
NOTES D'ÉTUDES

ET DE RECHERCHE

LES MÉTHODES

MICRO-ÉCONOMÉTRIQUES D'ÉVALUATION

Denis Fougère

Mars 2007

NER - E # 166



LES MÉTHODES
MICRO-ÉCONOMÉTRIQUES D'ÉVALUATION

Denis Fougère

Mars 2007

NER - E # 166

Les Notes d'Études et de Recherche reflètent les idées personnelles de leurs auteurs et n'expriment pas nécessairement la position de la Banque de France. Ce document est disponible sur le site internet de la Banque de France « www.banque-france.fr ».

Working Papers reflect the opinions of the authors and do not necessarily express the views of the Banque de France. This document is available on the Banque de France Website "www.banque-france.fr".

Les méthodes micro-économétriques d'évaluation

Denis Fougère¹

*CNRS, CREST-INSEE, CEPR, IZA
et Banque de France (DGEI-DIR)*

¹ Adresse : CREST-INSEE, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France. Courriel : fougere@ensae.fr

Résumé : Nous passons ici en revue les développements récents dans le domaine des méthodes micro-économétriques d'évaluation. Dans cette littérature, le modèle statistique de référence est le modèle causal de Rubin (modèle voisin du modèle économétrique d'auto-sélection de Roy), dont nous rappellerons la forme et les hypothèses. Ce modèle est un cadre adapté pour définir et étudier les conditions d'identification de l'effet, par exemple, du passage par un dispositif de la politique active de l'emploi. Nous insistons également sur la définition et l'identification des différents paramètres pertinents pour l'évaluation, tels que les effets moyens et marginaux du dispositif sur l'emploi ou les salaires des bénéficiaires. Nous explicitons en particulier les hypothèses qui permettent d'identifier l'effet moyen ou marginal du dispositif dans la situation virtuelle (dite contrefactuelle) où les bénéficiaires du dispositif ne pourraient y accéder. A chaque type d'hypothèses (sélection sur variables inobservables, indépendance conditionnelle des variables de résultats et de l'entrée en dispositif, etc.) correspond généralement un estimateur particulier. Nous mettons plus particulièrement l'accent sur les estimateurs par appariement pour le modèle de sélection sur variables observables, ainsi que sur les méthodes des doubles différences et de régression avec discontinuité pour le modèle de sélection sur variables observables et inobservables.

Mots clés : évaluation, biais de sélectivité, estimateurs par appariement, méthode des doubles différences, régression avec discontinuité.

Codes JEL : C13, C14, C41, J64.

Abstract : Our survey covers the recent developments of the microeconomic literature on evaluation methods. In this field, the canonical model is Rubin's causal model, which is close to Roy's selectivity model. This model is the relevant framework for defining and for examining the identifiability conditions of the parameters of interest in any evaluation study. We insist on the definition of these parameters, which include the average effect of the treatment on the treated and on the non-treated individuals. For each set of assumptions (selectivity on observable or unobservable characteristics, conditional independence between outcomes and treatment indicators, etc.), we present the most adapted estimation method. We put a special emphasis on matching estimators in the situation where the selectivity depends only on observables, and on differences-in-differences methods and on regression-discontinuity techniques when the selectivity depends both on observable and unobservable characteristics.

Keywords : evaluation, selectivity bias, matching estimators, differences-in-differences, regression-discontinuity.

JEL Codes : C13, C14, C41, J64.

Résumé non technique : Comment évaluer l'effet d'une intervention publique ou d'une réforme ? Comment cette intervention a-t-elle modifié les revenus des ménages, la rentabilité des entreprises, ou encore le niveau d'emploi des économies régionales ou nationales concernées ? Comment cette intervention aurait-elle modifié les résultats des ménages, entreprises, économies qui n'ont pu en bénéficier ? Durant les dernières années, ces questions ont été au centre d'une littérature économétrique riche en avancées méthodologiques. Les méthodes statistiques disponibles sont certes nombreuses, mais leurs spécificités et leurs hypothèses conditionnent fortement les résultats. Ces méthodes sont en général basées sur une comparaison entre des individus ou entreprises bénéficiant de la réforme ou de l'intervention publique que l'on souhaite évaluer et des individus ou entreprises n'en bénéficiant pas. Au cœur de la démarche statistique d'évaluation, est de ce fait souvent posée la question des biais de sélectivité : le recours à une subvention, à un crédit d'impôt ou à un dispositif public d'aide ou d'accompagnement est une décision qui met en œuvre un choix individuel, relevant au moins en partie du domaine de la rationalité ; en particulier, cette décision ne peut être indépendante de la façon dont l'agent économique évalue par lui-même les conséquences de son choix. Si l'on ne tient pas compte de ces biais de sélectivité, l'on risque de produire des estimations biaisées des effets de l'intervention publique en comparant directement les situations des deux groupes, bénéficiaires et non bénéficiaires de cette intervention.

Pour limiter les conséquences des biais de sélectivité, les analystes construisent généralement un groupe de contrôle dont les caractéristiques observables se rapprochent le plus possible de celles des agents bénéficiaires de l'intervention publique. Cette démarche n'est valide que dans le cas où la sélectivité opère seulement sur la base de caractéristiques observables (telles que l'âge des personnes ou l'effectif des entreprises). Elle est toutefois jugée trop restrictive par les statisticiens et économètres, qui lui préfèrent l'hypothèse d'une sélectivité provenant à la fois des caractéristiques individuelles observables et inobservables.

Dans cette synthèse, nous présentons tout d'abord le modèle canonique de l'évaluation, introduit par Rubin (1974). Ce modèle permet de définir clairement tout à la fois l'effet causal de la politique que l'on souhaite évaluer et la nature du biais de sélectivité. Les deux caractéristiques importantes de l'effet causal, tel qu'il est défini dans ce cadre, sont, d'une part, son hétérogénéité dans la population, et, d'autre part, son caractère inobservable. Il faut

comprendre que c'est cette non-observabilité qui ne permet pas d'identifier librement la distribution de l'effet causal dans la population.

Nous commençons par examiner la situation dans laquelle l'éligibilité à l'intervention évaluée dépend uniquement des caractéristiques individuelles des agents observables par l'analyste. Cette situation a été particulièrement examinée par les économètres. Elle les a conduits à proposer de nouveaux estimateurs, en particulier les estimateurs par appariement, introduits à l'origine par Rubin (1977) et développés récemment par Heckman, Ichimura et Todd (1998). Nous détaillons les étapes de la mise en œuvre pratique de ces estimateurs et nous présentons ensuite plusieurs études utilisant cette procédure.

Nous examinons ensuite la situation dans laquelle la sélectivité dépend tout à la fois des caractéristiques observables et inobservables des agents. Dans ce contexte nous mettons l'accent sur les méthodes les plus souvent utilisées, à savoir la méthode des doubles différences (*differences in differences*), la régression avec discontinuité (*regression discontinuity*), mais aussi le modèle Tobit de sélection. Plusieurs exemples d'application permettent une fois encore d'illustrer ces méthodes.

Nous concluons cette synthèse en soulignant l'utilité de chacune de ces méthodes mais aussi en insistant sur leurs limites. Tirant parti des travaux de Heckman et al. (1998), basés sur des échantillons de contrôle expérimentaux et non expérimentaux, nous tentons de dresser une carte des différentes situations possibles et des estimateurs à mettre en œuvre dans chacun des cas.

1. Introduction

Comment évaluer l'effet d'une intervention publique ou d'une réforme ? Comment cette intervention a-t-elle modifié les revenus des ménages, la profitabilité des entreprises, ou encore le niveau d'emploi des économies régionales ou nationales concernées ? Comment cette intervention aurait-elle modifié les résultats des ménages, entreprises, économies qui n'ont pu en bénéficier ? Durant les dernières années, ces questions ont été au centre d'une littérature économétrique riche en avancées méthodologiques.² Les méthodes statistiques disponibles sont certes nombreuses, mais leurs spécificités et leurs hypothèses conditionnent fortement les résultats.³ Ces méthodes sont en général basées sur une comparaison entre des individus ou entreprises bénéficiant de la réforme ou de l'intervention publique que l'on souhaite évaluer et des individus ou entreprises n'en bénéficiant pas. Au cœur de la démarche statistique d'évaluation, est de ce fait souvent posée la question des biais de sélectivité : le recours à une subvention, à un crédit d'impôt ou à un dispositif public d'aide ou d'accompagnement est une décision qui met en œuvre un choix individuel, relevant au moins en partie du domaine de la rationalité ; en particulier, cette décision ne peut être indépendante de la façon dont l'agent économique évalue par lui-même les conséquences de son choix. Si l'on ne tient pas compte de ces biais de sélectivité, l'on risque de produire des estimations biaisées des effets de l'intervention publique en comparant directement les situations des deux groupes, bénéficiaires et non bénéficiaires de cette intervention.

Pour limiter les conséquences des biais de sélectivité, les analystes construisent généralement un groupe de contrôle dont les caractéristiques observables se rapprochent le plus possible de celles des agents bénéficiaires de l'intervention publique. Cette démarche n'est valide que dans le cas où la sélectivité opère seulement sur la base de caractéristiques observables (telles que l'âge des personnes ou l'effectif des entreprises). Elle est toutefois jugée trop restrictive par les statisticiens et économètres, qui lui préfèrent l'hypothèse d'une sélectivité provenant à la fois des caractéristiques individuelles observables et inobservables.⁴

² Voir Heckman, Smith et Lalonde (1999) pour une synthèse très complète.

³ Sur la sensibilité des résultats de l'évaluation aux méthodes, le lecteur pourra consulter, entre autres exemples, les contributions de Lalonde (1986), Heckman et Hotz (1989), Heckman et Robb (1985), Heckman, Tobias et Vytlačyl (2000).

⁴ Les travaux français d'évaluation que l'on pourrait appeler « de première génération » reposent sur ce second type d'hypothèse. Ces travaux, recensés dans la synthèse rédigée par Fougère, Kramarz et Magnac (2000), sont fondés sur des modèles de transition en temps continu (Bonnal, Fougère et Sérandon, 1997), ou en temps discret (Magnac, 2000, Kramarz et Philippon, 2001). Ces études ont examiné les effets de deux types de politiques actives, d'une part les stages et dispositifs d'insertion des jeunes sur le marché du travail, et d'autre part, les

Il est important à ce stade de bien comprendre la différence entre l'évaluation *ex ante*, reposant sur la simulation des réformes possibles à partir d'un modèle économétrique structurel estimé avant la mise en place des réformes, et l'évaluation *ex post*, proposant une évaluation réalisée à l'aide de données expérimentales ou non, mais collectées après mise en place des réformes. Ces deux approches sont toutefois complémentaires comme l'ont montré notamment Ichimura et Taber (2002), Todd et Wolpin (2005), ou bien encore Lise, Seitz et Smith (2005), en utilisant les résultats de méthodes économétriques d'évaluation *ex post* pour élaborer et calibrer des exercices d'évaluation *ex ante*.

Le cadre statistique général adapté à la démarche évaluative a été à l'origine développé par les statisticiens dans un domaine très éloigné de celui de l'évaluation des politiques économiques, à savoir celui de l'évaluation des traitements dans le domaine biomédical. Mais, dans ce champ, les statisticiens ont proposé des outils, concepts et méthodes, dont le domaine d'application est bien plus large que celui des essais cliniques.⁵ En particulier, ces concepts et méthodes ont été fréquemment utilisés en micro-économétrie pour l'évaluation des programmes de formation et d'aide sociale. Mais, à n'en pas douter, leur champ d'application potentiel en économétrie est bien plus large. De ce point de vue, notre synthèse ne se propose pas de recenser l'ensemble des applications aujourd'hui réalisées, mais plutôt d'illustrer les méthodes passées en revue à l'aide d'exemples.

Dans la section 2, nous présentons le modèle canonique de l'évaluation, introduit par Rubin (1974). Le succès de ce modèle provient du fait qu'il permet de définir clairement tout à la fois l'effet causal de la politique que l'on souhaite évaluer et la nature du biais de sélectivité. Les deux caractéristiques importantes de l'effet causal, tel qu'il est défini dans ce cadre, sont, d'une part, son hétérogénéité dans la population, et, d'autre part, son caractère inobservable. Il faut comprendre que c'est cette non-observabilité qui ne permet pas d'identifier librement la distribution de l'effet causal dans la population.

allègements de charges patronales sur l'emploi des moins qualifiés. Les conclusions de ces études corroborent les résultats des travaux conduits dans d'autres pays : les programmes de formation et de stages plus particulièrement destinés aux jeunes chômeurs peu qualifiés n'ont en général pas d'effets à court terme sur leurs salaires et leur employabilité, à moins que ces stages et formations aient un contenu suffisant en formation. Toutefois, les politiques d'allègement des charges patronales ont des effets substantiels sur l'emploi notamment des travailleurs dont le salaire est proche du salaire minimum.

⁵ Un des concepts centraux dans cette littérature est celui de causalité statistique, ou plus exactement d'effet causal d'un traitement (sur ce point, voir Holland, 1989).

Dans la section 3, nous considérons la situation dans laquelle la sélection des agents éligibles à l'intervention évaluée dépend uniquement de caractéristiques individuelles observables par l'analyste. Cette situation a été particulièrement examinée par les économètres. Elle les a conduits à proposer de nouveaux estimateurs, en particulier les estimateurs par appariement, introduits à l'origine par Rubin (1977) et développés plus récemment par Heckman, Ichimura et Todd (1998). Ces estimateurs sont présentés dans la section 4. Dans la section 5, nous détaillons les étapes de la mise en œuvre pratique de l'estimation par appariement et nous présentons plusieurs études utilisant cette procédure.

Les sections 6 et 7 sont respectivement consacrées à la méthode des doubles différences (*differences in differences*) et à la régression avec discontinuité (*regression discontinuity*). Dans la section 8, nous examinons la situation dans laquelle la sélectivité dépend tout à la fois des caractéristiques observables et inobservables des agents. Dans ce cadre élargi, nous définissons d'abord les paramètres d'intérêt, puis nous rappelons la forme du modèle Tobit de sélection, dans lequel la loi des caractéristiques inobservables est supposée normale. Plusieurs exemples d'application permettent une fois encore d'illustrer la méthode. Nous présentons ensuite certaines extensions possibles du modèle Tobit à des lois paramétriques plus générales que la loi normale.

Dans la section 9, nous concluons cette synthèse en soulignant l'utilité de chacune de ces méthodes mais aussi en insistant sur leurs limites. Tirant parti des travaux de Heckman et al. (1998), basés sur des échantillons de contrôle expérimentaux et non expérimentaux, nous tentons de dresser une carte des différentes situations possibles et des estimateurs à mettre en œuvre dans chacun des cas.

2. Le modèle générique pour l'évaluation

2.1 Définitions et notations

Le modèle canonique de l'évaluation a été introduit par Rubin en 1974. Ce modèle, assez général, est adapté à la situation dans laquelle un *traitement* peut être administré ou non à un individu. Le terme de traitement se réfère aux premiers travaux ayant permis de développer ce cadre conceptuel, travaux qui concernaient l'évaluation de l'efficacité des traitements dans le domaine médical. Bien qu'il ne soit pas le plus approprié, il est utilisé en économétrie pour

qualifier une intervention publique, une réforme fiscale, une politique de subvention, un programme de formation, ou bien un programme d'aide sociale que l'on cherche à évaluer.

De manière formelle, nous supposons que, pour chacun des individus i d'un échantillon de taille N , nous observons l'ensemble suivant de variables aléatoires. L'accès au programme (i.e. au traitement) est représenté par une variable aléatoire T , qui prend la valeur 1 si l'individu accède au programme, 0 sinon. L'efficacité du programme est mesurée au travers d'une variable de résultat, notée Y . En fait, le modèle de Rubin repose sur l'existence de deux variables latentes de résultat, notées Y_1 et Y_0 , selon que l'individu reçoit le traitement ($T=1$) ou non ($T=0$). Ces variables correspondent aux *résultats potentiels* (*potential outcomes*) du programme. Elles ne sont jamais simultanément observées à la même date pour un même individu. Ainsi, pour un individu traité, Y_1 est observée tandis que Y_0 est inconnue. Dans ce cas, la variable Y_0 correspond au résultat qui aurait été réalisé si l'individu n'avait pas été traité. On dit aussi que la variable Y_0 représente le résultat *contrefactuel*. Pour un individu non traité, on observe au contraire Y_0 , tandis que Y_1 est inconnue. La variable de résultat observée peut donc se déduire des variables potentielles et de la variable de traitement par la relation :

$$Y = T Y_1 + (1-T)Y_0 \quad (1)$$

Seul le couple (Y, T) est observé pour chaque individu.

2.2 Paramètres d'intérêt

L'*effet causal* du traitement est défini pour chaque individu par l'écart $\Delta = Y_1 - Y_0$ qui représente la différence entre ce que serait la situation de l'individu s'il était traité et ce qu'elle serait s'il ne l'était pas. L'effet causal a ainsi deux caractéristiques importantes :

- il est *inobservable*, puisque seule une des deux variables potentielles est observée pour chaque individu,
- il est *individuel*, et de ce fait il existe une distribution de l'effet causal dans la population.

La distribution de l'effet causal n'est toutefois pas identifiable, tout simplement parce que l'effet causal est inobservable. Néanmoins, grâce à des hypothèses sur la loi jointe du triplet

(Y_0, Y_1, T) , on peut identifier certains paramètres de la distribution de l'effet causal à partir de la densité des variables observables (Y, T) . Deux paramètres font généralement l'objet d'un examen spécifique. Il s'agit de *l'effet moyen du traitement dans la population* :

$$\Delta^{ATE} = E(Y_1 - Y_0) \quad (2)$$

et de *l'effet moyen du traitement dans la population des individus traités* :

$$\Delta^{TT} = E(Y_1 - Y_0 | T=1) \quad (3)$$

Ces deux paramètres ne sont égaux que sous certaines hypothèses très restrictives. D'une façon générale, les conditions nécessaires à l'identification du premier paramètre sont plus exigeantes que celles nécessaires à l'identification du second. En effet, pour ce dernier, les hypothèses ne portent que sur la loi de Y_0 et T . En particulier, si les variables de résultat sont indépendantes de la variable d'accès au traitement, c'est-à-dire si $(Y_0, Y_1) \perp\!\!\!\perp T$, il est possible d'identifier les deux paramètres d'intérêt Δ^{ATE} et Δ^{TT} préalablement définis. En effet, si cette condition (suffisante) est satisfaite, ces deux paramètres d'intérêt deviennent :

$$\Delta^{ATE} = E(Y_1) - E(Y_0) = E(Y_1 | T=1) - E(Y_0 | T=0) = E(Y | T=1) - E(Y | T=0) \quad (4)$$

et

$$\Delta^{TT} = E(Y_1 | T=1) - E(Y_0 | T=1) = E(Y_1 | T=1) - E(Y_0 | T=0) = E(Y | T=1) - E(Y | T=0) \quad (5)$$

Dans ce cas, les deux paramètres sont égaux et peuvent être estimés simplement comme la différence des moyennes des variables de résultat observées dans le groupe des individus traités et dans le groupe des individus non traités. On peut remarquer que l'identification du second paramètre nécessite une hypothèse moins forte. En ce cas, il suffit en effet que les variables aléatoires Y_0 et T soient indépendantes (soit $Y_0 \perp\!\!\!\perp T$).

Dès lors que la propriété d'indépendance précédente n'est plus satisfaite, l'estimateur naturel formé par la différence des moyennes des variables de résultat est affecté d'un *biais de sélection*. En effet,

$$\begin{aligned}
E(Y|T=1) - E(Y|T=0) &= E(Y_1|T=1) - E(Y_0|T=0) \\
&= E(Y_1|T=1) - E(Y_0|T=1) + E(Y_0|T=1) - E(Y_0|T=0) \\
&= \Delta^{TT} + B^{TT}
\end{aligned}
\tag{6}$$

Le biais de sélection est le terme $B^{TT} = E(Y_0|T=1) - E(Y_0|T=0)$ dans l'expression précédente. Ce biais trouve son origine dans le fait que la situation moyenne des individus qui ont reçu le traitement n'aurait pas été la même en l'absence de traitement que celle des individus n'ayant pas reçu le traitement. Il en est ainsi parce que ces deux populations ne sont pas identiques, sauf dans le cas particulier d'une expérience contrôlée.

2. 3 Le principe des expérimentations contrôlées

L'évaluation par expérimentation contrôlée essaie de s'affranchir du biais de sélection en proposant d'affecter de manière aléatoire (par tirage au sort) les individus potentiellement intéressés par un dispositif d'aide ou de subvention, un programme de formation, etc., à un groupe de traitement qui bénéficiera de ce dispositif ou programme, ou bien à un groupe de contrôle qui n'en bénéficiera pas. Un bilan des expérimentations relatives à l'évaluation des programmes d'aide à l'emploi introduits durant les vingt dernières années est contenu dans les deux articles de synthèse rédigés par Fougère (2000a, 2000b).

Depuis le début des années 60, les autorités fédérales et locales des Etats-Unis d'Amérique ont recouru à des expériences contrôlées pour évaluer des politiques publiques dans des domaines aussi divers que la santé, le logement, l'impôt sur les revenus du travail, l'insertion et la ré-insertion sociale, les économies d'énergie ou la tarification de la consommation d'électricité. Mais c'est dans le domaine des programmes de formation et d'emploi que les expériences menées ont donné lieu aux débats les plus vifs et aux contributions méthodologiques les plus significatives.

Le fait que le tirage au sort ajoute de l'incertitude dans le processus d'accès au programme peut conduire les individus les plus adverses au risque à ne pas se porter candidats. Parallèlement, si les agences chargées de la mise en œuvre des programmes ou interventions

offrent un nombre de places limité ou si elles sont évaluées sur la base des résultats de l'expérimentation, elles peuvent être de ce fait incitées à sélectionner les candidats les plus qualifiés ou ceux pour lesquels le programme est potentiellement le plus bénéfique. Les expériences contrôlées peuvent par ailleurs être soumises à des biais de substitution, apparaissant dès lors que les membres du groupe de contrôle ont accès à des substituts proches du traitement ou du programme évalué. Dans le cas du programme américain *Job Training Partnership Act*, les enquêtes de suivi réalisées par l'administration indiquaient que 65% des membres du groupe de traitement reçurent des aides (aide à la recherche d'emploi, formation, etc.) durant les 18 mois suivant la procédure d'affectation par tirage au sort. Toutefois, des enquêtes complémentaires montrèrent que, durant la même période, 32% des membres du groupe de contrôle et 20% environ des individus éligibles ne participant pas à l'expérimentation reçurent une formation dispensée par un autre organisme. Au cours de cette expérimentation, les individus du groupe de contrôle eurent donc accès à un volume de formation supérieur à celui dont bénéficia la population éligible ne participant pas au programme.

Les conditions qui permettent de recourir à l'expérimentation et déterminent son protocole sont au nombre de trois :

1. les participants à l'expérience doivent être pleinement informés des avantages comme des risques que leur participation peut leur procurer ou leur faire courir (*principe de consentement éclairé*);
2. dans les cas où l'on anticipe que le programme peut provoquer des risques ou des torts, il faut prévoir de compenser les individus qui seront affectés au groupe de traitement; dans le cas où l'on anticipe que le programme peut être bénéfique, il faut prévoir de compenser ceux qui seront affectés au groupe de contrôle (*principe de compensation*);
3. enfin, une expérimentation n'est envisageable que dans les cas où l'on ignore a priori l'ampleur de l'effet du traitement; si le traitement s'avère préjudiciable, il est préférable d'observer ses effets sur un nombre limité d'individus consentants que sur un très grand nombre d'individus contraints (*principe de précaution*).

3. Identification sous l'hypothèse d'indépendance conditionnelle à des caractéristiques observables

Lorsque l'on souhaite évaluer des interventions publiques, par exemple des programmes de formation ou des politiques de subvention, à l'aide de données d'observation (non expérimentales), on fait face à deux populations, les bénéficiaires et les non – bénéficiaires, qui diffèrent de par la distribution des caractéristiques individuelles observables qui affectent vraisemblablement la participation à ces programmes. En pratique donc, l'indépendance (inconditionnelle) entre les variables latentes de résultat (Y_0, Y_1) et l'affectation au traitement T est une hypothèse très improbable. Une condition moins restrictive consiste à considérer qu'il existe un ensemble de variables observables X conditionnellement auquel la propriété d'indépendance entre les résultats latents et l'affectation au traitement est vérifiée.

Une façon alternative de formuler cette restriction est de considérer que conditionnellement aux variables observables individuelles X , on se situe dans le cadre d'une expérience contrôlée, c'est à dire avec une affectation aléatoire au traitement. La « randomisation » serait maintenant assurée par les covariables X plutôt que par un processus aléatoire d'affectation. Ainsi, la loi de distribution de chacun des résultats potentiels est identifiable, puisque leur distribution conditionnelle aux variables observables X l'est.

Proposition : Si $(Y_0, Y_1) \perp\!\!\!\perp T|X$, alors les distributions marginales $l(Y_0)$, $l(Y_1)$, et conditionnelles $l(Y_0|T=0)$, $l(Y_1|T=1)$ des variables de résultat sont identifiables.

En effet, sous l'hypothèse d'indépendance conditionnelle aux observables X , on a :

$$l(Y_k|X) = l(Y_k|X, T=k) = l(Y|X, T=k) \text{ pour } k=0, 1. \quad (7)$$

Comme la distribution des variables observables X est identifiable, il en résulte que les quatre distributions $l(Y_0)$, $l(Y_1)$, $l(Y_0|T=0)$ et $l(Y_1|T=1)$ sont identifiables. Il est alors possible d'identifier l'ensemble des paramètres de chacune de ces quatre distributions. On peut en particulier identifier leur espérance et donc l'effet moyen du traitement, de même que l'effet moyen du traitement pour le groupe des individus traités.

Comme précédemment, la condition d'identification pour ce dernier paramètre est moins forte, puisqu'elle ne nécessite que l'indépendance entre le résultat potentiel en l'absence de traitement et le traitement, soit $Y_0 \perp\!\!\!\perp T | X$. En effet, pour identifier l'effet moyen du

traitement sur les traités Δ^{TT} , il est suffisant d'identifier la loi conditionnelle $l(Y_1|T=1)$, ce qui est directement possible à partir des observations de l'échantillon des individus traités, et la loi conditionnelle $l(Y_0|T=1)$, ce qui par contre requiert la propriété d'indépendance conditionnelle.

Remarquons enfin que, dans un cas comme dans l'autre, la distribution des variables latentes est identifiée, mais leur loi jointe ne l'est pas. En particulier, les moments d'ordre supérieur à 1 de la loi de la différence $Y_1 - Y_0$ (en particulier sa variance) ne sont pas identifiés.

Le principe de l'estimation est d'utiliser les informations dont on dispose sur les individus non traités pour construire pour chaque individu traité un contre factuel, c'est à dire une estimation de ce qu'aurait été sa situation s'il n'avait pas été traité. Considérons par exemple l'effet causal du traitement sur les traités :

$$\begin{aligned}\Delta^{TT} &= E(Y_1 - Y_0 | T=1) = E(Y - Y_0 | T=1) \\ &= E[Y - E(Y_0|X, T=1) | T=1] = E[Y - E(Y_0|X, T=0) | T=1] \\ &= E[Y - E(Y|X, T=0) | T=1]\end{aligned}$$

Le problème est donc d'estimer pour chaque individu traité de caractéristiques x_i la quantité $E(Y | X = x_i, T=0) = g(x_i)$. L'estimateur final est alors obtenu comme la moyenne des écarts de la situation des individus traités et du contre factuel construit. Sa formule est donc :

$$\hat{\Delta}^{TT} = \frac{1}{N_1} \sum_{i \in I_1} \{y_i - \hat{g}(x_i)\} \quad (8)$$

où I_1 est le sous-échantillon des individus traités, défini par $I_1 = \{i | T_i=1\}$, N_1 est le nombre d'individus traités et $\hat{g}(x_i)$ est l'estimateur de la fonction $g(x_i) = E(Y | X = x_i, T=0)$.

4. Estimation des paramètres d'intérêt

4.1 Estimateur par appariement sur les caractéristiques observables

La méthode d'estimation par appariement a été initialement proposée par Rubin (1977). Cette méthode associe à chaque individu i traité un individu non traité, noté $\tilde{i}(i)$, dont les caractéristiques sont identiques à celles de l'individu i . Si l'on note X le vecteur des caractéristiques individuelles (supposées observables), cette méthode suppose que l'on observe, pour chaque individu traité, un individu non traité comparable, au sens des caractéristiques X , c'est-à-dire tel que $X_{\tilde{i}(i)} = X_i$. La quantité $Y_{\tilde{i}(i)}$ est un estimateur de l'espérance du résultat potentiel Y_{0i} conditionnellement au score de l'individu i :

$$Y_{\tilde{i}(i)} = \hat{E}(Y_{0i} | T_i=0, x_i) = \hat{E}(Y_{0i} | T_i=1, x_i) \quad (9)$$

L'estimateur de Rubin pour l'effet moyen du traitement sur les traités est alors égal à :

$$\hat{\Delta}_R^{TT} = \frac{1}{N_1} \sum_{i \in I_1} y_i - y_{\tilde{i}(i)} \quad (10)$$

La propriété d'indépendance conditionnelle nécessite en général la prise en compte d'un nombre important de variables de conditionnement. Pour cette raison, l'appariement peut être difficile à réaliser en pratique. En effet, pour certains individus traités, on peut n'observer aucun individu non traité ayant exactement les mêmes caractéristiques. L'estimateur proposé par Rubin consiste en fait à choisir un individu non traité aussi proche que possible de l'individu traité. En pratique, on peut se servir de la distance de Mahalanobis, et prendre pour contre factuel l'individu non traité dont l'indice vérifie :

$$\tilde{i}(i) = \arg \min_{T_j=0} \|x_i - x_j\|_{\Sigma^{-1}} \quad (11)$$

Dans cette expression, Σ est la matrice des variances-covariances des caractéristiques X dans la population des individus traités.

4.2 Estimation par appariement sur le score de propension

Le problème de dimension du vecteur X est fortement réduit par une propriété mise en évidence par Rosenbaum et Rubin (1983). La propriété d'indépendance conditionnellement à des variables observables implique celle d'indépendance conditionnellement à un résumé de

dimension un, qui est la probabilité de traitement, ou score de propension à être traité, noté $P(X) = \Pr (T = 1 | X)$.

Proposition (Rosenbaum et Rubin, 1983) : Si la variable de résultat Y_0 est indépendante de l'accès au traitement T conditionnellement aux observables X , alors elle est également indépendante de T conditionnellement au score de propension $P(X)$:

$$Y_0 \perp\!\!\!\perp T | X \Rightarrow Y_0 \perp\!\!\!\perp T | P(X)$$

En raison de cette propriété, il n'est pas nécessaire d'apparier les individus sur l'ensemble des variables de conditionnement. Il suffit de les apparier sur leur score de propension, lequel constitue un résumé unidimensionnel de l'ensemble de ces variables. L'individu non traité noté \tilde{i} , qui est apparié avec l'individu traité i , est alors défini par $P(x_{\tilde{i}}) = P(x_i)$.

Néanmoins, les propriétés asymptotiques de cet estimateur restent inconnues : l'estimateur est-il convergent ? Est-il asymptotiquement normal ? Quelle est sa vitesse de convergence ? Aucun résultat ne permet de décrire le comportement de l'estimateur lorsque le nombre d'individus traité devient grand.

4.3 Estimateur par appariement avec fonction noyau (*kernel matching*)

Les travaux d'Heckman, Ichimura et Todd (1997, 1998) ont permis de répondre à ces questions. L'idée principale est que la quantité $y_{\tilde{i}}$ est un estimateur non paramétrique de l'espérance du résultat potentiel Y_0 conditionnellement au score de l'individu i :

$$y_{\tilde{i}} = \hat{E}[Y_0 | P(x) = P(x_i)] \tag{12}$$

L'erreur quadratique de cet estimateur peut toutefois être améliorée en prenant une moyenne pondérée des observations des n voisins les plus proches de i (*n nearest neighbours matching*). D'autres estimateurs non paramétriques peuvent être envisagés. Ils consistent tous en une moyenne pondérée des observations de l'échantillon de contrôle (la population des individus non traités). Heckman et al. (1998) proposent en particulier des estimateurs à noyau, dont le plus simple pourrait s'écrire :

$$\hat{E}(Y_0|P(x)=P(x_i)) = \sum_{j \in I_0} \frac{K((P(x_j)-P(x_i))/h)}{\sum_{j \in I_0} K((P(x_j)-P(x_i))/h)} y_j \quad (13)$$

où I_0 est l'ensemble des individus non traités, défini par $I_0 = \{i|T_i = 0\}$, N_0 est le nombre d'individus non traités, K est une fonction noyau (*kernel function*), continûment différentiable, symétrique par rapport à 0, et telle que $\int_{-\infty}^{+\infty} K(u)du = 1$, et h la fenêtre d'estimation (*bandwidth function*).⁶

Chaque individu non traité participe ainsi à la construction du contre factuel de l'individu i , avec une importance qui varie selon la distance entre son score et celui de l'individu considéré. L'estimateur final de l'effet du traitement conditionnellement au fait d'être traité est ainsi :

$$\hat{\Delta}_{KM}^{TT} = \frac{1}{N_1} \sum_{i \in I_1} \left\{ y_i - \sum_{j \in I_0} \frac{K[(P(x_j)-P(x_i))/h]}{\sum_{j \in I_0} K[(P(x_j)-P(x_i))/h]} y_j \right\} \quad (14)$$

Heckman, Ichimura et Todd (1998) montrent que, sous certaines hypothèses de régularité, cet estimateur est convergent, asymptotiquement normal, avec une vitesse de convergence en racine carrée de N .

5. Mise en œuvre des procédures d'estimation par appariement d'échantillons

5.1 Le support de la distribution du score de propension

La question du support des distributions du score conditionnel au traitement est essentielle dans ce type d'analyse. Son importance a été soulignée par Heckman, Ichimura et Todd (1998) qui ont montré qu'elle constitue une source prépondérante de biais dans l'estimation de l'effet causal du traitement. Dans les méthodes d'estimation par appariement ou par

⁶ La fonction « noyau » est un opérateur qui pondère symétriquement les valeurs des observations y_j , en accordant un poids plus élevé aux observations telles que la valeur de $P(x_j)$ est « proche » de celle de $P(x_i)$. Cette distance est inversement proportionnelle à la largeur de la « fenêtre » h , qui peut être choisie de façon optimale. Les fonctions « noyau » les plus usuelles sont le noyau gaussien ; le noyau quartique ou le noyau d'Epanechnikov. Le lecteur trouvera dans Härdle et Linton (1994) une très bonne présentation des méthodes d'estimation non-paramétrique.

régression, il est nécessaire de pouvoir construire pour chaque individu traité de score p_0 un contre factuel à partir des individus non traités, c'est-à-dire de pouvoir estimer $E(Y|P(X) = p_0, T = 0)$ pour déterminer l'effet causal du traitement dans la population des individus traités.

Une estimation non paramétrique de cette quantité impose que l'on dispose pour chaque individu traité d'individus non traités dont les scores ont des valeurs proches du score de l'individu traité. En d'autres termes, la densité du score des individus non traités ne doit pas être nulle au voisinage des valeurs du score de l'individu traité que l'on considère.

On ne peut donc construire de contre factuel que pour les individus dont le score appartient à l'intersection des supports des distributions des scores des individus traités et des individus non traités. Ceci conduit à la conclusion que même sous l'hypothèse d'indépendance conditionnelle à des observables, on ne peut pas systématiquement estimer Δ^{ATE} , puisque $E(\Delta|P(X))$ ne peut être estimé que pour les individus dont le score appartient au support commun de la distribution des scores. *L'estimateur obtenu in fine est alors un estimateur local* : on est seulement en mesure d'estimer $E(\Delta|P(X) \in S_\cap, T = 1)$ ou $E(\Delta|P(X) \in S_\cap, T = 0)$, S_\cap étant le support commun défini par $S_\cap = S_T \cap S_{NT}$, S_T étant le support de la distribution du score des individus traités et S_{NT} celui des individus non traités. On présente en annexe deux exemples de distribution conditionnelle du score conduisant à des restrictions de support commun très différentes, l'une impliquant notamment de ne faire l'évaluation que sur une très faible proportion des individus traités.

Cette condition de support a une autre implication : le modèle servant à la construction du score, c'est à dire expliquant le traitement à partir des variables de conditionnement, ne doit pas être trop « bon ». Dans le cas extrême où l'on expliquerait parfaitement le traitement, les densités du score conditionnellement au traitement seraient toutes deux des masses de Dirac, l'une en zéro pour les individus non traités, l'autre en 1 pour les individus traités. Les supports seraient alors disjoints et aucun appariement ne serait possible.

Pour bien comprendre cette condition de support, il faut garder présente à l'esprit l'idée initiale de Rubin : conditionnellement à un ensemble de variables explicatives X (ou conditionnellement au score de propension), on se trouve dans le cas d'une expérience

contrôlée. La persistance de cette composante aléatoire de l'affectation au traitement conditionnellement à des observables est ainsi essentielle dans la procédure d'appariement.

5.2 Les étapes de l'estimation

Les différentes méthodes d'évaluation précédemment exposées se font toutes en plusieurs étapes. La première consiste à « expliquer » la variable T d'affectation au traitement par les caractéristiques observables X . Pour cela, une solution commode est d'estimer un modèle de type Logit (ou un modèle Logit multinomial lorsqu'il existe plusieurs catégories de traitement). Cette étape est informative dans la mesure où elle propose une description de l'affectation au traitement. Le choix des variables de conditionnement est essentiel dans cette étape, et il faut conserver à l'esprit que ce qui importe n'est pas une description aussi fidèle que possible de la probabilité du traitement, mais simplement la détermination des variables nécessaires à l'obtention de la propriété d'indépendance. Introduire un trop grand nombre de variable peut avoir des conséquences néfastes sur l'estimation à plusieurs titres. D'abord, la description de la variable de traitement étant meilleure, les supports des distributions des scores des individus traités et des individus non traités risque de se dissocier davantage ; les possibilités d'appariement seront alors plus restreintes. Mais surtout, introduire trop de variables de conditionnement peut conduire à biaiser les estimations. Il est possible en effet que la propriété d'indépendance soit satisfaite pour un ensemble de variables de conditionnement, mais qu'elle ne le soit plus lorsque l'on ajoute d'autres variables de conditionnement, quand bien même celles-ci seraient significatives dans l'estimation du score.

La deuxième étape consiste à déterminer le support commun des densités des scores des deux groupes d'individus. Pour cela, plusieurs techniques peuvent être employées :

1. on peut par exemple exclure les observations dont le score de propension estimé est proche de 1 ou de 0 ;
2. Dehejia et Wabba (1999) suggèrent de supprimer toutes les observations du groupe de contrôle pour lesquelles le score de propension estimé est inférieur au minimum des scores de propension estimés dans le groupe de traitement, et appliquer la même règle pour le maximum;
3. Heckman, Ichimura et Todd (1998) préfèrent supprimer les observations du groupe de contrôle dont les covariables ont une densité inférieure à un certain seuil. Cette

deuxième étape est importante, et l'oublier pourrait là encore biaiser les estimations. C'est en outre une importante source d'imprécision des estimations, l'estimation du contre factuel pour les individus n'appartenant pas au support commun étant biaisée et très imprécise.

Mais il faut en cet endroit vérifier une autre conséquence de l'hypothèse. L'appariement réalisé sur la base du score de propension doit en principe équilibrer la distribution des variables X dans les deux groupes, traitement et contrôle. Cette propriété du score de propension, ainsi défini comme un score d'équilibrage (*balancing score*), est un corollaire du théorème de Rosenbaum et Rubin (1983). Elle signifie que :

$$X \perp\!\!\!\perp T \mid P(X)$$

En d'autres termes, si après conditionnement par le score $P(X)$, la variable T dépend encore de X , le modèle utilisé pour estimer $P(X)$ est mal spécifié, ou bien l'hypothèse CIA ($Y_0 \perp\!\!\!\perp T \mid X$) n'est pas valide. Pour tester la validité de la propriété d'équilibrage des variables X dans les deux groupes, plusieurs procédures ont été proposées : parmi elles, citons le calcul du biais standardisé (Rosenbaum et Rubin, 1985), le t -test (Rosenbaum et Rubin, 1985), le pseudo- R^2 (Sianesi, 2004) et le test de stratification (Dehejia et Wahba, 1999, 2002)

La dernière étape est celle de l'estimation à proprement parler. Il suffit pour cela d'appliquer les formules, par exemple celle proposée par Heckman, Ichimura et Todd (1998) pour l'estimateur par appariement fondé sur les noyaux. Les résultats d'Heckman et al. (1998) permettent en outre de garantir la convergence de l'estimateur et sa normalité asymptotique, avec une vitesse de convergence en \sqrt{N} . L'écart type de l'estimateur est obtenu en appliquant les méthodes du bootstrap, ce qui consiste à répliquer l'ensemble de la procédure d'estimation sur un échantillon tiré aléatoirement avec remise dans l'échantillon initial, et à déterminer l'écart type de la distribution de l'ensemble des estimateurs ainsi obtenus. Notons que l'écart type doit aussi prendre en compte le fait que le score n'est pas connu et est de ce fait estimé. Chaque étape du bootstrap doit ainsi comprendre non seulement l'appariement sur l'échantillon tiré, mais aussi l'estimation du score. L'estimation de l'écart type peut être de ce fait coûteuse en temps de calcul.

5.3 Cas avec plusieurs traitements

Supposons maintenant qu'il existe non plus un seul, mais K traitements, et que l'accès à l'un de ces traitements soit représenté par la variable aléatoire $T \in \{1, \dots, K\}$, $K > 1$. À ces K traitements correspondent K variables de résultats Y_1, \dots, Y_K . On ajoute à ces K variables la variable de résultat associée à l'absence de traitement Y_0 . Seule l'une de ces $(K+1)$ variables de résultats potentiels est observée

Comme dans le cas avec un seul traitement, on suppose que les variables de résultat sont conditionnellement indépendantes du processus d'affectation aux traitements, étant données les valeurs prises par l'ensemble de covariables observables X :

$$(Y_0, Y_1, \dots, Y_K) \perp\!\!\!\perp T \mid X$$

Imbens (2000) et Lechner (2001) ont montré que sous cette hypothèse :

$$(Y_k, Y_{k'}) \perp\!\!\!\perp T \mid \Pi^{k/k'}(X), \forall \{k, k'\}$$

où :

$$\Pi^{k/k'}(X) = \frac{\Pi^k(X)}{\Pi^k(X) + \Pi^{k'}(X)}$$

et

$$\Pi^k(X) = \Pr(T = k \mid X)$$

Cette généralisation produit plusieurs paramètres d'intérêt qui ne sont pas symétriques :

$$E(Y_k - Y_{k'} \mid T = k, X) \neq E(Y_{k'} - Y_k \mid T = k', X)$$

L'estimation de ces paramètres est réalisée en appariant le résultat pour chaque bénéficiaire du traitement k à la moyenne pondérée des résultats des bénéficiaires du traitement k' sur la base du score conditionnel $\Pi^{k/k'}(X)$.

5.4 Quelques exemples d'application

De nombreuses études ont été consacrées à l'évaluation des programmes de formation offerts aux chômeurs. Deux études françaises récentes ont utilisé les méthodes d'estimation par appariement pour évaluer les effets de politiques actives de l'emploi. Ces deux études reposent sur l'hypothèse d'une sélection sur caractéristiques observables, et sur celle d'indépendance entre la variable de résultat et la variable d'accès au dispositif conditionnellement aux variables observables.

L'étude de Brodaty, Crépon et Fougère (2002) examine les relations entre la durée de chômage avant l'entrée dans un dispositif d'aide à l'emploi et l'effet de ce dispositif sur l'employabilité ultérieure des bénéficiaires. Elle met en évidence l'existence d'effets différenciés sur l'employabilité des chômeurs de courte et de longue durée. A cette fin, Brodaty, Crépon et Fougère (2002) développent un outil général d'évaluation qui permet de tenir compte à la fois de la multiplicité des dispositifs et de celle des dates d'entrée dans ces dispositifs. Un modèle de durée de chômage à risques concurrents sert à obtenir les scores de propension qui sont ensuite utilisés pour mettre en œuvre les estimations par appariement de l'effet des dispositifs. L'application concerne certaines mesures d'aide à l'emploi mises en place durant les quinze dernières années. L'analyse empirique exploite deux ensembles de données individuelles, qui ont été collectés au cours de deux périodes distinctes mais relativement similaires du point de vue de la conjoncture macro-économique. Il s'agit de l'enquête « Suivi des chômeurs » conduite par l'INSEE entre 1986 et 1988, et de l'enquête « Trajectoires des Demandeurs d'Emploi » réalisée par la DARES entre 1995 et 1998. Les résultats montrent que, pour ce qui est de l'employabilité des jeunes chômeurs, les dispositifs d'aide à l'emploi ont été généralement moins efficaces à la fin des années 90. Toutefois, certains dispositifs semblent avoir bénéficié plus aux jeunes chômeurs de longue durée. Cela a été particulièrement le cas des stages de formation à la fin des années 80, et des emplois aidés dans le secteur marchand dix ans plus tard.

L'étude de Crépon et Desplatz (2002) procède à une évaluation ex post des créations ou sauvegardes d'emplois induites par les dispositifs d'allègements de charges sociales sur les bas salaires introduits en France entre 1994 et 1997. L'évaluation repose sur la comparaison de l'évolution de l'emploi entre entreprises similaires du point de vue de la taille, du secteur, de l'évolution passée des effectifs, du niveau de productivité, de la rentabilité, de la structure des qualifications, de l'exposition à la concurrence étrangère, etc., mais ayant bénéficié de réductions ex ante du coût du travail différentes. Cette réduction ex ante mesure l'ampleur

avec laquelle une entreprise est concernée par les dispositifs de réductions de charges. Il s'agit de la subvention reçue par les entreprises concernées par l'extension des dispositifs, avant même tout ajustement de l'emploi réalisé par l'entreprise. Elle est basée sur la distribution des salaires observés dans les entreprises en 1994, avant la mise en place du premier dispositif d'allègement. Plus l'entreprise est intensive en main d'œuvre faiblement rémunérée, plus cette réduction ex ante est importante. L'évaluation repose sur l'hypothèse selon laquelle la situation d'une entreprise de caractéristiques observables données ne bénéficiant pas des réductions de charges aurait été la même que la situation moyenne des entreprises bénéficiaires de mêmes caractéristiques si le dispositif n'avait pas été introduit. La procédure d'appariement statistique concerne les entreprises bénéficiaires et celles qui n'ont pas bénéficié des réductions de charges. Selon cette étude, les réductions de charges auraient eu des effets très importants sur l'emploi: 460 000 emplois auraient été créés ou sauvegardés entre 1994 et 1997 (la fourchette donnée par l'intervalle de confiance à 95% est toutefois très large, puisque celui-ci va de 255 000 à 670 000 emplois).

Mais les méthodes d'appariement ont été appliquées à de nombreux autres contextes. Ainsi, Atoyán et Conway (2006) ont évalué de cette manière les programmes mis en place par le FMI de 1993 à 2003 dans 95 pays en voie de développement ou en transition. Leur échantillon contient 913 observations pays - années, 181 correspondant au groupe de traitement (les pays ayant appliqué les programmes du FMI durant ces années), 732 au groupe de contrôle (les pays n'ayant pas participé à ces programmes durant les mêmes années). La variable de résultat est le taux de croissance réel de l'économie mesuré l'année suivant la participation au programme du FMI. Les données proviennent de la base « World Economic Outlook » (FMI). Chaque observation du groupe de traitement est appariée aux observations du groupe de contrôle sur la base de la proximité des scores de propension. Atoyán et Conway (2006) trouvent que les programmes du FMI ont eu un effet positif mais statistiquement non significatif sur la croissance économique de ces pays. Mais leur échantillon est de faible taille : seules 76 observations du groupe de traitement ont pu être appariées, celles correspondant à des valeurs élevées du score de propension n'ayant pu l'être en raison de l'absence d'un support commun.

Pimenta Monteiro (2004) a utilisé l'estimation par appariement pour mesurer l'effet des privatisations sur les salaires dans le secteur bancaire au Portugal. Elle utilise pour cela des données individuelles recueillies de 1989 à 1997. Les individus "traités" sont les

salariés employés dans une banque publique qui a été privatisée au cours de la période. Le groupe de contrôle est constitué des salariés employés dans une banque restée publique tout au long de la période. L'échantillon a été restreint aux salariés n'ayant pris pas leur retraite au cours de la période. Les salaires dans les banques privatisées sont restés en moyenne plus bas durant les deux premières années, puis ont rattrapé lentement ceux des banques publiques. Les salariés des banques privatisées dont les rémunérations ont évolué le moins rapidement sont ceux les plus éduqués, les plus expérimentés, et qui étaient préalablement les mieux payés dans le secteur public.

6. L'estimateur des doubles différences

6.1 Le cadre général

La méthode des doubles différences est très souvent utilisée dans les applications économétriques. Son principe est très simple. Il se déduit directement du cadre des modèles à effets fixes individuels et temporels, utilisés en économétrie des données de panel. La forme générale du modèle est la suivante :

$$Y_{it} = \gamma T_{it} + \alpha_i + \lambda_t + \varepsilon_{it} \quad \text{avec } i = 1, \dots, N \text{ et } t = 1, \dots, T \quad (15)$$

Dans cette expression, Y_{it} représente la variable de résultat pour l'agent i à la date t , T_{it} est la variable indicatrice du traitement pour l'agent i à la date t (qui vaut 1 s'il est traité, 0 sinon), γ est un paramètre à estimer, qui représente l'effet du traitement (ici supposé *constant*), α_i est un effet fixe individuel, λ_t est un effet temporel *commun* à tous les agents. Les termes T_{it} , α_i et λ_t sont potentiellement corrélés, alors qu' ε_{it} est un aléa centré, homoscedastique, non corrélé à T_{it} , α_i et λ_t .

Seuls les agents appartenant à un groupe particulier (par exemple, une classe d'âge, les habitants d'une ville ou d'une région donnée, etc.) sont traités à partir de la date $t = \tau$. Ceci implique que :

$$T_{it} = 1 \text{ pour } i \in \mathbf{Tr} \text{ et } t \geq \tau$$

$$T_{it} = 0 \text{ pour } i \in C \text{ ou } t < \tau$$

où Tr est le groupe de traitement et C le groupe de contrôle. Le principe de la méthode est d'éliminer les effets fixes et temporels à l'aide de deux différences successives. La première différence permet d'éliminer les effets fixes :

$$\Delta Y_{it} = \gamma \Delta T_{it} + \Delta \lambda_t + \Delta \varepsilon_{it}$$

où Δ est l'opérateur de différences premières défini par :

$$\Delta Y_{it} = Y_{it} - Y_{it-1}, \Delta \lambda_t = \mu_t = \lambda_t - \lambda_{t-1}, \Delta \varepsilon_{it} = v_{it} = \varepsilon_{it} - \varepsilon_{it-1}$$

Posons maintenant $t = \tau$ et $t - 1 = \tau - 1$ (ou $t \geq \tau$ et $t - 1 \leq \tau - 1$). Si $i \in Tr$, $\Delta T_{it} = 1$, ce qui implique que $\Delta Y_{it}^{Tr} = \gamma + \mu_t^{Tr} + v_{it}^{Tr}$. Si $i \in C$, $\Delta T_{it} = 0$, ce qui implique que $\Delta Y_{it}^C = \mu_t^C + v_{it}^C$. La deuxième différence élimine les effets temporels communs :

$$\gamma = E(\Delta Y_{it}^{Tr}) - E(\Delta Y_{it}^C)$$

puisque $\mu_t^{Tr} = \mu_t^C$ et $E(v_{it}^{Tr}) = E(v_{it}^C) = 0$. L'estimateur des doubles différences est alors défini comme :

$$\hat{\gamma} = \left(\overline{\Delta Y_{it}^{Tr}} \right) - \left(\overline{\Delta Y_{it}^C} \right) \quad (16)$$

avec

$$\overline{\Delta Y_{it}^k} = \frac{1}{N_k} \sum_{i=1}^{N_k} (Y_{it}^k - Y_{it-1}^k), k \in \{Tr, C\}$$

L'extension au cas avec covariables se fait en posant :

$$Y_{it} = \gamma T_{it} + Z_{it} \delta + \alpha_i + \lambda_t + \varepsilon_{it} \text{ avec } i = 1, \dots, N \text{ et } t = 1, \dots, T \quad (17)$$

L'estimateur des doubles différences est équivalent à l'estimateur « *within* » dans le modèle projeté sur l'espace orthogonal aux effets fixes et temporels :

$$WY = WX\beta + W\varepsilon$$

où $X_{it} = [T_{it} | Z_{it}]$, $\beta = (\gamma \ \delta)'$, $W = I_{NT} - G - B_N - B_T$, $G = J_{NT} / NT$, J étant une matrice carrée composée de 1, $B_N = [(I_N \otimes J_T) / T] - G$ et $B_T = [(J_N \otimes I_T) / N] - G$. L'estimateur within du vecteur de paramètres β est alors :

$$\hat{\beta}_{\text{cov}} = (X'WX)^{-1}(X'WY) \quad (18)$$

et

$$V(\hat{\beta}_{\text{cov}}) = \sigma^2 (X'WX)^{-1}$$

Les hypothèses cruciales pour la mise en œuvre de l'estimateur des doubles différences sont les suivantes:

1. les effets temporels sont supposés être communs aux deux groupes de traitement et de contrôle ($\lambda_t^{\text{Tr}} = \lambda_t^{\text{C}}$) au moins en $t = \tau$ et $t - 1 = \tau - 1$;
2. il ne peut y avoir ni attrition, ni sélection endogène entre $(\tau - 1)$ et τ ; ainsi, cette méthode ne peut être appliquée à une réforme fiscale (telle que la Prime pour l'emploi) qui peut influencer la participation au marché du travail, variable par nature discrète ; en ce cas, certaines personnes peuvent être sans emploi en $(\tau - 1)$ mais employées en τ (ou l'inverse) ;
3. les termes d'erreur et les covariables sont supposées non auto-corrélés ; dans le cas contraire, l'écart-type de l'effet du traitement est systématiquement sous-estimé, comme l'ont montré Bertrand, Duflo et Mullainathan (2004) ; en ce cas, l'hypothèse nulle d'absence d'effet du traitement, i.e. $H_0: \gamma = 0$, est trop souvent rejetée.

6.2 Exemples

Card et Krueger (1994) ont appliqué cette méthode pour estimer les effets du salaire minimum sur l'emploi. Ils utilisent pour cela le fait qu'en 1992, le salaire minimum a été augmenté de 80% au New Jersey. Leur idée consiste à mesurer l'évolution de l'emploi dans les fast-foods situés des deux côtés de la frontière entre les états du New Jersey et de Pennsylvanie. Appliquant la méthode des doubles différences aux deux années 1991 et 1993, ils trouvent de

manière inattendue que la hausse de l'emploi (en équivalent à temps complet) a été de 0.47 points dans le New Jersey, alors que l'emploi (en équivalent à temps complet) a baissé de 2.28 points en Pennsylvanie dans le même temps. Dans un article ultérieur, Card (1990) a mesuré les effets de l'immigration sur l'emploi non qualifié en mettant en oeuvre une fois encore cette méthode. Il utilise comme « *expérience naturelle* » l'arrivée massive d'immigrés cubains non qualifiés à Miami en 1979. Il estime l'évolution du taux de chômage des travailleurs non qualifiés blancs, noirs et hispaniques à Miami, mais aussi dans quatre autres villes américaines (Atlanta, Houston, Los Angeles et Tampa) qui n'ont pas connu de flux migratoires significatifs au cours de la même période (1979-1981). Il constate une hausse de 1.3 points du taux de chômage des noirs à Miami, mais une hausse de 2.6 points dans les autres villes durant ces années.

Plus récemment, Giavazzi et Tabellini (2005) ont mesuré les effets de l'ouverture des économies et de la démocratisation des institutions politiques sur le taux de croissance du PIB, sur le taux d'inflation et le déficit public de différents pays. Ils utilisent pour cela des données agrégées relatives à 140 pays, observés de 1960 à 2000, et des indicateurs dichotomiques d'ouverture des économies et de démocratisation des institutions politiques pour chaque pays et chaque année. Dans cette analyse, les observations du groupe des « traités » sont les pays ayant connu une réforme économique ou politique majeure durant la période. Les observations du groupe de contrôle sont les pays n'ayant connu aucune réforme majeure durant la période (toujours fermés ou ouverts, toujours démocratiques ou jamais). L'ordre des réformes semble important: les pays qui ont d'abord libéralisé leur économie avant de démocratiser leurs institutions ont fait mieux au regard des variables macroéconomiques retenues.

Draca, Machin et Van Reenen (2005) ont estimé les effets de l'introduction d'un salaire minimum en 1999 au Royaume-Uni. Les effets de cette mesure étaient a priori différents pour les entreprises employant beaucoup de salariés au niveau du salaire minimum avant la réforme (les entreprises « traitées ») et celles en employant peu (les « contrôles »). Draca, Machin et Van Reenen (2005) trouvent un effet significatif du salaire minimum sur la rentabilité des entreprises : la marge de profit, définie comme le rapport du profit brut aux ventes, a été réduite de 8 à 11% après introduction du salaire minimum. Mais ils ne trouvent pas d'effet significatif sur la probabilité de défaillance des entreprises.

6.3 L'estimateur des doubles différences avec appariement (difference in difference kernel matching)

La validité de la condition d'identification $Y_0 \perp\!\!\!\perp T|X$, nécessaire pour la mise en œuvre des techniques d'estimation par appariement, peut être légitimement mise en cause par la présence d'un effet fixe individuel inobservable, noté u , affectant à la fois les variables potentielles de résultat et la variable de traitement. Une hypothèse plus vraisemblable pourrait donc être $Y_0 \perp\!\!\!\perp T|X, u$. Cette hypothèse stipule donc qu'une partie des variables nécessaires à la propriété d'indépendance est inobservée.

Ce cadre peut se ramener au cadre développé dans les sections 3 et 4 dès lors qu'il est possible d'éliminer l'information contenue dans l'élément inobservé par une transformation adaptée de celles-ci. En partant de l'hypothèse d'indépendance conditionnellement à des observables et à des inobservables, on peut en effet montrer le résultat suivant :

$$Y_0 \perp\!\!\!\perp T|X, u \text{ et } g(Y_0, X) \perp\!\!\!\perp u|X \Rightarrow g(Y_0, X) \perp\!\!\!\perp T|X \quad (19)$$

Un cas intéressant est celui dans lequel la liste des variables de conditionnement comprend les observations des variables de résultat à une date antérieure au traitement, notées Y^p , avec $Y^p \subset X$. Alors, par soustraction, il est possible d'éliminer l'effet individuel : $(Y_0 - Y^p) \perp\!\!\!\perp u|X$, à l'instar de ce qui est communément effectué avec les méthodes de différences de différences présentée dans la sous-section précédente.⁷ On obtient alors la propriété d'indépendance conditionnellement à des observables pour les évolutions des variables de résultat, propriété qui s'écrit $(Y_0 - Y^p) \perp\!\!\!\perp T|X$. Ceci a conduit Heckman, Ichimura et Todd (1998) à une généralisation de l'estimateur par différences de différences, largement utilisé dans le cas de l'évaluation des politiques publiques. La forme de l'estimateur de l'effet moyen du traitement sur les traités est alors :

$$\hat{\Delta}_{DDKM}^{TT} = \frac{1}{N_1} \sum_{i \in I_1} \left\{ \Delta y_i - \sum_{j \in I_0} \frac{K[(P(x_j) - P(x_i))/h]}{\sum_{j \in I_0} K[(P(x_j) - P(x_i))/h]} \Delta y_j \right\} \quad (20)$$

où $\Delta Y = Y - Y^p$.

⁷ L'article pionnier en la matière est celui d'Ashenfelter (1978).

7. La régression avec discontinuité

La méthode de la régression avec discontinuité est appliquée en général dans les cas où la probabilité de traitement est une *fonction discontinue* de la valeur prise par une covariable autre que la variable de résultat (par exemple, l'âge ou le revenu de l'individu, l'effectif de l'entreprise, etc.). Deux situations doivent être distinguées :

- le cas avec discontinuité nette ("*sharp design*") : ici l'accès au traitement T dépend de manière déterministe des valeurs prises par un vecteur de variables observables Z , soit $T = f(Z)$, et le point z_0 auquel il y a discontinuité est connu avec certitude. Par exemple, si Z est de dimension 1 :

$$T = 1 \text{ si } Z > z_0 \text{ et } T = 0 \text{ si } Z < z_0 \text{ (ou l'inverse)}$$

- le cas avec discontinuité floue ("*fuzzy design*") : ici l'accès au traitement T est une variable aléatoire conditionnée par Z , et la probabilité conditionnelle :

$$\Pr(T = 1 | Z = z) = E(T | Z = z) \equiv f(z)$$

est discontinue en $Z = z_0$.

Le problème provient du fait que Z peut être corrélé avec la variable de résultat. Une comparaison des résultats des agents ayant reçu le traitement et de ceux n'y ayant pas eu accès est de ce fait généralement biaisée. Toutefois, il est vraisemblable que les agents proches du point de discontinuité (situés des deux côtés de la limite ou du seuil) soient très semblables.

Posons $Y = Y_0 + (Y_1 - Y_0) T = \alpha + \beta T$. Dans le cas d'une discontinuité nette ("*sharp design*"), et si ε est un nombre arbitrairement faible, alors :

$$E(Y | Z = z_0 + \varepsilon) - E(Y | Z = z_0 - \varepsilon) = E(\beta | Z = z_0 + \varepsilon) + E(\alpha | Z = z_0 + \varepsilon) - E(\alpha | Z = z_0 - \varepsilon)$$

Si les agents proches du seuil sont identiques, on devrait avoir :

$$E(\alpha|Z = z_0 + \varepsilon) \cong E(\alpha|Z = z_0 - \varepsilon)$$

Les conditions d'identification de la régression avec discontinuité sont donc :

1. **(C1)** $E(\alpha|Z = z)$ est continue en $Z = z_0$
2. **(C2)** La limite $\lim_{\varepsilon \rightarrow 0} E(\beta|Z = z_0 + \varepsilon)$ existe et est définie.

Sous ces conditions,

$$\lim_{\varepsilon \rightarrow 0} [E(Y|Z = z_0 + \varepsilon) - E(Y|Z = z_0 - \varepsilon)] = E(\beta|Z = z_0)$$

Remarquons que l'effet moyen du traitement sur les traités n'est identifié que pour les agents proches du seuil (sauf dans le cas où cet effet est constant dans la population). Toutefois, dans le cas de points de discontinuité multiples, il est possible d'approcher l'identification de cet effet « pour tout Z » : il est donc possible de tester l'hypothèse d'effet constant du traitement.

Considérons maintenant le modèle plus général :

$$Y = \alpha + \beta T = g(Z) + \beta T + u \quad (21)$$

où $g(Z) = E(\alpha|Z)$ est une fonction de contrôle flexible pour la régression de Y sur T . Dans le cas d'une discontinuité floue (*fuzzy design*), Van der Klaauw (2001) propose de remplacer T par un estimateur de première étape de $E(T|Z)$. Dans le cas général (*sharp* ou *fuzzy design*), le rapport :

$$\frac{\lim_{\varepsilon \rightarrow 0^+} E(Y|Z = z_0 + \varepsilon) - \lim_{\varepsilon \rightarrow 0^+} E(Y|Z = z_0 - \varepsilon)}{\lim_{\varepsilon \rightarrow 0^+} E(T|Z = z_0 + \varepsilon) - \lim_{\varepsilon \rightarrow 0^+} E(T|Z = z_0 - \varepsilon)} \quad (22)$$

identifie l'effet du traitement en $Z = z_0$

A titre d'exemple, considérons l'étude de Saez (2002). Celui-ci souhaite savoir si l'on observe une concentration des foyers fiscaux au voisinage des limites des tranches du barème fiscal aux U.S.A.. Les modèles standards, notamment d'offre de travail en présence de contrainte budgétaire discontinue, prévoient que l'on devrait observer ce type de phénomène.

Mais les données fiscales collectées entre 1960 et 1997 contredisent cette prédiction: à l'aide de régressions avec discontinuité (au voisinage des seuils de tranche), Saez conclut qu'il n'y a pas de « *bunching* », sauf au seuil délimitant les deux premières tranches

8. Le modèle de sélectivité sur variables inobservables

8.1 Définitions

Une solution alternative pour résoudre le problème de sélectivité est de recourir à une modélisation jointe des résultats potentiels (Y_0, Y_1) et de l'affectation au traitement T , en faisant l'hypothèse que ces trois variables dépendent de termes d'erreur inobservables, potentiellement corrélés entre eux. Ce modèle repose sur un formalisme simple. La variable d'affectation au traitement est supposée être déterminée par un indice latent T^* , appelé la propension à être traitée, cet indice étant lui-même linéairement dépendant d'un vecteur de variables explicatives Z et d'un résidu V , de sorte que l'on peut écrire :

$$T = \mathbf{1}(T^* > 0) = \mathbf{1}(Z\gamma + V > 0) \quad (23)$$

où γ est un vecteur de paramètres inconnu, et $\mathbf{1}(\cdot)$ est une fonction indicatrice prenant la valeur 1 si l'expression entre parenthèses est vraie, 0 sinon. Les variables de résultat latentes sont supposées être chacune engendrée par un modèle de régression linéaire de la forme

$$Y_j = a_j + R b_j + U_j, \quad j=0,1 \quad (24)$$

où R est un vecteur de variables explicatives a priori différent de Z , b_j est un vecteur de paramètres associé à R , et U_j est un résidu centré. Dans cette écriture on fait l'hypothèse que les éléments inobservés sont indépendants des variables explicatives R et Z . Les paramètres b_0, b_1 et γ sont des paramètres structurels et ont donc une interprétation économique. Dans ce cadre, les expressions des paramètres d'intérêt $\Delta^{TT}(R, Z) = E(Y_1 - Y_0 | T=1, R, Z)$ et $\Delta^{ATE}(R, Z) = E(Y_1 - Y_0 | R, Z)$ deviennent :

$$\Delta^{TT}(R, Z) = a_1 - a_0 + R(b_1 - b_0) + E(U_1 - U_0 | T=1, R, Z) \quad (25)$$

$$\Delta^{ATE}(R, Z) = a_1 - a_0 + R(b_1 - b_0) \quad (26)$$

L'intérêt central de cette modélisation est de reposer sur l'existence d'une dépendance entre les éléments inobservés affectant le traitement et les résultats potentiels. C'est la raison pour laquelle ce modèle est appelé modèle de sélection sur inobservables. On voit que si la modélisation des résultats potentiels est structurelle, en revanche les paramètres d'intérêt ne le sont pas, puisqu'ils font intervenir la distribution des éléments inobservés V , U_0 et U_1 .

8.2 Le modèle Tobit de sélection endogène

Dès lors que l'on spécifie la loi jointe des éléments inobservés, l'estimation d'un tel modèle par la méthode du maximum de vraisemblance ne pose en général pas de problème particulier. En pratique, on fait très souvent l'hypothèse que les résidus V , U_0 et U_1 suivent une loi normale de moyenne 0 et de matrice de variances et covariances Σ , soit

$$\begin{pmatrix} V \\ U_0 \\ U_1 \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma = \begin{bmatrix} 1 & & \\ \rho_0 \sigma_0 & \sigma_0^2 & \\ \rho_1 \sigma_1 & \sigma_{01} & \sigma_1^2 \end{bmatrix} \right) \quad (27)$$

où σ_j^2 est la variance de U_j ($j = 0, 1$), σ_{01} est la covariance entre U_0 et U_1 (qui n'est pas identifiable en coupe puisqu'en ce cas, pour chaque unité statistique, seule l'une des deux variables Y_0 et Y_1 est observée), et ρ_j est le coefficient de corrélation linéaire entre V et U_j .

Sous cette hypothèse de normalité, on peut écrire l'espérance de la variable de résultat pour les individus traités et les individus non traités à l'aide des formules suivantes :

$$\begin{aligned} \Pr(T = 1|Z) &= \Phi(Z\gamma) \\ E(Y|R, Z, T = 0) &= a_0 + Rb_0 - \rho_0 \sigma_0 \phi(Z\gamma)/(1 - \Phi(Z\gamma)) \\ E(Y|R, Z, T = 1) &= a_1 + Rb_1 + \rho_1 \sigma_1 \phi(Z\gamma)/\Phi(Z\gamma) \end{aligned} \quad (28)$$

où ϕ est la fonction de densité de la loi normale standard, et Φ est sa fonction de répartition.

Par conséquent, les paramètres d'intérêt sont dans ce cas égaux à :

$$\begin{aligned}\Delta^{TT} &= a_1 - a_0 + R(b_1 - b_0) + (\rho_1\sigma_1 - \rho_0\sigma_0) \frac{\phi(Z\gamma)}{\Phi(Z\gamma)} \\ \Delta^{ATE} &= a_1 - a_0 + R(b_1 - b_0)\end{aligned}\tag{29}$$

8.3 Deux exemples d'application

Dans la période récente, deux études françaises ont utilisé le modèle Tobit, que nous venons de présenter, pour évaluer des dispositifs de la politique d'emploi.

Cavaco et Fougère (2007) ont estimé l'effet d'un passage par une convention de conversion sur la durée du chômage. Le dispositif des conventions de conversion, qui a pris fin au 30 juin 2001, avait pour but de favoriser le reclassement des licenciés économiques. Les données utilisées par Cavaco et Fougère (2007) proviennent de deux enquêtes longitudinales conduites par la DARES. Les résultats qu'ils obtiennent montrent qu'il existe bien un biais de sélection à l'entrée du dispositif dû aux variables observables et inobservables. La probabilité d'adhérer à une convention de conversion est plus élevée pour les travailleurs ayant subi un licenciement collectif sans fermeture d'établissement, et ceux dont la qualification et l'ancienneté dans l'emploi préalable étaient plus élevées. L'estimation de la variation moyenne de la durée de chômage résultant d'un passage par une convention révèle qu'un tiers, ou au plus 41%, seulement des bénéficiaires de convention voient leur durée moyenne de chômage baisser après passage par le dispositif. A l'inverse, la durée de chômage de plus de la moitié des chômeurs licenciés n'ayant pas adhéré à une convention de conversion aurait pu être réduite par un passage en convention.

De leur côté, Fougère, Goux et Maurin (2001) ont évalué l'impact des formations financées par les employeurs sur la mobilité et les rémunérations des salariés. Pour cela, ils ont utilisé l'enquête sur la Formation et la Qualification Professionnelle (FQP) réalisée en 1993 par l'INSEE. Cette enquête permet d'identifier les employeurs avant et après les actions de formation continue réalisées entre 1988 et 1993. Elle permet également d'apparier le fichier des données individuelles aux fichiers de données disponibles sur les entreprises. De ce fait, les auteurs de l'étude ont pu disposer d'instruments originaux pour l'identification des effets structurels d'un système d'équations simultanées reliant formation continue, mobilité et salaires. Le modèle économétrique est un modèle à double sélection. Tout d'abord, les employeurs sélectionnent les candidats à la formation continue et/ou les salariés décident de

suivre une formation continue (l'accès à la formation continue correspond à une première équation de sélection). A l'issue de la formation continue, certains salariés peuvent quitter leur emploi ou être licenciés. Les éventuels changements d'entreprise observés au cours de la période correspondent à une seconde équation de sélection. La spécification de cette seconde équation permet de tester si le passage préalable par une formation continue a un effet direct sur les décisions de séparation. Les résidus de ces deux équations et celui de l'équation de salaire sont supposés pouvoir être corrélés. Si les valeurs estimées de ces corrélations sont significativement différentes de zéro, alors les mécanismes de sélection sont déterminés par des variables non observées par l'économètre. Le signe de ces corrélations permet par ailleurs de préciser la nature de ces sélections. Si, par exemple, le résidu de l'équation probit d'accès à la formation continue est estimé être négativement corrélé avec les déterminants non observés des salaires dans les entreprises autres que l'entreprise formatrice, alors, *toutes choses observables égales par ailleurs, les entreprises forment moins les travailleurs dont les salaires potentiels dans les autres entreprises sont élevés*. Si la corrélation entre le résidu de l'équation d'accès à la formation continue et celui de l'équation de changement d'entreprise sont négativement corrélés, cela signifie que, *toutes choses observables égales par ailleurs, les entreprises forment plus les travailleurs les plus stables, ceux dont elle ne souhaite pas se séparer ou ceux les moins enclins à la quitter*. L'estimation valide cette dernière hypothèse. En même temps, elle montre que (1) la formation continue n'a pas d'impact significatif sur les carrières salariales dans les entreprises formatrices, (2) la formation continue atténue toutefois l'impact salarial négatif associé en moyenne aux transitions entre deux emplois, (3) les décisions de mobilité des salariés sont assez peu sensibles aux variations de salaires offerts.

8.4 Extensions paramétriques du modèle Tobit

D'autres estimateurs paramétriques peuvent être mis en œuvre de manière assez simple. Il suffit pour cela de considérer des familles de lois plus générales que la loi normale. Lee (1983) propose ainsi deux types de généralisation. L'une porte sur la spécification du modèle de sélection, l'autre porte sur la spécification jointe des perturbations du modèle. La première modification consiste à remarquer que si V a pour fonction de répartition $F(V) = P(v \leq V)$, alors la fonction $u = F(v)$ suit une loi uniforme sur l'intervalle $[0,1]$. Grâce à ce résultat, on voit que la variable aléatoire transformée $\Phi^{-1} \circ F(v)$ suit une loi normale. Pour se retrouver dans le cadre du modèle Tobit présenté plus haut, il suffit donc de faire l'hypothèse que U_0 ,

U_i , et $\Phi^{-1} \circ F(v)$ suivent une loi normale. En d'autres termes, les expressions précédentes se généralisent au cas où $P(Z)$ n'est pas directement de la forme probit $\Phi(Z\gamma)$. Le modèle transformé est donc :

$$\begin{aligned} \Pr(T=1|Z) &= P(Z) \\ E(Y|R, Z, T=0) &= a_0 R b_0 - \rho_0 \sigma_0 \phi \circ \Phi^{-1}(P(Z)) / (1 - P(Z)) \\ E(Y|R, Z, T=1) &= a_1 + R b_1 + \rho_1 \sigma_1 \phi \circ \Phi^{-1}(P(Z)) / P(Z) \end{aligned} \quad (30)$$

Les paramètres d'intérêt ont alors des expressions proches des précédentes. En particulier, Δ^{TT} est maintenant égal à :

$$\Delta^{TT} = E\left(Y - a_0 - R b_0 - \rho_0 \sigma_0 \frac{\phi \circ \Phi^{-1} P(Z)}{P(Z)} \middle| T=1 \right) \quad (31)$$

La seconde généralisation consiste à adopter un ensemble de lois paramétriques plus flexibles que la loi normale pour représenter la distribution jointe des perturbations. Lee (1983) a proposé de considérer les lois de Student de degré v choisi (le cas normal correspond à la situation dans laquelle $v = +\infty$). On peut montrer dans ce cas que les expressions précédentes se réécrivent sous la forme :

$$\begin{aligned} \Pr(T=1|Z) &= P(Z) \\ E(Y|R, Z, T=0) &= a_0 + R b_0 - \rho_0 \sigma_0 \left[\frac{v + (T_v^{-1}(P(Z)))^2}{v-1} \right] \frac{t_v \circ T_v^{-1}(P(Z))}{1 - P(Z)} \\ E(Y|R, Z, T=1) &= a_1 + R b_1 + \rho_1 \sigma_1 \left[\frac{v + (T_v^{-1}(P(Z)))^2}{v-1} \right] \frac{t_v \circ T_v^{-1}(P(Z))}{P(Z)} \end{aligned} \quad (32)$$

Dans ces expressions, t_v représente la densité d'une loi de Student à v degrés de liberté et T_v est sa fonction de répartition. On remarque que cette généralisation n'est pas incompatible avec la première. En effet, la fonction P correspondant à la fonction de répartition des résidus de l'équation de sélection peut être une loi paramétrique spécifiée de manière plus générale encore. Le paramètre Δ^{TT} peut donc s'écrire sous la forme plus générale :

$$\Delta^{TT} = E \left(a_1 - a_0 + R(b_1 - b_0) + (\rho_1 \sigma_1 - \rho_0 \sigma_0) \left[\frac{\nu + (T_\nu^{-1}(P(Z)))^2}{\nu - 1} \right] \frac{t_\nu \circ T_\nu^{-1}(P(Z))}{P(Z)} \middle| T=1 \right) \quad (33)$$

On peut ainsi proposer toute une gamme d'estimateurs paramétriques du modèle de sélection sur inobservables. A partir d'expérience de simulation et de mise en œuvre dans un cas concret, Heckman, Tobias et Vytlačil (2000) montrent que les biais en cas de mauvaise spécification peuvent être importants et que les résultats sont assez sensibles aux hypothèses sur la loi des perturbations. On peut remarquer que l'une comme l'autre des généralisations envisagées permettent sans difficulté d'identifier les constantes a_0 et a_1 qui jouent un rôle fondamental dans l'estimation des deux paramètres d'intérêt. C'est le fait d'imposer des formes fonctionnelles aux lois des termes inobservables qui rend possible l'identification de ces constantes.

9. Conclusion : quelle méthode adopter?

Il y a donc plusieurs façons de traiter le problème de la sélectivité. Chacune correspond à des hypothèses bien spécifiques. En pratique, il est important de savoir laquelle de ces hypothèses est la mieux adaptée. Les travaux d'Heckman, Ichimura et Todd (1997, 1998) apportent quelques éléments de réponse à cette question. Ils utilisent les données du programme de formation *JTPA (Job Training Partnership Act)* mis en place aux Etats-Unis. Ce programme a été conçu comme une expérience contrôlée. Il est donc possible d'obtenir une évaluation de son effet en utilisant le groupe de contrôle expérimental. En utilisant un groupe de contrôle non expérimental, il est par ailleurs possible de mettre en œuvre les différents estimateurs non expérimentaux que nous venons de présenter. Il s'agit ensuite d'examiner ceux qui se rapprochent le plus de l'évaluation sur données expérimentales. Les enseignements de cet exercice sont les suivants :

1. Les méthodes d'appariement sur les niveaux des variables de résultat (plutôt que sur leurs différences premières) sont mal adaptées : il y a persistance d'effets individuels inobservés, même lorsque l'on introduit un grand nombre de caractéristiques individuelles observables ;

2. Les méthodes d'appariement appliquées aux variations des variables de résultat, telles que l'estimation par « différences de différences », marchent bien, mais leur bonne performance dépend néanmoins de la richesse des variables de conditionnement. Ces méthodes ne sont pas pour autant la panacée : il subsiste en général un biais qui, bien qu'il ne soit pas le plus important, représente une fraction significative de la valeur de l'effet du traitement. Ceci suggère que les méthodes par appariement ne doivent pas être utilisées lorsqu'il n'est pas possible d'observer l'évolution de la situation des individus, ou lorsqu'il n'est pas possible d'éliminer les effets individuels par simple application des différences premières.
3. Les problèmes relatifs au support des scores de propension ont une réelle importance dans la pratique. Ils reflètent l'hétérogénéité de l'effet du traitement dans la population et mettent l'accent sur les difficultés d'extrapolation des résultats obtenus.
4. L'étude réalisée par Mueser, Troske et Gorislawsky (2003) à l'aide de données non expérimentales complète les résultats obtenus par Heckman, Ichimura et Todd (1997, 1998) ; en outre, elle montre que les choix techniques liés aux méthodes d'appariement (choix de la fonction noyau, choix de la fenêtre, etc.) donnent des résultats assez peu différents les uns des autres, mais aussi de ceux obtenus avec des méthodes plus classiques, telles que les régressions à variables de contrôle; toutefois, ces choix affectent la précision des estimations ;
5. L'hypothèse de normalité dans les modèles de sélection sur inobservables conduit en général à des estimations biaisées ;⁸ en particulier, l'étude d'Heckman, Tobias et Vytlačil (2000) montre que les paramètres déduits du modèle de sélection sur inobservables sont très sensibles à la spécification de la loi des termes d'erreur inobservés.

Au total, il existe aujourd'hui des méthodes alternatives à la démarche expérimentale, souvent jugée coûteuse au double point de vue juridique et politique. Néanmoins, l'utilisation des méthodes non expérimentales dépend fortement du contexte propre à l'observation. Ce contexte détermine en particulier le recours à tel ou tel type de méthode dans le but de réduire

⁸ Bien évidemment, cette critique ne s'adresse pas aux seuls modèles de sélection sur inobservables utilisés en évaluation, mais à beaucoup d'autres modèles économétriques.

les biais spécifiques au problème de l'évaluation. Mais il conditionne aussi la précision des estimations et donc leur caractère informatif et opérationnel. L'absence de données non expérimentales, mais suffisamment riches, explique aujourd'hui pour une grande part le retard pris par les pays européens (à part quelques exceptions) en matière d'évaluation quantitative des politiques publiques.

Les méthodes d'évaluation sur micro-données ont par ailleurs une limite évidente : elles permettent seulement d'identifier l'effet de programmes qui n'affectent qu'une faible part de la population. Dès lors que ces programmes ont une ampleur plus importante, on ne peut ignorer l'existence d'effets sur l'équilibre général du marché du travail.⁹

⁹ Sur les effets d'équilibre général d'une politique publique, le lecteur pourra consulter l'article d'Heckman, Lochner et Taber (1998).

Références bibliographiques

Ashenfelter O. (1978) : “Estimating the Effects of Training Programs on Earnings”, *Review of Economics and Statistics*, vol. 60, pp. 47-57.

Atoyan R. et Conway C. (2006) : “Evaluating the Impact of IMF Programs: A Comparison of Matching and IV Estimators”, *The Review of International Organizations*, vol. 1, pp. 99-124.

Bertrand M., Duflo E. et Mullainathan S. (2004) : “How Much Should We Trust Differences-in-Differences Estimates?”, *Quarterly Journal of Economics*, vol. 119, pp. 249-75.

Bonnal L., Fougère D. et Sérandon A. (1997) : “Evaluating the Impact of French Employment Policies on Individual Labour Market Histories”, *Review of Economic Studies*, vol. 64, pp. 683-713.

Brodaty T., Crépon B. et Fougère D. (2001) : “Using Kernel Matching Estimators to Evaluate Alternative Youth Employment Programs: Evidence from France, 1986-1988”, dans *Econometric Evaluations of Labour Market Policies*, édité par M. Lechner et F. Pfeiffer, Physica Verlag, Heidelberg, pp. 85-124.

Brodaty T., Crépon B. et Fougère D. (2002) : “Do Long-Term Unemployed Workers Benefit from Active Labour Market Policies ? Evidence from France 1986-1998”, mimeo, CREST-INSEE.

Card D. (1990): “The Impact of the Mariel Boatlift on the Miami Labor Market”, *Industrial and Labor Relations Review*, vol. 43, pp. 245-257.

Card D. et Krueger A. (1994): “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”, *The American Economic Review*, vol. 84, pp. 772-793

Cavaco S. et Fougère D. (2007) : « Conventions de conversion et durées de chômage », à paraître dans *Cahiers Travail et Emploi*, La Documentation Française, Paris.

Crépon B. et Desplatz R. (2001) : « Une nouvelle évaluation des effets des allègements de charges sociales sur les bas salaires », *Economie et Statistique*, n°348, pp. 3-24.

Dehejia R. et Wabba S. (1999) : “Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs”, *Journal of the American Statistical Association*, vol. 94, pp. 1053-1062.

Dehejia R. et Wabba S. (2002) : “Propensity Score Matching Methods for Non-Experimental Causal Studies”, *Review of Economics and Statistics*, vol. 84, pp. 151-161.

Draca M., Machin S. et Van Reenen J. (2005): “Minimum Wages and Firm Profitability”, *CEPR Discussion Paper No. 1913*, Londres.

Fougère D. (2000a) : « Expérimenter pour évaluer les politiques d'aide à l'emploi: les exemples anglo-saxons et nord-européens », *Revue Française des Affaires Sociales*, vol. 54, pp. 111-144.

Fougère D. (2000b) : « Accompagnement des chômeurs et sanctions : leurs effets sur le retour à l'emploi », dans *Plein Emploi*, Rapport du Conseil d'Analyse Economique n° 30, La Documentation Française, Paris, pp. 313-340.

Fougère D., Goux D. et Maurin E. (2001) : « Formation continue et carrières salariales: une évaluation sur données individuelles », *Annales d'Economie et de Statistique*, n° 62, pp. 49-70.

Fougère D., Kramarz F. et Magnac T. (2000) : “Youth Employment Policies in France”, *European Economic Review*, vol. 44, pp. 928-942.

Giavazzi F. et Tabellini G. (2005): “Economic and Political Liberalizations”, *Journal of Monetary Economics*, vol. 52, pp. 1297-1330.

Härdle W. et O. Linton (1994) : “Applied Nonparametric Methods”, dans *Handbook of Econometrics*, Vol. 4, édité par R. F. Engle et D. McFadden, North-Holland, Amsterdam, pp. 2295-2339.

Heckman J. (1997): “Instrumental Variables : A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations”, *Journal of Human Resources*, vol. 32, pp. 441-462.

Heckman J. et Honoré B. (1990) : “The Empirical Content of the Roy Model”, *Econometrica*, vol. 50, pp. 1121-1149.

Heckman J. et Hotz J. (1989) : “Choosing Among Alternative Methods of Estimating the Impact of Social Programs : The Case of Manpower Training”, *Journal of the American Statistical Association*, vol. 84, pp. 862-874.

Heckman J., Ichimura H., Smith J. et Todd P. (2000) : “Characterizing Selection Bias Using Experimental Data”, *Econometrica*, vol. 66, pp. 1017-1098.

Heckman J., Ichimura H. et Todd P. (1997) : “Matching as an Econometric Evaluation Estimator : Evidence from Evaluating a Job Training Program”, *Review of Economic Studies*, vol. 64, pp. 605-654.

Heckman J., Ichimura H. et Todd P. (1998) : “Matching as an Econometric Evaluation Estimator”, *Review of Economic Studies*, vol. 65, pp. 261-294.

Heckman J., Lalonde R. et Smith J. (1999) : “The Economics and Econometrics of Active Labor Market Programs”, dans *Handbook of Labor Economics*, vol. III, édité par O. Ashenfelter et D. Card, Amsterdam : North Holland.

Heckman J., Lochner L. et Taber C. (1998) : “General-Equilibrium Treatment Effects: A Study of Tuition Policy”, *The American Economic Review*, vol. 88, pp. 381-386.

Heckman J. et Robb R. (1985) : “Alternative Methods for Evaluating the Impact of Interventions : An Overview”, *Journal of Econometrics*, vol. 30, pp. 239-267.

Heckman J., Tobias J. et Vytlacil E. (2000) : “Simple Estimators for Treatment Parameters in a Latent Variable Framework With an Application to Estimating the Return to Schooling”, *NBER Working paper* n°7950.

Heckman J. et Vytlacil E. (2000) : “Local Instrumental Variables”, *NBER Working Paper* n°T0252.

Holland P. (1986) : “Statistics and Causal Inference”, *Journal of the American Statistical Association*, vol. 81, pp. 945-960.

Ichimura H. et Taber C. (2002) : “Semiparametric Reduced Form Estimation of Tuition Subsidies”, CeMMAP Working Paper CWP01/02, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Imbens, G. W. (2000): “The Role of the Propensity Score in Estimating Dose-Response Functions”. *Biometrika*, Vol. 87, pp. 706-710.

Imbens G. et Angrist J. (1994) : “Identification and Estimation of Local Average Treatment Effects”, *Econometrica*, vol. 62, pp. 467-476.

Kramarz F. et Philippon T. (2001) : “The Impact of Differential Payroll Tax Subsidies on Minimum Wage Employment”, *Journal of Public Economics*, vol. 82, pp. 115-146.

Lalonde R. (1986) : “Evaluating the Econometric Evaluations of Training Programs with Experimental Data”, *American Economic Review*, vol. 76, pp. 604-620.

Lechner M. (2001) : “Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption”, dans *Econometric Evaluations of Labour Market Policies*, édité par M. Lechner et F. Pfeiffer, Heidelberg : Physica Verlag, pp. 43-58.

Lee L. (1983) : “Generalized Econometric Models with Selectivity”, *Econometrica*, vol. 51, pp 507-512.

Lise J., Seitz S. et Smith J. (2005) : “Equilibrium Policy Experiments and the Evaluation of Social Programs”, Working Paper, Department of Economics, University of Maryland.

Magnac T. (2000a) : “State Dependence and Heterogeneity in Youth Employment Histories”, *The Economic Journal*, vol. 110, pp. 805-837.

Mueser P., Troske K. and Gorislavsky A. (2003) : “Using State Administrative Data to Measure Program Performance”, *IZA Discussion Paper* N° 786.

Pimenta Monteiro N. (2004): “Using Propensity Matching Estimators to Evaluate the Impact of Privatization on Wages”, Working Paper No. 12/2004, Universidade do Minho.

Rosenbaum P. et Rubin D. (1983) : “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika*, vol. 70, pp. 41-55.

Rosenbaum P. et Rubin D. (1985) : “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score”, *The American Statistician*, Vol. 39, pp. 33-38.

Rubin D. (1974) “Estimating Causal Effects of Treatments in Randomized and Non Randomized Studies”, *Journal of Educational Psychology*, vol. 66, pp. 688-701.

Rubin D. (1977) : “Assignment to Treatment Group on the Basis of a Covariate”, *Journal of Educational Statistics*, vol. 2, pp. 1-26.

Saez E. (2002): “Do Taxpayers Bunch at Kink Points?”, NBER Working Paper No. 7366.

Sianesi B. (2004): “An Evaluation of the Swedish System of Active Labour Market Programmes in the 1990s”, *Review of Economics and Statistics*, vol. 86, pp. 133-155.

Todd P. et Wolpin K. (2005) : “Ex Ante Evaluations of Social Programs”, Working Paper, University of Pennsylvania.

Van der Klaauw W. (2002): “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach”, *International Economic Review*, vol 43, pp. 1249-1287.

Notes d'Études et de Recherche

1. C. Huang and H. Pagès, "Optimal Consumption and Portfolio Policies with an Infinite Horizon: Existence and Convergence," May 1990.
2. C. Bordes, « Variabilité de la vitesse et volatilité de la croissance monétaire : le cas français », février 1989.
3. C. Bordes, M. Driscoll and A. Sauviat, "Interpreting the Money-Output Correlation: Money-Real or Real-Real?," May 1989.
4. C. Bordes, D. Goyeau et A. Sauviat, « Taux d'intérêt, marge et rentabilité bancaires : le cas des pays de l'OCDE », mai 1989.
5. B. Bensaïd, S. Federbusch et R. Gary-Bobo, « Sur quelques propriétés stratégiques de l'intéressement des salariés dans l'industrie », juin 1989.
6. O. De Bandt, « L'identification des chocs monétaires et financiers en France : une étude empirique », juin 1990.
7. M. Boutillier et S. Dérangère, « Le taux de crédit accordé aux entreprises françaises : coûts opératoires des banques et prime de risque de défaut », juin 1990.
8. M. Boutillier and B. Cabrillac, "Foreign Exchange Markets: Efficiency and Hierarchy," October 1990.
9. O. De Bandt et P. Jacquinot, « Les choix de financement des entreprises en France : une modélisation économétrique », octobre 1990 (English version also available on request).
10. B. Bensaïd and R. Gary-Bobo, "On Renegotiation of Profit-Sharing Contracts in Industry," July 1989 (English version of NER n° 5).
11. P. G. Garella and Y. Richelle, "Cartel Formation and the Selection of Firms," December 1990.
12. H. Pagès and H. He, "Consumption and Portfolio Decisions with Labor Income and Borrowing Constraints," August 1990.
13. P. Sicsic, « Le franc Poincaré a-t-il été délibérément sous-évalué ? », octobre 1991.
14. B. Bensaïd and R. Gary-Bobo, "On the Commitment Value of Contracts under Renegotiation Constraints," January 1990 revised November 1990.
15. B. Bensaïd, J.-P. Lesne, H. Pagès and J. Scheinkman, "Derivative Asset Pricing with Transaction Costs," May 1991 revised November 1991.
16. C. Monticelli and M.-O. Strauss-Kahn, "European Integration and the Demand for Broad Money," December 1991.
17. J. Henry and M. Phelipot, "The High and Low-Risk Asset Demand of French Households: A Multivariate Analysis," November 1991 revised June 1992.
18. B. Bensaïd and P. Garella, "Financing Takeovers under Asymmetric Information," September 1992.

19. A. de Palma and M. Uctum, "Financial Intermediation under Financial Integration and Deregulation," September 1992.
20. A. de Palma, L. Leruth and P. Régibeau, "Partial Compatibility with Network Externalities and Double Purchase," August 1992.
21. A. Frachot, D. Janci and V. Lacoste, "Factor Analysis of the Term Structure: a Probabilistic Approach," November 1992.
22. P. Sicsic et B. Villeneuve, « L'afflux d'or en France de 1928 à 1934 », janvier 1993.
23. M. Jeanblanc-Picqué and R. Avesani, "Impulse Control Method and Exchange Rate," September 1993.
24. A. Frachot and J.-P. Lesne, "Expectations Hypothesis and Stochastic Volatilities," July 1993 revised September 1993.
25. B. Bensaid and A. de Palma, "Spatial Multiproduct Oligopoly," February 1993 revised October 1994.
26. A. de Palma and R. Gary-Bobo, "Credit Contraction in a Model of the Banking Industry," October 1994.
27. P. Jacquinet et F. Mihoubi, « Dynamique et hétérogénéité de l'emploi en déséquilibre », septembre 1995.
28. G. Salmat, « Le retournement conjoncturel de 1992 et 1993 en France : une modélisation VAR », octobre 1994.
29. J. Henry and J. Weidmann, "Asymmetry in the EMS Revisited: Evidence from the Causality Analysis of Daily Eurorates," February 1994 revised October 1994.
30. O. De Bandt, "Competition Among Financial Intermediaries and the Risk of Contagious Failures," September 1994 revised January 1995.
31. B. Bensaid et A. de Palma, « Politique monétaire et concurrence bancaire », janvier 1994 révisé en septembre 1995.
32. F. Rosenwald, « Coût du crédit et montant des prêts : une interprétation en terme de canal large du crédit », septembre 1995.
33. G. Cette et S. Mahfouz, « Le partage primaire du revenu : constat descriptif sur longue période », décembre 1995.
34. H. Pagès, "Is there a Premium for Currencies Correlated with Volatility? Some Evidence from Risk Reversals," January 1996.
35. E. Jondeau and R. Ricart, "The Expectations Theory: Tests on French, German and American Euro-rates," June 1996.
36. B. Bensaid et O. De Bandt, « Les stratégies "stop-loss" : théorie et application au Contrat Notionnel du Matif », juin 1996.
37. C. Martin et F. Rosenwald, « Le marché des certificats de dépôts. Écarts de taux à l'émission : l'influence de la relation émetteurs-souscripteurs initiaux », avril 1996.

38. Banque de France - CEPREMAP - Direction de la Prévision - Erasme - INSEE - OFCE, « Structures et propriétés de cinq modèles macroéconomiques français », juin 1996.
39. F. Rosenwald, « L'influence des montants émis sur le taux des certificats de dépôts », octobre 1996.
40. L. Baumel, « Les crédits mis en place par les banques AFB de 1978 à 1992 : une évaluation des montants et des durées initiales », novembre 1996.
41. G. Cette et E. Kremp, « Le passage à une assiette valeur ajoutée pour les cotisations sociales : Une caractérisation des entreprises non financières “gagnantes” et “perdantes” », novembre 1996.
42. S. Avouyi-Dovi, E. Jondeau et C. Lai Tong, « Effets “volume”, volatilité et transmissions internationales sur les marchés boursiers dans le G5 », avril 1997.
43. E. Jondeau et R. Ricart, « Le contenu en information de la pente des taux : Application au cas des titres publics français », juin 1997.
44. B. Bensaïd et M. Boutillier, « Le contrat notionnel : efficience et efficacité », juillet 1997.
45. E. Jondeau et R. Ricart, « La théorie des anticipations de la structure par terme : test à partir des titres publics français », septembre 1997.
46. E. Jondeau, « Représentation VAR et test de la théorie des anticipations de la structure par terme », septembre 1997.
47. E. Jondeau et M. Rockinger, « Estimation et interprétation des densités neutres au risque : Une comparaison de méthodes », octobre 1997.
48. L. Baumel et P. Sevestre, « La relation entre le taux de crédits et le coût des ressources bancaires. Modélisation et estimation sur données individuelles de banques », octobre 1997.
49. P. Sevestre, “On the Use of Banks Balance Sheet Data in Loan Market Studies : A Note,” October 1997.
50. P.-C. Hautcoeur and P. Sicsic, “Threat of a Capital Levy, Expected Devaluation and Interest Rates in France during the Interwar Period,” January 1998.
51. P. Jacquinet, « L'inflation sous-jacente à partir d'une approche structurelle des VAR : une application à la France, à l'Allemagne et au Royaume-Uni », janvier 1998.
52. C. Bruneau et O. De Bandt, « La modélisation VAR structurel : application à la politique monétaire en France », janvier 1998.
53. C. Bruneau and E. Jondeau, “Long-Run Causality, with an Application to International Links between Long-Term Interest Rates,” June 1998.
54. S. Coutant, E. Jondeau and M. Rockinger, “Reading Interest Rate and Bond Futures Options' Smiles: How PIBOR and Notional Operators Appreciated the 1997 French Snap Election,” June 1998.
55. E. Jondeau et F. Sédillot, « La prévision des taux longs français et allemands à partir d'un modèle à anticipations rationnelles », juin 1998.

56. E. Jondeau and M. Rockinger, "Estimating Gram-Charlier Expansions with Positivity Constraints," January 1999.
57. S. Avouyi-Dovi and E. Jondeau, "Interest Rate Transmission and Volatility Transmission along the Yield Curve," January 1999.
58. S. Avouyi-Dovi et E. Jondeau, « La modélisation de la volatilité des bourses asiatiques », janvier 1999.
59. E. Jondeau, « La mesure du ratio rendement-risque à partir du marché des euro-devises », janvier 1999.
60. C. Bruneau and O. De Bandt, "Fiscal Policy in the Transition to Monetary Union: A Structural VAR Model," January 1999.
61. E. Jondeau and R. Ricart, "The Information Content of the French and German Government Bond Yield Curves: Why Such Differences?," February 1999.
62. J.-B. Chatelain et P. Sevestre, « Coûts et bénéfices du passage d'une faible inflation à la stabilité des prix », février 1999.
63. D. Irac et P. Jacquinot, « L'investissement en France depuis le début des années 1980 », avril 1999.
64. F. Mihoubi, « Le partage de la valeur ajoutée en France et en Allemagne », mars 1999.
65. S. Avouyi-Dovi and E. Jondeau, "Modelling the French Swap Spread," April 1999.
66. E. Jondeau and M. Rockinger, "The Tail Behavior of Stock Returns: Emerging Versus Mature Markets," June 1999.
67. F. Sédillot, « La pente des taux contient-elle de l'information sur l'activité économique future ? », juin 1999.
68. E. Jondeau, H. Le Bihan et F. Sédillot, « Modélisation et prévision des indices de prix sectoriels », septembre 1999.
69. H. Le Bihan and F. Sédillot, "Implementing and Interpreting Indicators of Core Inflation: The French Case," September 1999.
70. R. Lacroix, "Testing for Zeros in the Spectrum of an Univariate Stationary Process: Part I," December 1999.
71. R. Lacroix, "Testing for Zeros in the Spectrum of an Univariate Stationary Process: Part II," December 1999.
72. R. Lacroix, "Testing the Null Hypothesis of Stationarity in Fractionally Integrated Models," December 1999.
73. F. Chesnay and E. Jondeau, "Does correlation between stock returns really increase during turbulent period?," April 2000.
74. O. Burkart and V. Coudert, "Leading Indicators of Currency Crises in Emerging Economies," May 2000.

75. D. Irac, "Estimation of a Time Varying NAIRU for France," July 2000.
76. E. Jondeau and H. Le Bihan, "Evaluating Monetary Policy Rules in Estimated Forward-Looking Models: A Comparison of US and German Monetary Policies," October 2000.
77. E. Jondeau and M. Rockinger, "Conditional Volatility, Skewness, and Kurtosis: Existence and Persistence," November 2000.
78. P. Jacquinot et F. Mihoubi, « Modèle à Anticipations Rationnelles de la CONjoncture Simulée : MARCOS », novembre 2000.
79. M. Rockinger and E. Jondeau, "Entropy Densities: With an Application to Autoregressive Conditional Skewness and Kurtosis," January 2001.
80. B. Amable and J.-B. Chatelain, "Can Financial Infrastructures Foster Economic Development?," January 2001.
81. J.-B. Chatelain and J.-C. Teurlai, "Pitfalls in Investment Euler Equations," January 2001.
82. M. Rockinger and E. Jondeau, "Conditional Dependency of Financial Series: An Application of Copulas," February 2001.
83. C. Florens, E. Jondeau and H. Le Bihan, "Assessing GMM Estimates of the Federal Reserve Reaction Function," March 2001.
84. J.-B. Chatelain, "Mark-up and Capital Structure of the Firm facing Uncertainty," June 2001.
85. B. Amable, J.-B. Chatelain and O. De Bandt, "Optimal Capacity in the Banking Sector and Economic Growth," June 2001.
86. E. Jondeau and H. Le Bihan, "Testing for a Forward-Looking Phillips Curve. Additional Evidence from European and US Data," December 2001.
87. G. Clette, J. Mairesse et Y. Kocoglu, « Croissance économique et diffusion des TIC : le cas de la France sur longue période (1980-2000) », décembre 2001.
88. D. Irac and F. Sédillot, "Short Run Assessment of French Economic Activity Using OPTIM," January 2002.
89. M. Baghli, C. Bouthevillain, O. de Bandt, H. Fraisse, H. Le Bihan et Ph. Rousseaux, « PIB potentiel et écart de PIB : quelques évaluations pour la France », juillet 2002.
90. E. Jondeau and M. Rockinger, "Asset Allocation in Transition Economies," October 2002.
91. H. Pagès and J.A.C. Santos, "Optimal Supervisory Policies and Depositor-Preferences Laws," October 2002.
92. C. Loupias, F. Savignac and P. Sevestre, "Is There a Bank Lending Channel in France? Evidence from Bank Panel Data," November 2002.
93. M. Ehrmann, L. Gambacorta, J. Martínez-Pagés, P. Sevestre and A. Worms, "Financial Systems and The Role in Monetary Policy Transmission in the Euro Area," November 2002.
94. S. Avouyi-Dovi, D. Guégan et S. Ladoucette, « Une mesure de la persistance dans les indices boursiers », décembre 2002.

95. S. Avouyi-Dovi, D. Guégan et S. Ladoucette, "What is the Best Approach to Measure the Interdependence between Different Markets?," December 2002.
96. J.-B. Chatelain and A. Tiomo, "Investment, the Cost of Capital and Monetary Policy in the Nineties in France: A Panel Data Investigation," December 2002.
97. J.-B. Chatelain, A. Generale, I. Hernando, U. von Kalckreuth and P. Vermeulen, "Firm Investment and Monetary Policy Transmission in the Euro Area," December 2002.
98. J.-S. Mésonnier, « Banque centrale, taux de l'escompte et politique monétaire chez Henry Thornton (1760-1815) », décembre 2002.
99. M. Baghli, G. Cette et A. Sylvain, « Les déterminants du taux de marge en France et quelques autres grands pays industrialisés : Analyse empirique sur la période 1970-2000 », janvier 2003.
100. G. Cette and Ch. Pfister, "The Challenges of the "New Economy" for Monetary Policy," January 2003.
101. C. Bruneau, O. De Bandt, A. Flageollet and E. Michaux, "Forecasting Inflation using Economic Indicators: the Case of France," May 2003.
102. C. Bruneau, O. De Bandt and A. Flageollet, "Forecasting Inflation in the Euro Area," May 2003.
103. E. Jondeau and H. Le Bihan, "ML vs GMM Estimates of Hybrid Macroeconomic Models (With an Application to the "New Phillips Curve")," September 2003.
104. J. Matheron and T.-P. Maury, "Evaluating the Fit of Sticky Price Models," January 2004.
105. S. Moyen and J.-G. Sahuc, "Incorporating Labour Market Frictions into an Optimising-Based Monetary Policy Model," January 2004.
106. M. Baghli, V. Brunhes-Lesage, O. De Bandt, H. Fraise et J.-P. Villette, « MASCOTTE : Modèle d'Analyses et de préviSion de la Conjoncture TrimésTrielle », février 2004.
107. E. Jondeau and M. Rockinger, "The Bank Bias: Segmentation of French Fund Families," February 2004.
108. E. Jondeau and M. Rockinger, "Optimal Portfolio Allocation Under Higher Moments," February 2004.
109. C. Bordes et L. Clerc, « Stabilité des prix et stratégie de politique monétaire unique », mars 2004.
110. N. Belorgey, R. Lecat et T.-P. Maury, « Déterminants de la productivité par employé : une évaluation empirique en données de panel », avril 2004.
111. T.-P. Maury and B. Pluyaud, "The Breaks in per Capita Productivity Trends in a Number of Industrial Countries," April 2004.
112. G. Cette, J. Mairesse and Y. Kocoglu, "ICT Diffusion and Potential Output Growth," April 2004.

113. L. Baudry, H. Le Bihan, P. Sevestre and S. Tarrieu, "Price Rigidity. Evidence from the French CPI Micro-Data," September 2004.
114. C. Bruneau, O. De Bandt and A. Flageollet, "Inflation and the Markup in the Euro Area," September 2004.
115. J.-S. Mésonnier and J.-P. Renne, "A Time-Varying "Natural" Rate of Interest for the Euro Area," September 2004.
116. G. Clette, J. Lopez and P.-S. Noual, "Investment in Information and Communication Technologies: an Empirical Analysis," October 2004.
117. J.-S. Mésonnier et J.-P. Renne, « Règle de Taylor et politique monétaire dans la zone euro », octobre 2004.
118. J.-G. Sahuc, "Partial Indexation, Trend Inflation, and the Hybrid Phillips Curve," December 2004.
119. C. Loupias et B. Wigniolle, « Régime de retraite et chute de la natalité : évolution des mœurs ou arbitrage micro-économique ? », décembre 2004.
120. C. Loupias and R. Ricart, "Price Setting in France: new Evidence from Survey Data," December 2004.
121. S. Avouyi-Dovi and J. Matheron, "Interactions between Business Cycles, Stock Markets Cycles and Interest Rates: the Stylised Facts," January 2005.
122. L. Bilke, "Break in the Mean and Persistence of Inflation: a Sectoral Analysis of French CPI," January 2005.
123. S. Avouyi-Dovi and J. Matheron, "Technology Shocks and Monetary Policy in an Estimated Sticky Price Model of the US Economy," April 2005.
124. M. Dupaigne, P. Fève and J. Matheron, "Technology Shock and Employment: Do We Really Need DSGE Models with a Fall in Hours?," June 2005.
125. P. Fève and J. Matheron, "Can the Kydland-Prescott Model Pass the Cogley-Nason Test?," June 2005.
126. S. Avouyi-Dovi and J. Matheron, "Technology Shocks and Monetary Policy in an Estimated Sticky Price Model of the Euro Area," June 2005.
127. O. Loisel, "Central Bank Reputation in a Forward-Looking Model," June 2005.
128. B. Bellone, E. Gautier et S. Le Coent, « Les marchés financiers anticipent-ils les retournements conjoncturels ? », juillet 2005.
129. P. Fève, « La modélisation macro-économétrique dynamique », juillet 2005.
130. G. Clette, N. Dromel and D. Méda, "Opportunity Costs of Having a Child, Financial Constraints and Fertility," August 2005.
131. S. Gouteron et D. Szpiro, « Excès de liquidité monétaire et prix des actifs », septembre 2005.
132. J. Baude, « L'impact des chocs boursiers sur le crédit en France depuis le milieu des années quatre-vingt-dix », septembre 2005.

133. R. Broulès and G. Clette, "A Comparison of Structural Productivity Levels in the Major Industrialised Countries," October 2005.
134. T. Grunspan, "The Fed and the Question of Financial Stability: An Empirical Investigation," October 2005.
135. S. Fabiani, M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl and A. Stockman, "The Pricing Behaviour of Firms in the Euro Area: New Survey Evidence," November 2005.
136. E. Dhyne, L. Alvarez, H. Le Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lünemann, F. Rumler and J. Vilminen, "Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data," November 2005.
137. D. Fougère, H. Le Bihan and P. Sevestre, "Heterogeneity in Consumer Price Stickiness: A Microeconomic Investigation," November 2005.
138. L. Alvarez, E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilminen, "Sticky Prices in the Euro Area: a Summary of New Micro Evidence," November 2005.
139. E. Kharroubi, "Illiquidity, Financial Development and the Growth-Volatility Relationship," February 2006.
140. M. Baghli, C. Cahn and H. Fraise, "Is the Inflation-Output Nexus Asymmetric in the Euro Area," April 2006.
141. E. Jondeau and J-G. Sahuc, "Optimal Monetary Policy in an Estimated DSGE Model of the Euro Area with Cross-country Heterogeneity," April 2006.
142. S. Avouyi-Dovi, M. Brun, A. Dreyfus, F. Drumetz, V. Oung et J.-G. Sahuc, « La fonction de demande de monnaie pour la zone euro : un réexamen », mai 2006.
143. C. Jarret, "Term Structure Anomalies : Term Premium or Peso Problem?" May 2006.
144. S. Avouyi-Dovi, R. Kierzenkowski and C. Lubochinsky, "Are Business and Credit Cycles Converging or Diverging? A comparison of Poland, Hungary, the Czech Republic and the Euro Area", May 2006.
145. O. De Bandt, C. Bruneau and A. Flageollet, "Assessing Aggregate Comovements in France, Germany and Italy. Using a Non Stationary Factor Model of the Euro Area" June 2006.
146. M. Baghli, C. Cahn and J-P. Villetelle, "Estimating Potential Output with a Production Function for France, Germany and Italy", June 2006.
147. E. Fonteny, « La désaisonnalisation des séries d'agrégats monétaires et de crédit à la Banque de France : aspects théoriques et mise en œuvre », juin 2006.
148. J. Matheron and C. Poilly, "How Well Does a Small Structural Model with Sticky Prices and Wages Fit Postwar U.S. Data", July 2006.
149. E. Kharroubi, "Financial (Dis)Integration," July 2006.

150. J. Carrillo, P. Fève and J. Matheron, “Monetary Policy Inertia or Persistent Shocks?” July 2006.

151. C. Ewerhart, N. Cassola and N. Valla, "Declining Valuations and Equilibrium Bidding in Central Bank Refinancing Operations," August 2006.
152. D. Fougère, « Réformes structurelles sur le marché du travail : quels enseignements peut-on tirer des études existantes », août 2006.
153. D. Irac, "Revisiting the proximity-concentration trade-off: Distance and Horizontal Foreign Direct Investment in OECD Countries," August 2006.
154. D. Irac and C. Minoiu, "Risk insurance in a Transition Economy: Evidence from Rural Romania," August 2006.
155. H. Lustig and A. Verdelhan, "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk," August 2006.
156. R. Boulès and G. Cette, "Trends in "structural" productivity levels in the major industrialized countries," September 2006.
157. J.-S. Mésonnier, "The Reliability of Macroeconomic Forecasts based on Real Interest Rate Gap Estimates in Real Time: an Assessment for the Euro Area," October 2006.
158. O. de Bandt, C. Bruneau and W. El Amri, "Convergence in Household Credit Demand across Euro Area Countries: Evidence from Panel Data," October 2006.
159. J. Idier, "Stock Exchanges Industry Consolidation and Shock Transmission," December 2006.
160. E. Gautier, "The Behaviour of Producer Prices: Some Evidence from the French PPI Micro Data," December 2006.
161. O. Loisel, "Bubble-free interest-rate rules," December 2006.
162. J. Boivin and M. P. Giannoni, "DSGE Models in a Data-Rich Environment," January 2007.
163. J. Coffinet, J. Matheron et C. Poilly, « Une évaluation structurelle du ratio de sacrifice dans la zone euro », janvier 2007.
164. P. Vermeulen, D. Dias, M. Dossche, E. Gautier, I. Hernando, R. Sabbatini and H. Stahl, "Price setting in the euro area: Some stylised facts from Individual Producer Price Data," February 2007.
165. C. Bordes, L. Clerc and V. Marimoutou, "Is there a structural break in equilibrium velocity in the euro area?" February 2007.
166. D. Fougère, « Les méthodes micro-économétriques d'évaluation », mars 2007.

Pour tous commentaires ou demandes sur les Notes d'Études et de Recherche, contacter la bibliothèque de la direction de la recherche à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Research Directorate at the following address :

BANQUE DE FRANCE
41- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 92 49 55 ou 62 65
fax :0033 (0)1 42 92 62 92
email : thierry.demoulin@banque-france.fr
jeannine.agoutin@banque-france.fr