OCDE certains programmes sont évalués, parfois plus d'une fois, tandis que d'autres ne le sont pas. Comme on l'a déjà signalé, certains sont en outre évalués selon des techniques considérablement plus perfectionnées que d'autres.

Bien entendu, cela tient en partie au fait que les pouvoirs publics établissent un lien entre les budgets alloués aux opérations d'évaluation et l'envergure des programmes. L'inconvénient de cette démarche est toutefois qu'elle complique la comparaison entre les résultats des évaluations de petite et de grande ampleur. Or, pour que la politique d'aide aux PME satisfasse aux critères de « *supériorité* » et

« *d'efficacité systémique* » de l'OCDE, l'efficacité marginale de l'ensemble des politiques doit être identique, ce qui ne pourra être déterminé si les politiques sont évaluées par des méthodes différentes.

Une solution pragmatique serait de convenir au minimum que tous les programmes doivent être évalués, de manière à ce qu'aucun ne passe au travers des mailles du filet. Il semble en outre approprié d'établir un lien entre l'importance du budget de l'évaluation et l'envergure du programme. Enfin, pour faciliter les comparaisons générales, on pourrait suggérer qu'aucun programme ne soit évalué selon des méthodes d'un niveau inférieur à l'étape IV, et que plus les programmes seront importants, plus il y aura de probabilité qu'ils soient évalués selon les méthodes des étapes V et VI.

iii) *Faut-il confier l'évaluation à des organismes internes ou externes ?*

Il y a lieu de déterminer si les programmes doivent être évalués par les services responsables de leur mise en œuvre, par d'autres instances de l'administration ou par des organismes extérieurs. Il existe des arguments en faveur d'évaluateurs de tous ces groupes.

Il est clair que les Ministères responsables de l'exécution des programmes disposent d'une compréhension et d'informations de première main auxquelles aucun tiers n'aura accès. Ils seront aussi probablement plus à même de mobiliser le soutien des exécutants du programme. Par contre, cette proximité risque de porter atteinte à leur « indépendance », et ils pourraient faire l'objet d'une « récupération » ou être assujettis à des influences politiques.

Une autre solution consisterait à confier les évaluations à une branche spécialisée du gouvernement, indépendante du ministère chargé de l'exécution du programme. L'avantage réside ici dans le fait que cette branche aura plus de spécialistes compétents en matière d'évaluation à sa disposition, et jouira d'une plus grande indépendance que le ministère responsable du programme. Elle aura toutefois une connaissance moins approfondie du programme et devra miser davantage sur une coopération avec les exécutants.

Une troisième solution consiste à engager des spécialistes « extérieurs », par exemple des consultants ou des universitaires. Ces deux groupes présentent l'avantage d'être moins manifestement susceptibles de « récupération », et d'être des spécialistes dans leur domaine¹⁴. Le principal inconvénient au recours à des intervenants extérieurs est qu'une évaluation censée s'inscrire dans un *processus* d'amélioration des politiques doit conduire, comme l'illustre la figure 2, à une analyse. Celle-ci risque d'être moins mobilisatrice si elle est conduite par des tiers, éventuellement jugés moins bien informés, que si elle découle d'une évaluation conduite par les responsables du programme.

Il convient donc de procéder à un examen de ces arguments contradictoires afin d'aboutir à la décision importante de savoir qui conduira les évaluations.

¹⁴ On pourrait soutenir que les cabinets d'expertise-conseil ne sont pas totalement indépendants, puisque leur activité dépend du renouvellement de leurs contrats avec les pouvoirs publics. Il convient de rappeler ici l'argument avancé par Kuusisto et al. (1999) selon lequel les évaluations extérieures officielles se sont dans l'ensemble montrées plus favorables que celles réalisées par des entités indépendantes. Cela dit, il est tout aussi judicieux de souligner que les consultants ont pour réputation de mener leur mission à terme dans les délais impartis, caractéristique qui ne se retrouve pas toujours chez les universitaires !

iv) Mise en œuvre des conclusions de l'évaluation

Comme l'évaluation n'est pas l'étape finale, il est fondamental de savoir dans quels domaines elle a conduit à des réorientations de l'action menée. Ces réorientations peuvent se traduire soit par des modifications aux objectifs du programme, soit par des aménagements dans la manière dont le programme est mis en œuvre ou exécuté. Il arrive par exemple qu'un programme soit élargi et devienne accessible à des entreprises appartenant à des secteurs auparavant exclus de son champ d'application, ou que des modifications soient apportées aux modalités, conditions ou résultats escomptés du programme.

La procédure de mise en œuvre des modifications aux programmes variera selon l'ampleur et la notoriété du programme, et en fonction des processus politiques. Quoiqu'il en soit, les évaluations ont pour finalité première d'influencer l'action publique. Dans certains pays, il serait utile que les informations émanant de l'évaluation entrent dans le domaine public, surtout dans le cas de programmes très en vue, ce qui a plus de chances de se produire si l'évaluation a été conduite par des intervenants extérieurs.

Plusieurs solutions sont envisageables. L'une d'elle consisterait à attribuer à tous les programmes une « durée de vie » définie, leur reconduction étant fonction d'une évaluation satisfaisante. On pourrait voir dans cette démarche une variante à la loi « de caducité automatique ». Elle a pour avantage de garantir que tous les participants conviennent des objectifs et du mode d'évaluation du programme lors de son élaboration initiale. Elle permettrait par ailleurs de réorienter les programmes en fonction de l'évolution du climat économique.

v) Répercussions pour les pays à revenu faible et intermédiaire.

Il importe certes que les contribuables de tous les pays puissent vérifier l'efficacité des programmes publics, mais il convient également d'admettre que les ressources nécessaires aux évaluations feront considérablement défaut aux pays à revenu faible et intermédiaire. Les organismes donateurs peuvent ici remplir une fonction essentielle ; il leur faut établir une piste de vérification pour les financements, mais ils peuvent aussi montrer la voie en mettant en évidence les avantages d'une évaluation appropriée. On citera pour exemple les travaux de la Banque interaméricaine de développement (BID) et de FUNDES International. Ces deux organismes ont cherché à élaborer des instruments capables de concevoir, de suivre et d'évaluer des instruments de politique en faveur des PME en Amérique latine, et tout particulièrement au Mexique, au Chili et en Colombie. On trouvera un résumé des principaux résultats de ce projet à l'Annexe 2.

Les organisations internationales telles que l'OCDE pourraient aussi jouer un rôle en fournissant des études de cas portant sur des évaluations adaptées aux pays à revenu faible et intermédiaire. Enfin, des bourses et des stages d'étude à l'étranger pourraient être accordés aux économistes de l'Administration en début de carrière afin qu'ils se familiarisent avec les nouvelles techniques de recherche.

III. CONCLUSIONS

Ce document a présenté un examen approfondi des méthodes employées pour évaluer les composantes de la politique à l'égard des PME. Il s'inspire essentiellement de l'expérience des États Membres de l'OCDE mais, le cas échéant, fait aussi référence à des États non membres. Les conclusions principales sont présentées ci-après.

En premier lieu, cette évaluation ne peut être conduite correctement avant que les objectifs et cibles de la politique à l'égard des PME ne soient clairement spécifiés. Le document propose d'utiliser le cadre COCE pour formuler les politiques. La première lettre « C » y représente à la fois la clarté et la cohérence de l'action publique. « Clarté », dans ce contexte, signifie que la politique en question est pleinement comprise par ceux qui l'appliquent et par ceux qui sont censés en bénéficier. Le concept de « cohérence », quant à lui, est un peu plus complexe. Il a trait au besoin de veiller à ce que tous les secteurs de l'Administration s'emploient à travailler collectivement en faveur des PME. Plus précisément, les politiques à l'égard des PME ont peu de chance d'être « cohérentes » si elles sont appliquées, sans coordination, par de multiples administrations ou si quelques administrations travaillent à la réglementation des PME et d'autres à « l'assistance », sans mettre en place un dialogue approprié.

La lettre « O » signifie « Objectifs ». Pour garantir la clarté des politiques, il convient d'énoncer précisément leurs objectifs ; concrètement, ceux-ci peuvent consister à augmenter le nombre d'entreprises, le taux de création des entreprises ou leur valeur ajoutée. Il peut aussi s'agir de chercher à accroître le nombre d'entrepreneurs parmi des groupes définis en termes d'appartenance ethnique, de sexe, de désavantage ou de localisation.

La deuxième lettre « C » correspond à « cibles ». Celles-ci doivent être spécifiées de manière à ce qu'une évaluation ultérieure puisse être réalisée pour définir si elles sont atteintes ou pas. A titre d'exemple de transformation d'un « objectif » en « cible », prenons l'objectif consistant à « augmenter le taux de création des entreprises ». Dans ce cas, la cible pourrait être « d'augmenter la création d'entreprises de 10 % dans les trois prochaines années ». Ce n'est qu'ainsi que l'on pourra définir si la cible est atteinte ou pas.

La lettre « E » signifie « Évaluation », qui est le thème central du présent document. Celui-ci conclut au premier chef qu'il est impossible de conduire une évaluation en l'absence de cibles, car une évaluation ne peut s'effectuer que dans un cadre où les effets escomptés des politiques sont clairement spécifiés.

Le document conclut en deuxième lieu que l'évaluation doit être au cœur du processus d'élaboration des politiques. Elle ne doit pas être engagée uniquement à titre d'exercice comptable rétrospectif pour déterminer si l'argent public a été judicieusement dépensé, bien que cette fonction soit utile. L'évaluation ne doit pas être « gardée pour la fin », mais éclairer les politiques en vigueur de sorte que les objectifs et cibles puissent être modifiés en fonction de l'efficacité des mesures. De ce fait, les réflexions sur les méthodes d'évaluation devraient être intégrées à la formulation des politiques au stade de l'élaboration des nouvelles idées.

Veiller à ce que l'évaluation reçoive une priorité plus élevée signifie que tous les programmes, et pas quelques-uns seulement, sont évalués. Le document ne précise toutefois pas si tous les programmes doivent être évalués de manière aussi approfondie. L'argument fondamental en faveur d'un degré uniforme d'évaluation est qu'il permet d'établir des comparaisons valables entre les différents programmes. L'argument inverse avance qu'il est déraisonnable de prévoir des budgets d'évaluation similaires pour des programmes d'envergures très diverses.

Quoi qu'il en soit, le document souligne qu'il est désormais possible, sur le plan technique, de mesurer précisément l'impact des programmes. Il montre, dans le cadre de son approche « Six étapes vers une méthode d'évaluation idéale » que, si la méthode la plus perfectionnée est parfois appliquée, c'est là l'exception plutôt que la règle. Le problème vient de ce que les méthodes moins approfondies apportent généralement des réponses fallacieuses, le plus souvent parce qu'elles surestiment l'effet des politiques. La troisième conclusion principale du document est que, lorsque cela se justifie, il convient d'employer les méthodes les plus perfectionnées (étapes 5 ou 6).

Au moment d'engager ces évaluations, il convient de ne pas perdre de vue les éventuelles incompatibilités. D'abord, dans l'idéal, ceux qui conduisent les évaluations doivent être indépendants des responsables des programmes, comme dans tout audit. Mais si l'évaluation a pour fonction essentielle de concourir au perfectionnement constant de la politique, les évaluateurs ont besoin de la collaboration et de la participation actives de ceux qui l'ont élaborée et de ceux qui l'appliquent. Cela risque d'être plus difficile si l'évaluateur est vu comme un intrus. L'un dans l'autre, le document conclut, mais de justesse, que l'indépendance de l'évaluation revêt une importance cruciale.

Enfin, le document conclut que l'OCDE elle-même pourrait jouer un plus grand rôle dans la diffusion des « bonnes pratiques » en matière de techniques d'évaluation des politiques à l'égard des PME.

ANNEXE 1

Tableau A.1. Dépenses du gouvernement britannique consacrées à la prestation de services aux petites entreprises, 2001

(en millions de GBP)

Dépenses		5 342
<i>dont</i>		
Budget du <i>Small Business Service</i>	349	
Autres dépenses du ministère du Commerce et de l'Industrie et de ses agences	275	
BTI	44	
DCMS et ses agences	332	
DFES et ses agences	138	
RDA	274	
DWP	71	
Pouvoirs publics locaux	300	
Administrations fiscales (services de soutien)	81	
Prêts, conseils, soutien en nature du DEFRA (ministère de l'Agriculture)	3120	
DTLR et ses agences	107	
CE et BEI	246	
Mesures fiscales (Crédit d'impôt à la R-D ; 10 ; Impôt sur les sociétés ; Taux d'imposition sur les petites sociétés : 20 % ; Fonds de capital-risque ; « Enterprise Investment Scheme »)		2 590
Total		7 932
(en pourcentage du PIB)		(0,8)

Source : "Cross Cutting Review of Government Services for Small Businesses",
Ministère britannique du Commerce et de l'Industrie, septembre 2002, URN 02/1324.

ANNEXE 2

GUIDE A L'ÉVALUATION DES POLITIQUES EN FAVEUR DES PETITES ENTREPRISES⁺,**Maria Vega, Genente Internacional, Entorno Empresarial, Fundes International****Résumé analytique**

La Banque interaméricaine de développement (BID) et FUNDES International ont entrepris, dans le cadre d'une coopération technique, de mettre au point un instrument susceptible d'aider les responsables de la conception, de l'évaluation et du suivi des interventions d'aide. La méthodologie exposée dans ce guide sera donc utile à tous ceux qui engagent une évaluation exhaustive d'une politique d'aide aux microentreprises et aux PME (à partir d'un jeu commun d'indicateurs). Elle s'applique également à l'évaluation de politiques particulières (par domaine) à l'échelon national, régional et/ou local. Grâce à des indicateurs quantitatifs et qualitatifs des résultats, il est possible, en la mesurant à l'aune de la meilleure pratique possible, de se former une idée claire d'une politique dans un contexte donné.

Le guide comporte cinq grands chapitres : (i) l'objectif, les définitions et le contenu d'une politique à l'égard des petites entreprises ; (ii) la méthodologie en soi : critères, champ d'application et sources d'informations ; (iii) la politique générale et son incidence sur les microentreprises et les PME ; (iv) une évaluation par domaine ; et (v) une évaluation du dispositif d'aide dans son ensemble.

Pour les besoins du guide, nous considérons qu'une politique d'aide aux petites entreprises a pour objectif ultime de renforcer le secteur, et doit donc accroître le nombre d'entreprises et améliorer la concurrence entre elles. De ce fait, la politique à l'égard des PME repose sur trois composantes : (i) un ensemble de domaines stratégiques où les mesures ont un effet direct sur les microentreprises et les PME ; (ii) le degré de coordination, de coopération et de relation entre elles et les diverses institutions concernées ; et (iii) ce que nous dénommons « politiques générales » et qui, bien qu'elles n'aient pas pour objet de modifier le comportement des microentreprises et PME, l'influencent concrètement.

A cet égard, étant donné les lacunes patentées dans le domaine de l'évaluation, la méthodologie est structurée « de bas en haut ». Elle commence donc par une analyse et une évaluation générales des composantes de la politique d'aide, à savoir les programmes, les domaines auxquels ils sont associés, les personnes et les organismes participants, leur cohésion interne, le degré de coordination entre eux, et enfin leur fonctionnement global. Elle aborde aussi brièvement d'autres mesures sectorielles publiques (les mesures générales et leur impact sur les microentreprises et les PME). S'agissant des domaines concernés, on suppose que les plus problématiques sont le financement, la technologie et l'information (y compris les centres d'aide aux entreprises), la commercialisation et le commerce extérieur, le cadre juridique, réglementaire et concurrentiel, les ressources humaines, l'éducation et la formation, et la coopération entre les microentreprises et les PME.

Le Guide prend en considération trois critères d'évaluation : **final** (qui tient compte du bien-être économique de la société dans son ensemble) ; **intermédiaire** (plus simple et orienté sur l'incidence des politiques sur les résultats des microentreprises et des PME) et **institutionnel** (qui mesure la qualité des institutions participant à l'application des mesures d'aide). Les critères couvrent donc l'économie en général, les bénéficiaires d'un service donné et l'institution qui fournit ce service. On trouvera ci-après un exemple de critères dans le domaine de la technologie.

⁺ Ici comme dans le Guide, les termes « microentreprises et PME » ou « petites entreprises » sont utilisés de manière synonyme pour définir le groupe des petites, moyennes et micro-entreprises ayant un potentiel de développement.

TABLEAU A. 2. CRITÈRES APPLICABLES A L'ANALYSE DES DOMAINES STRATÉGIQUES

Domaine Critère	TECHNOLOGIE
CRITERE INTERMEDIAIRE	Le service (programme ou autre type d'assistance) a-t-il couvert un nombre satisfaisant (par rapport à la cible) de petites entreprises ? Par exemple, quel est le pourcentage des petites entreprises examinées dont les dirigeants ont visité un centre de développement technologique ?
	Le service a-t-il répondu aux besoins pour lesquels il a été conçu ? Par exemple, quel est le pourcentage des entreprises dont les dirigeants qui ont visité le centre ont jugé la visite utile ?
	Le service ou l'aide ont-ils eu l'effet favorable souhaité sur la croissance, la productivité, l'emploi, les bénéfices ou d'autres objectifs de l'entreprises (ou groupe d'entreprises) ? Ainsi, est-il prouvé que les entreprises qui font appel aux centres de développement technologique augmentent plus rapidement leur productivité que d'autres petites entreprises comparables ?
CRITERE FINAL	Le train de mesures (ou le dispositif d'aide global) a-t-il eu l'effet escompté sur la croissance générale du produit, de la productivité et de l'emploi dans les petites entreprises, et sur la distribution des revenus dans le pays ? A titre d'exemple, est-il statistiquement démontrable que la présence des centres d'aide aux entreprises a des effets tangibles en termes de croissance et de création d'emplois dans le secteur des petites entreprises ?
CRITERE INSTITUTIONNEL	L'engagement des pouvoirs publics envers l'élaboration du système de soutien technologique aux microentreprises et aux PME et son efficacité, qui se traduit concrètement par des moyens économiques, des ressources humaines qualifiées, des cibles et des objectifs clairs et l'évaluation des résultats.
	L'appui du secteur non gouvernemental (organismes intermédiaires et secteur privé) sous forme de coopération à l'élaboration, l'exécution et/ou le suivi des opérations d'aide du secteur public, au travers de cofinancements, de la prestation directe de services d'assistance technique, etc.
	La qualité des institutions qui participent à la prestation des services, qui transparait dans les qualités et l'expérience professionnelles, par exemple, des membres des conseils scientifiques et technologiques nationaux, dans le degré et le mode de décentralisation des activités, etc.
	Le choix judicieux des instruments : les technologies sont-elles adaptées aux besoins spécifiques du secteur des microentreprises et des PME du pays ou de la région ? L'équilibre dans ce domaine est-il adéquat ?

Selon les domaines à couvrir et les critères de décision, il est possible de recueillir des informations pertinentes sur les résultats des politiques et des programmes en exploitant les informations quantitatives et qualitatives disponibles (celles que réunissent par exemple les organismes d'aide) et en créant de nouvelles sources selon les besoins. Le Guide comporte un questionnaire qui a pour objet d'aider à obtenir les informations auprès des clients, des services d'aide et des spécialistes du sujet. Le tableau qui suit permet de planifier les types d'indicateurs qu'il convient de recueillir dans chaque domaine.

TABLEAU A. 3. INDICATEURS A RECUEILLIR DANS CHAQUE DOMAINE

Domaine Variables	<i>Technologie</i>
Objectif(s) du domaine	<ul style="list-style-type: none"> • Un ou plusieurs • Concrets ou vagues (réalisables ?) • Complémentaires ou incompatibles ?
Institution(s) responsable(s)	<ul style="list-style-type: none"> • Une ou plusieurs • S'il y en a plusieurs, travaillent-elles en coordination ?
Description du domaine	<ul style="list-style-type: none"> • Brève description des politiques, programmes et/ou interventions concernés (sont-ils redondants, complémentaires, etc. ?). Ressources humaines et économiques. Degré de coordination avec les entreprises privées, les pouvoirs locaux, etc.
Dispositifs de diffusion du programme	<ul style="list-style-type: none"> • Presse (journaux, magazines) • Audiovisuel (radio, télévision, Internet) • Brochures, dépliants, documents des institutions, etc.. • Chambres de commerce, guildes locales, universités, écoles techniques • Divers
Mécanismes opérationnels	<ul style="list-style-type: none"> • Publics • Privés • Mixtes
Mécanismes d'évaluation	<ul style="list-style-type: none"> • Évaluation des domaines/de programmes spécifiques • Variables prises en compte (couverture, satisfaction, gains de productivité, économies, etc.) • Suivi après un délai donné
Coût par domaine	<ul style="list-style-type: none"> • Coût total des interventions dans un domaine (en monnaie locale et en USD) • Coût par programme (en monnaie locale et en USD) • Ressources humaines et pourcentage des coûts qu'elles représentent dans les interventions dans un domaine - peuvent être ventilées par programme. • Les programmes sont-ils cofinancés ? (dans quelle proportion et par qui ?)
Degré de coordination avec d'autres domaines	<ul style="list-style-type: none"> • Y a-t-il une coordination avec d'autres domaines dans le cadre d'une politique d'aide intégrée ? • Coordination avec le secteur privé / des organisations internationales sous forme de : <ul style="list-style-type: none"> ○ Sous-traitance ○ Services de conseil ○ Projets communs ○ Divers
Bénéficiaires (critères d'admissibilité)	<p>En théorie :</p> <ul style="list-style-type: none"> • Tous les types d'entreprises • Seulement les PME • Les microentreprises • Autres <p>Dans la pratique, qui sont les bénéficiaires ?</p>
Label de qualité	<p>Existe-t-il des labels de qualité en ce qui concerne les programmes portant sur des domaines</p> <ul style="list-style-type: none"> • à l'échelon local • à l'échelon international

Ces informations, complétées d'une brève description des microentreprises et des PME dans la région et d'un point sur les politiques générales, forment la base de l'évaluation et de la production du rapport. L'évaluation peut être plus ou moins raffinée, selon les besoins des usagers, les informations et les ressources économiques et humaines disponibles.

Enfin, le Guide à l'évaluation des politiques d'aide n'est pas un instrument statique. Il doit être constamment actualisé à mesure que le dispositif d'aide s'adapte aux réalités nouvelles et aux besoins des entreprises. Il s'agit donc d'un processus continu d'adaptation et d'amélioration, tant dans la structure générale que dans les domaines particuliers de l'évaluation et des instruments spécifiques à chacun. Avec l'amélioration de la qualité des informations, sur le plan général et pour les interventions spécifiques, et celui du niveau de qualification des ressources humaines, de la disponibilité des ressources économiques, etc., la méthodologie sera constamment perfectionnée.

BIBLIOGRAPHIE

- Angrist, Joshua D. et Alan B. Krueger (2001), « Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments » *Journal of Economic Perspectives*, 15(4), pp. 69 à 85.
- CEAP (2002), « Evaluation Framework for Economic and Technical co-operation work across APEC Fora », Document présenté par l'Australie au Groupe restreint sur l'évaluation de l'ESC.
- Bondonio, Daniele et John Engberg, (2000) « Enterprise Zones and Local Employment: Evidence from the States' Programs. » *Regional Science and Urban Economics* 30: 519–549.
- Bager-Sjogren, L. (2003), « On the Job Training and Small Business competitiveness – EU Structural funds Impact in Sweden », Swedish Institute of Growth Policy Studies (ITPS), présenté à la Conférence de l'ICSB, Belfast 16 – 18 juin 2003.
- Bartik, T., (2002) « Evaluating the Impacts of Local Economic Development Policies on Local Economic Outcomes: What has been done and what is doable? », communication à la Conférence sur l'évaluation du développement économique et la création d'emploi au niveau local, OCDE, Vienne, 20-21 novembre 2002.
- Fraser, S., Storey, D.J. et Westhead, P., (2003), « Small firm work placements for Undergraduates: Does it pay off? », Université de Warwick, SME Centre.
- Greenbaum, Robert T, (1998), « An Evaluation of State Enterprise Zone Policies: Measuring the Impact on Business Decisions and Housing Market Outcomes. » Thèse doctorale, Université Carnegie Mellon, Décembre.
- Harrison, R.T et Leitch, C.M. (1994), « Whatever you hit call the target: an alternative approach to small business policy » in Danson, M (ed.) « Small Firm Formation and Regional Economic Development », Routledge Series in Small Business, Londres.
- Heckman, James J., Hidehiko Ichimura, et Petra Todd, (1997), « Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. » *Review of Economic Studies*, 64(4), pp. 605 à 654.
- Holzer, H. J., R. N. Block, M. Cheatham et J. H. Knott, (1993), « Art Training Subsidies For Firms Effective? The Michigan Experience. », *Industrial and Labor Relations Review*, 46, (juillet) pp. 625 à 636.
- Jarmin, Ronald S. (1999) « Evaluating the Impact of Manufacturing Extension on Productivity Growth. » *Journal of Policy Analysis and Management*, 18(1), pp 99 à 119.
- Kuusisto, J., Berney, R. et Blackburn, R (1999) « A Critical Evaluation of SME Support Policies in the United Kingdom and the Republic of Ireland - An in depth Delphi Study of Selected SME Support

Policies and their Evaluation », Ministère du Commerce et de l'Industrie, Finlande, Studies and Reports 6/1999.

Lerner, J. (1999), « The Government as Venture Capitalist: The Long Run Impact of the SPIR Program », *Journal of Business*, Vol. 72, No3, pp 285-318.

Meyer Bruce D, (1995), « Natural and Quasi-Experiments in Economics », *Journal of Business and Economic Statistics*, 13(2), pp. 151 à 161.

Murnane, R. J., S. Newstead, et R. J. Olsen, (1985) « Comparing Public and Private Schools: The Puzzling Role of Selectivity Bias. » *Journal of Business and Economic Statistics*, 3, pp 23 à 35.

Oldsman, E., (2002), « Evaluating Business Assistance Programme », communication à la Conférence sur l'évaluation du développement économique et la création d'emploi au niveau local, OCDE, Vienne, 20-21 novembre.

OCDE (1997) *Policy Evaluation in Innovation and Technology: Towards best Practices*, OCDE (Paris).

OCDE (2000) *Perspectives de l'OCDE sur les PME*, OCDE (Paris).

OCDE (2003) *La cohérence des politiques : un impératif pour le développement mondial*, Synthèses, septembre, OCDE, Paris.

Papaconstantinou, G et Polt, W. (1997), « Policy Evaluation in Innovation and Technology: An Overview », in « *Policy Evaluation in Innovation and Technology: Towards Best Practices* », OCDE (Paris).

Roper S et Hewitt-Dundas, N. (2001), « Grant Assistance and Small Firm Development in Northern Ireland and the Republic of Ireland », *Scottish Journal of Political Economy*, Vol. 48 n° 1, pp. 99 à 117.

Smith, Jeffrey A. et Petra E. Todd, (2001), « Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods. » *The American Economic Review*, Vol. 91(2), pp 112 à 118.

Storey, D. J. (2000) « Six Steps to Heaven: evaluating the impact of public policies to support small businesses in developed economies ». in Landstrom H and Sexton D.L. (eds.) « *Handbook of Entrepreneurship* », Blackwells, Oxford, pp 176 à 194.

Thomas, T et Landry B (2001) « Évaluation des répercussions des politiques de l'APECA - Programmes fédéraux de développement économique du Canada atlantique », ACOA, Moncton, New Brunswick.