

Estimation de l'incidence des mesures actives d'aide à l'emploi au moyen de données administratives et de méthodes d'appariement¹

Andy Handouyahia², Stéphanie Roberge², Yves Gingras², Tony Haddad³, Georges Awad²

Résumé

Dans le présent document, nous abordons les répercussions des prestations d'emploi et mesures de soutien versées au Canada, en vertu des Ententes sur le développement du marché du travail. Nous utilisons un ensemble riche de données administratives longitudinales couplées englobant tous les participants aux EDMT de 2002 à 2005. Sur la base d'un appariement par score de propension, comme dans Blundell et coll. (2002), Gerfin et Lechner (2002), et Sianesi (2004), nous avons produit des estimations de l'impact différentiel à l'échelle nationale à partir d'un estimateur de l'écart des différences et d'un estimateur par la méthode du noyau (Heckman et Smith, 1999). Les résultats laissent supposer que, tant pour les Services d'aide à l'emploi que pour les programmes de prestations d'emploi, comme le Développement des compétences et les Subventions salariales ciblées, des effets positifs se font sentir sur la rémunération et l'emploi.

Mots-clés : appariement par score de propension, PEMS, évaluation des programmes de mesures actives d'emploi, emploi, rémunération

1. Introduction

La présente étude applique des méthodes d'appariement au moyen d'un panel de données administratives pour l'analyse de l'incidence des prestations d'emploi et mesures de soutien (PEMS) versées au Canada, en vertu des Ententes sur le développement du marché du travail (EDMT). Les PEMS examinées comprennent les Services d'aide à l'emploi (SAE), le Développement des compétences (DC) et les Subventions salariales ciblées (SSC). À partir de données administratives riches couvrant toutes les personnes qui ont commencé à participer entre avril 2002 et mars 2005, l'analyse évalue les répercussions des programmes de mesures actives d'emploi sur l'employabilité et la rémunération des personnes participantes, à court et à moyen terme (jusqu'à cinq ans), suivant leur participation. C'est la première fois qu'une telle analyse est effectuée à l'échelle nationale au Canada.

Un guide utile est fourni par Card, Kluge et Weber, 2010, à savoir une méta-analyse des évaluations microéconométriques récentes des politiques de mesures actives d'emploi, découlant de 97 études distinctes sur les répercussions des programmes dans les pays d'Europe, les pays nordiques et les pays anglophones (Card et coll., 2010). Il est important de souligner que les services offerts au Canada en vertu des PEMS ressemblent à ceux

¹ Les points de vue exprimés dans les documents de recherche sont ceux des auteurs et ne représentent pas forcément les opinions d'Emploi et développement social Canada (EDSC) ou du gouvernement fédéral. Le présent document est fondé sur des méthodologies et analyses élaborées dans le contexte des évaluations des EDMT. Au moment de l'élaboration de ces méthodologies, les évaluateurs d'EDSC ont profité des conseils de divers experts universitaires et d'un examen par les pairs. Par ailleurs, en 2014, les évaluateurs ont tenu un panel d'experts externes pour passer en revue et examiner les méthodologies d'évaluation. De façon plus particulière, nous aimerions remercier Walter Nicholson, Jeff Smith, Guy Lacroix et David Gray pour les conseils fournis concernant ces études d'évaluation.

² Andy Handouyahia, gestionnaire, Données et méthodologie; Stéphanie Roberge, agente principale d'évaluation; Yves Gingras, directeur général; Georges Awad, gestionnaire, Évaluation, Direction d'évaluation, Emploi et développement social, 140, promenade du Portage, Gatineau (Québec) K1A 0J9, Canada.

³ Tony Haddad, Ph. D., gestionnaire, Évaluation, Division de l'évaluation et de la mesure du rendement, Statistique Canada, 170, promenade Tunney's Pasture, Ottawa (Ontario) K1A 0T6, Canada.

des autres pays de l'OCDE. Afin de limiter l'analyse et la discussion, le présent document est axé sur les trois interventions suivantes des PEMS :

- Les SAE fournissent une gamme variée de services de soutien pour aider les personnes en chômage qui ont besoin d'aide pour entrer ou pour revenir sur le marché du travail. Les services offerts vont de l'aide à la recherche d'emploi pour les clients prêts à travailler à l'élaboration de plans d'action approfondis pour le retour au travail, dans le cas des clients qui font face à plusieurs obstacles à l'emploi.
- Le DC aide les personnes à acquérir des compétences en emploi, qui vont des niveaux de base aux niveaux avancés. Il fournit une aide directe aux personnes pour la formation et, au besoin, des contributions aux provinces et aux territoires ou aux établissements de formation financés par les provinces et les territoires pour couvrir les coûts qui ne sont pas compris dans les frais de scolarité.
- Les SSC encouragent les employeurs à recruter des personnes en chômage, en leur fournissant une aide financière qui couvre une partie des salaires des nouvelles recrues, ainsi que certains coûts liés à l'emploi.

La grande majorité des ouvrages publiés ont été axés sur les programmes de formation en bonne et due forme; et un résumé exhaustif est fourni par Heckman, LaLonde et Smith (1999). Le présent document tire parti des ouvrages publiés par le passé en reliant les discussions théoriques, méthodologiques et empiriques pour examiner les répercussions et l'efficacité des interventions en mesures actives d'emploi au Canada.

L'utilisation de données de panel informatives, qui relient toutes les données administratives de l'assurance-emploi (a.-e.) et de l'Agence du revenu du Canada, contribue de façon importante à la présente analyse. Cette dernière couvre les prestataires actifs des 13 provinces et territoires qui ont commencé à participer aux SAE, DC et SSC entre le 1^{er} avril 2002 et le 31 mars 2005. L'évaluation des répercussions des SAE et du DC se limite à un échantillon aléatoire de 10 % et 50 %, respectivement, des prestataires actifs, en raison du nombre très important de personnes participant à ces interventions et pour des considérations de capacité informatique (38 564 participants pour les SAE, 64 283 participants pour le DC et 18 767 participants pour les SSC). Le groupe de comparaison comprend des prestataires actifs de l'a.-e.⁴ qui étaient admissibles à participer au cours de cette période, mais qui ne l'ont pas fait. Nous estimons les effets séparément pour les SAE, le DC et les SSC. Les groupes de comparaison finaux sélectionnés comprenaient 263 176 cas de comparaison pour les SAE, 274 062 pour le DC et 146 284 pour les SSC. Les principaux indicateurs des résultats sur le marché du travail pour la présente analyse sont l'incidence moyenne de l'emploi (c.-à-d. la probabilité annuelle moyenne d'avoir une rémunération d'emploi), la rémunération moyenne d'emploi et la somme de prestations d'a.-e. reçues. Tous ces indicateurs sont mesurés sur une base annuelle.

2. Données

L'analyse de l'impact différentiel est effectuée à partir de données de panel administratives couplées de la banque de données de l'a.-e. (données de la partie I de l'a.-e. sur les prestations d'a.-e. et données de la partie II de l'a.-e. sur la participation aux PEMS), ainsi que des fichiers de l'impôt de l'Agence du revenu du Canada. Les données disponibles pour l'étude couvrent de grands échantillons aléatoires et, dans certains cas, jusqu'à 100 % des participants aux PEMS au Canada, 20 % des prestataires d'a.-e. sans participation aux PEMS dans les trois territoires et les provinces de l'Atlantique, et 10 % des prestataires d'a.-e. sans participation aux PEMS dans les provinces qui restent, de 1990 à 2011. À cet égard, il convient de souligner que les données utilisées pour les politiques, l'analyse, la recherche et l'évaluation sont régies par des règles et processus stricts mis en place au niveau des ministères pour assurer la protection des renseignements personnels et la confidentialité. De façon plus particulière, les données utilisées pour l'évaluation sont assujetties à un processus qui permet de masquer les données d'identification personnelle (p. ex., les numéros d'assurance sociale, les noms ou les adresses) avant qu'elles soient analysées. Ce processus sert à prévenir l'identification de personnes au moment de l'élaboration des politiques, de l'analyse, de la recherche et de l'évaluation.

L'étude porte sur l'impact différentiel pour les participants qui ont commencé à recevoir des PEMS entre 2002 et 2005, et les suit pendant une période pouvant aller jusqu'à six années consécutives, entre 2002 et 2011. Le choix de la période de référence est influencé par le désir de mesurer les répercussions du programme sur une période d'au

⁴« Prestataire actif de l'a.-e. » signifie une personne qui touche des prestations d'assurance-emploi au moment de la participation.

moins cinq ans suivant la fin de la participation⁵. L'année la plus récente pour laquelle nous disposons des données de l'Agence du revenu du Canada est 2011. Par conséquent, l'analyse est axée sur des équivalents de plan d'action⁶ qui ont commencé entre le 1^{er} avril 2002 et le 31 mars 2005. L'équivalent de plan d'action est l'unité d'analyse utilisée dans l'étude. Elle regroupe toutes les PEMS versées à une personne à moins de six mois d'intervalle. L'objectif de l'équivalent de plan d'action est de rendre compte de la série de programmes et de services d'emplois offerts aux participants aux PEMS. Étant donné que les équivalents de plan d'action comprennent plus d'une intervention, aux fins de la production des rapports, l'effet observé est attribué à l'intervention la plus longue comprise dans l'équivalent de plan d'action, à moins que ce dernier ne soit composé que de SAE.

Dans l'ensemble, la grande majorité des participants compris dans l'étude ont terminé leurs équivalents de plan d'action au cours de la période de 2003 à 2006.

3. Approche d'évaluation et méthodologie

L'étude examine les répercussions des prestations d'emploi, comme les SAE, le DC et les SSC, en comparant des personnes qui ont profité de ces interventions et celles qui ne l'ont pas fait.

Les méthodes d'appariement combinées à l'écart des différences servent à déterminer les répercussions. Le concept de base de ces méthodes est que, selon les caractéristiques observées, deux groupes comportant la même probabilité de participation se retrouveront dans des proportions égales dans les échantillons de participants et de non-participants. Ainsi, ils peuvent être combinés aux fins de la comparaison des résultats (avant et après la participation). Le principal avantage de ces méthodes est que l'appariement ne nécessite pas d'hypothèses de forme fonctionnelle pour l'équation des résultats (c.-à-d. qu'elles sont non paramétriques) et qu'elles ne sont par conséquent pas sujettes à un biais de spécification inexacte pour cet aspect.

La présente étude a adopté la structure courante de Rubin, appelée approche des résultats possibles, ou le modèle de Neyman-Roy-Rubin (Neyman, 1935; Roy, 1951; Rubin, 1974). La méthode d'appariement utilisée est l'appariement pas score de propension (Rosenbaum et Rubin (1983)), qui est défini comme la probabilité d'affectation au traitement conditionnelle aux covariables de base observées et habituellement désignée comme l'hypothèse d'indépendance conditionnelle (Rosenbaum et Rubin, 1983; Dehejia et Wahba, 2002; Heckman et coll., 1998; Caliendo et Kopeinig, 2008; Galdo, Smith et Black, 2007; Smith et Todd, 2005; Imbens, 2004; Caliendo et coll., 2005).

Les données utilisées pour l'analyse englobent un nombre important de caractéristiques, qui rendent compte de l'expérience de la personne sur le marché du travail, y compris les caractéristiques sociodémographiques des participants et des cas de comparaison (p. ex., âge, sexe, état matrimonial, incapacité), leur région économique et leur province, leurs qualifications (p. ex., groupe professionnel, niveaux de compétence liés au dernier emploi avant la demande de prestations d'a.-e., codes d'industrie), ainsi que leurs antécédents sur le marché du travail (p. ex., prestations d'a.-e. et semaines, rémunération d'un emploi/travail autonome, recours à l'aide sociale, incidence de l'emploi au cours ces cinq années précédant la participation).

La richesse des données administratives disponibles permet l'inclusion des variables les plus pertinentes qui influencent la décision de participer aux PEMS et les résultats sur le marché du travail dans le modèle de score de propension. L'utilisation de caractéristiques personnelles courantes, des résultats avant le traitement, des transitions entre différents états sur le marché du travail, des données régionales et des antécédents sur le marché du travail est très importante pour diminuer la probabilité d'un estimateur biaisé de l'effet de traitement véritable.

Une évaluation de l'hypothèse d'indépendance conditionnelle montre que l'analyse tient compte de la majorité, sinon de la totalité, des facteurs qui déterminent la participation et les résultats sur le marché du travail. Dans ce contexte, on présume que l'hypothèse d'indépendance conditionnelle est satisfaite et que l'estimateur d'appariement peut être utilisé avec confiance pour estimer les effets additionnels des PEMS sur certains indicateurs de résultat. Les résultats

⁵Les répercussions sont mesurées jusqu'à sept ans après l'année du début de la participation.

du soutien commun montrent clairement un chevauchement important entre les participants et le groupe de comparaison pour chaque type de PEMS.

Afin de bien illustrer la qualité de l'appariement, nous présentons dans le tableau 1 les résultats sommaires des tests d'équilibrage des covariables, avant et après l'appariement dans chaque programme, grâce à l'utilisation de la différence moyenne uniformisée entre les participants et les non-participants (DiPrete et Gang, 2004). Les estimations montrent que la valeur moyenne pour les différences uniformisées avant l'appariement pour l'écart-type est d'environ 8,3. Après l'appariement, cette valeur passe à 0,9. Il est clair que la réduction des différences dans la répartition des covariables, entre le groupe de traitement et le groupe de comparaison, est très importante. L'appariement par score de propension a permis d'obtenir un équilibre dans les covariables. Les détails de la liste des covariables utilisées dans le modèle de score de propension et les scores d'équilibrage peuvent être fournis sur demande.

Tableau 1
Biais uniformisé moyen et pseudo R² - Résultats sommaires du test d'équilibrage sommaire

PEMS	Participants		Groupe de comparaison		Pseudo R ²		Biais moyen		Soutien commun ¹
	Avant	Après	Avant	Après	Avant	Après	Avant	Après	
Services d'aide à l'emploi (SAE)	37 382	37 364	146 28 4	146 28 4	0,193	0,003	8,6	1	18
Développement des compétences (DC)	63 481	63 450	274 06 2	274 06 2	0,178	0,002	8,3	0,9	31
Subventions salariales ciblées (SSC)	18 767	18 762	263 17 6	263 17 6	0,197	0,015	8,4	1,9	5

Nota :- Les calculs ont été effectués au moyen du progiciel PSMATCH2, par Leuven et Sianesi (2003), ainsi que PSTEST et PSGRAPH.
- Le nombre d'observations pour les participants des SAE et du DC correspond respectivement à l'échantillon aléatoire de 10 % et de 50 % des participants des fichiers originaux. Dans le cas des SSC, on utilise l'ensemble de la population des participants. Pour le groupe de comparaison, un échantillon aléatoire de 5 % a été utilisé pour les SAE et le DC, et un échantillon aléatoire de 10 % pour les SSC.

Les estimations d'impact différentiel sont produites au moyen de l'écart des différences et de l'appariement par la méthode du noyau, ainsi que deux autres méthodes à des fins de robustesse. Dans la présente analyse, les résultats différents avant et après la participation sont calculés dans les deux groupes en examinant la valeur du résultat d'intérêt au cours d'une période particulière suivant la participation, et en soustrayant la valeur du même résultat d'intérêt au cours d'une période correspondante précédant le programme, appelée période de base. La validité de la méthode de l'écart des différences dépend de la stabilité du biais entre les participants et les non-participants au cours de la période précédant le programme. Cela signifie qu'une période suffisamment longue précédant le programme est requise pour contrôler de façon appropriée la similitude entre les participants et les non-participants, ainsi que pour déterminer le fléchissement d'Ashenfelter⁷. L'étude actuelle utilise une période de cinq ans pour observer le comportement précédant le programme.

Une analyse rigoureuse de la sensibilité est aussi effectuée en appliquant des estimateurs d'appariement par la méthode du plus proche voisin et de la pondération par probabilité inverse. Cela révèle que les effets estimés ne sont pas sensibles aux différents estimateurs d'appariement. En outre, afin de vérifier le biais caché découlant des caractéristiques non observables, nous vérifions la robustesse de nos résultats et de nos modèles au moyen de ce que

⁷ Ashenfelter, O. (1978), « Estimating the Effect of Training Programs on Earnings », Review of Economics and Statistics, vol. 60, 47 à 57. Ashenfelter a déterminé que, parmi les participants, une période d'instabilité transitoire sur le marché du travail (appelée fléchissement de Ashenfelter) précède souvent la participation au programme. Étant donné que la même instabilité ne touche pas les non-participants, il est important de déterminer si une telle période de transition existe et de l'éviter lorsque l'on choisit une période de base appropriée précédant le programme. Une période précédant le programme suffisamment longue est requise pour contrôler de façon appropriée la similitude entre les participants et les non-participants, et pour déterminer et traiter le phénomène d'Ashenfelter.

l'on appelle les limites de Rosenbaum (voir Rosenbaum, 2002, 2005, 2010)⁸. À noter que ce test ne permet pas de déterminer si les résultats sont biaisés ou s'il y a des variables omises dans le modèle par score de propension, mais mesure plutôt la sensibilité des estimations à la détérioration du modèle. Les résultats de Rosenbaum, qui sont disponibles sur demande, sont partagés, mais ne modifient pas les répercussions et les directions.

La principale limite de la présente étude est la possibilité de différences existant au préalable, mais non observées, entre les participants et les cas de comparaison, qui n'ont pas été mesurées dans le processus d'appariement. Cela pourrait avoir eu une influence sur les résultats. Par exemple, des facteurs comme la capacité, la santé, la scolarité et la motivation pour trouver un emploi n'étaient pas mesurés directement, sauf lorsqu'ils étaient saisis dans les caractéristiques de revenu antérieur et de participation au marché du travail. Des données complètes existent en ce qui a trait aux professions, au niveau de compétences lié à la dernière profession, à l'a.-e. et aux caractéristiques reliées à rémunération au cours des cinq ans précédant la participation. Ces données sont utilisées avec confiance comme approximation des différences existant au préalable entre les participants et les non-participants. Dans ce contexte, il est impossible d'énoncer de façon définitive que les impacts différentiels n'ont pas été influencés par les différences existant au préalable entre les participants et les groupes de comparaison.

4. Résultats

Comme il a été indiqué précédemment, l'objectif principal du présent document est d'estimer les effets de trois programmes actifs sur le marché du travail (SAE, DC et SSC) sur la rémunération et la probabilité d'emploi des participants. L'effet global du programme en ce qui a trait à l'augmentation de l'emploi et de la rémunération chez les participants est positif et très significatif.

Le tableau 2 ci-après présente les effets de traitement estimés des différents programmes, jusqu'à cinq ans après le début. Des estimations sont aussi produites pour des sous-échantillons plus homogènes selon le groupe d'âge.

⁸ L'analyse de sensibilité est effectuée en appliquant la commande de stata *rbound* et *mhbounds* fournie par Diprete et Gangl, 2004, ainsi que Becker et Caliendo, 2007, respectivement. La commande *rbound* est exécutée sur des variables de résultats continues et est fondée sur le test de Wilcoxon pour observation appariée et l'estimation ponctuelle d'Hodges-Lehmann pour le test pour observations appariées, tandis que la commande *mhbounds* est axée sur le cas des variables de résultats binaires et est fondée sur le test statistique de Mantel et Haenszel (MH, 1959).

Tableau 2
Analyse des répercussions additionnelles pour les prestataires actifs

Incremental Impacts for Active Claimants									
Indicators	In-program period		Post-program period					Total post-program	Total in- and post-program
	Program start year	Additional program year	1st year	2nd year	3rd year	4th year	5th year		
Employment Assistance Services (EAS) (n=38,564)									
Employment earnings (\$)	-2,913***		-1,097***	-279***	347*	645***	742***	358	-2,555***
Incidence of employment (percentage points)	-0.5***		0.6*	0.8***	1.7***	1.8***	1.7***	N/A	N/A
EI benefits (\$)	697***		-451***	-312***	-251***	-222***	-136***	-1,375***	-677***
Skills Development (SD) (n=64,283)									
Employment earnings (\$)	-4,747***	-4,211***	204***	2,052***	3,077***	3,761***	4,059***	13,156***	4,197***
Incidence of employment (percentage points)	-4.5***	-4.7***	2.4***	3.7***	4***	4.2***	4.4**	N/A	N/A
EI benefits (\$)	1,847***	222***	-470***	-218***	-128***	-89***	-69***	-976***	1,093***
Targeted Wage Subsidies (TWS) (n=18,767)									
Employment earnings (\$)	-1,404***	752***	661***	971***	1,747***	1,815***	1,930***	7,125***	6,473***
Incidence of employment (percentage points)	4.4***	7.2***	5.0***	4.9***	5.1***	5.0***	5.1***	N/A	N/A
EI benefits (\$)	100***	-208***	-2	52	39	104***	146***	339***	231

Significance level *** 1%; ** 5%; * 10%

L'analyse de l'impact différentiel montre les effets positifs des SAE, du DC et des SSC pour les prestataires actifs. Les participants ont connu des hausses de leur rémunération d'emploi et de leur incidence d'emploi, et une réduction de leurs prestations d'a.-e. après leur participation. Les effets les plus prononcés touchent les participants du DC. Au cours des cinq années suivant la participation, la rémunération d'emploi annuelle des prestataires actifs était de 204 \$ à 4 059 \$ plus élevée que s'ils n'avaient pas participé au DC. En outre, les hausses de la rémunération sont accompagnées de hausses dans l'incidence de l'emploi allant de 2,4 points de pourcentage au cours de la première année à 4,4 au cours de la cinquième année suivant la participation au programme. De même, les prestataires actifs ont connu des hausses de rémunération et d'incidence de l'emploi après avoir participé aux SSC.

Les participants aux SAE ont connu des augmentations de l'incidence de l'emploi et des baisses du recours à l'a.-e. toutes les années suivant la participation, ce qui laisse supposer qu'ils sont retournés au travail après la participation. Les participants aux SAE ont connu des baisses de rémunération à court terme. Toutefois, les SAE ne sont pas axées sur le développement du capital humain et on ne s'attend pas nécessairement à ce qu'ils améliorent la rémunération d'emploi des participants.

5. Conclusion

L'objectif du présent document est d'évaluer les effets à court et à moyen terme des programmes de mesures actives d'emploi compris dans les EDMT au Canada, au moyen de méthodes non paramétriques. Il s'agit du premier document qui analyse ces programmes pour la période la plus récente et qui estime les effets à moyen terme au niveau du Canada. Le document repose sur un ensemble de données administratives riches et a déterminé des répercussions positives pour les trois programmes, après la participation.

Le DC est efficace pour augmenter la rémunération d'emploi et l'incidence de l'emploi pour les prestataires actifs. Qui plus est, dans le cas des prestataires actifs, la participation au DC a mené aux hausses additionnelles les plus importantes de la rémunération d'emploi parmi toutes les PEMS. Il convient aussi de souligner que les impacts différentiels sur la rémunération ont augmenté de façon continue au cours des cinq ans suivant la fin de la participation. On a déterminé la même tendance pour le programme des SSC, dont il a été démontré qu'il est efficace pour augmenter la rémunération d'emploi et l'incidence de l'emploi. Parmi toutes les PEMS, la participation aux SSC a entraîné les

hausse les plus importantes de l'incidence de l'emploi pour les prestataires actifs. Les SAE ont atteint leur objectif d'aider les participants à retourner sur le marché du travail, en augmentant l'incidence de l'emploi et en diminuant le recours à l'a.-e.

Bibliographie

- Becker, S.O. and Caliendo, M. (2007), « Sensitivity Analysis for Average Treatment Effects », *Stata Journal*, 7(1), 71-83. [<http://www.stata-journal.com>]
- Blundell et al. (2002), « Alternative Approaches to Evaluation in Empirical Microeconomics », *Portuguese Economic Journal*, 1, 91–115.
- Caliendo, M., and Kopeinig, S. (2008), « Some Practical Guidance for the Implementation of Propensity Score Matching », *Journal of Economic Surveys*, 2008, Vol.22(1), pp.31-72
- Card, D., Kluve, J. and Weber A. (2010), « Active Labour Market Policy Evaluations: A Meta-Analysis », *Economic Journal, Royal Economic Society*, vol. 120(548), pages F452-F477.
- Caliendo, M., Hujer, R. and Thomsen, S.L. (2005), « The Employment Effects of Job Creation Schemes in Germany - A Microeconomic Evaluation », *IDEAS Working Paper Series from RePEc*, 2005.
- Dehejia, R.H. and Wahba, S. (2002), « Propensity score matching methods for non-experimental causal studies », *Review of Economics and Statistics*, 84(1): 151–161.
- DiPrete, T. A. and Gangl, M. (2004), « Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments », *Sociological Methodology*, 34(1), 271-310.
- Galdo, J., Smith, J., and Black, D. (2007), « Bandwidth selection and the estimation of treatment effects with unbalanced data », *Annales d'Économie et de Statistique*, 1 July 2008, Issue 91/92, pp.189-216.
- Gerfin, M. and Lechner, M. (2002): « A Microeconomic Evaluation of the Active Labour Market Policy in Switzerland, » *The Economic Journal*, 112, 854–893.
- Heckman, J. J., Ichimura, H., Smith, J.A. and Todd, P. (1998), « Characterizing Selection Bias Using Experimental Data » *Econometrica*, 66, 1017–1098.
- Heckman, J. J., Ichimura, H. and Todd, P. (1997), « Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program » *Review of Economic Studies*, 64, 605–654.
- Heckman, J. J., Lalonde, R.J., and Smith, J.A. (1999): « The Economics and Econometrics of Active Labour Market Programs », *Handbook of Labour Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 31, pp. 1865–2097.
- Imbens, G. W., (2004), « Nonparametric estimation of average treatment effects under exogeneity: A review », *Review of Economics and statistics*, 86(1), 4-29.
- Lechner, M. and Wunsch, C. (2011), « Sensitivity of Matching-Based Program Evaluations of the Availability of Control Variables », *Labour Economics*, Apr 2013, Vol.21, p.111.
- Leuven, E. and Sianesi, B. (2003), «PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing », *IDEAS Working Papers*, <https://ideas.repec.org/c/boc/bocode/s432001.html>.
- Mantel, N. and Haenszel, W. (1959), « Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease », *Journal of the National Cancer Institute*, 22(4), 719-748.
- Neyman, J. (1935), « Statistical Problems in Agricultural Experimentation », *Journal of the Royal Statistical Society, Ser. B*, 2 (suppl.), 107–108.
- Rosenbaum, P. R. (2002), « Observational Studies », *Springer*, New York.
- Rosenbaum, P. R. and Rubin, D.B. (1983), « The Central Role of the Propensity Score in Observational Studies for Causal Effects », *Biometrika*, 70(1), 41–50.
- Roy, A. D. (1951), « Some Thoughts on the Distribution of Earnings », *Oxford Economic Papers*, 3, 135–145.
- Rubin, D. B. (1974), « Estimating Causal Effects to Treatments in Randomised and Nonrandomised Studies », *Journal of Educational Psychology*, 66, 688–701.

- Sianesi, B. (2004), « An Evaluation of the Active Labour Market Programs in Sweden », *The Review of Economics and Statistics*, 86(1), 133–155.
- Smith, J. A., and Todd, P. (2005), « Does Matching Overcome LaLonde’s Critique of Nonexperimental Estimators? », *Journal of Econometrics*, 125(1/2), 305–354.