Expérimentation contrôlée contre appariement : le cas d'un dispositif d'accompagnement de jeunes diplômés demandeurs d'emploi

Marianne Bléhaut, Roland Rathelot

Dans Économie & prévision 2014/1-2 n° 204-205 , pages 163 à 181
Éditions Direction générale du Trésor

ISSN 0249-4744
DOI 10.3917/ecop.204.0163
Date de mise en ligne : 14/01/2016

Article disponible en ligne à l'adresse
https://shs.cairn.info/revue-economie-et-prevision-2014-1-page-163?lang=fr

CAIRN.INFO

Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.

Distribution électronique Cairn.info pour Direction générale du Trésor.
Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur cairn.info/copyright.
Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

Expérimentation contrôlée contre appariement : le cas d'un dispositif d'accompagnement de jeunes diplômés demandeurs d'emploi

Marianne Bléhaut(*)

Roland Rathelot(**)

Le début des années 2000 a vu l'intérêt pour l'évaluation des politiques publiques croître de manière considérable. Dans un contexte budgétaire tendu, les décideurs publics sont demandeurs d'évaluations rapides, non coûteuses et non invasives. Du point de vue de l'économètre, les expérimentations aléatoires paraissent être l'outil le plus naturel. Toutefois, elles ne sont pas toujours acceptées. Les méthodes d'appariement sont faciles à mettre en œuvre et à appréhender et apparaissent comme des solutions qui s'adaptent à toutes les situations. En comparant les résultats expérimentaux aux résultats qui sont obtenus par appariement pour un même programme, cet article met en évidence leur divergence très nette. Cette divergence nous paraît très probablement liée aux limites de l'approche par appariement et nous semble inviter à la prudence aussi bien les économètres qui les mettent en œuvre que les destinataires de ces évaluations.

La validité des estimateurs expérimentaux repose sur un tirage au sort qui induit le caractère aléatoire de la variable d'assignation au traitement. Les estimateurs par appariement reposent quant à eux sur l'hypothèse dite d'indépendance conditionnelle : une fois les caractéristiques observables des individus prises en compte, leur décision de participer au traitement est supposée aléatoire. Il est en général impossible de tester la validité de cette hypothèse. Néanmoins, elle peut raisonnablement être remise en cause dans les cas où la décision d'être traité est susceptible d'être influencée par des variables inobservables pour l'économètre. Par exemple, la motivation d'un individu peut à la fois influencer sa décision de participer à un programme et sa réussite professionnelle.

À l'instar de travaux antérieurs, nous tirons parti d'une expérimentation aléatoire pour comparer ses résultats avec ceux qui découlent de l'approche par appariement. Nous utilisons ici les données issues de l'expérimentation "jeunes diplômés" (Crépon et alii*, 2011, 2013) mise en œuvre entre mi-2007 et mi-2009 pour évaluer un programme d'accompagnement renforcé*

(*) RITM, Univ. Paris-Sud, Université Paris Saclay et Centre de recherche en économie et statistique (Crest).
RITM : Réseaux, Innovation, Territoires, Mondialisation.
(**) Université de Warwick.
Email : r.rathelot@warwick.ac.uk

Nous remercions Bruno Crépon, Esther Duflo, Marc Gurgand et Philippe Zamora qui ont mené, avec Roland Rathelot, l'expérimentation sur laquelle cette étude est fondée, ainsi que la direction de l'animation de la recherche, des études et des statistiques (Dares) pour l'accès aux données. Ce texte doit également à Sylvain Chabé-Ferret, Hélène Erkel-Rousse et un rapporteur anonyme pour leurs judicieux conseils, qui ont amené de substantiels changements, ainsi qu'à Christel Colin et Béatrice Sédillot, pour leur relecture de versions antérieures.

Cet article n'engage que ses auteurs et non les institutions auxquelles ils appartiennent. Il n'engage *a fortiori* ni la Direction générale du Trésor, ni le ministère des Finances et des Comptes publics, ni le ministère de l'Économie, de l'Industrie et du Numérique.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

visant des jeunes de moins de 30 ans diplômés du supérieur et en situation de chômage de longue durée. Menée dans dix régions de France, cette expérimentation aléatoire reposait sur un mécanisme d'encouragement : chaque jeune potentiellement éligible au programme était orienté ou non vers celui-ci de manière aléatoire. S'il était orienté, la décision d'y participer était ensuite sienne. S'il n'était pas orienté ou décidait de ne pas y participer, il bénéficiait de tous les autres accompagnements de droit commun proposés par Pôle emploi. Le groupe non orienté constitue le groupe de contrôle expérimental. L'estimation de l'effet expérimental de l'accompagnement renforcé sur les bénéficiaires est réalisée en instrumentant la participation au dispositif par l'assignation aléatoire. Par construction de l'expérimentation, cette variable instrumentale possède les propriétés nécessaires à l'identification.

Pour simuler une situation non expérimentale, dans laquelle tous les individus potentiellement éligibles auraient pu prétendre au programme, nous restreignons artificiellement l'analyse à ceux qui ont été orientés vers le dispositif. Le groupe de contrôle non expérimental devient alors celui des jeunes orientés vers le programme mais qui n'ont pas choisi d'y participer. Cette stratégie permet d'exploiter la même source de données pour les deux méthodes et donc de s'affranchir des potentiels biais résultants de collectes ou de mesures différentes. Par ailleurs, le tirage au sort garantit également que les individus du groupe orienté vers le traitement sont comparables à ceux de l'échantillon total.

Les résultats que nous obtenons sont clairs et robustes à divers changements de spécification. Si ce programme avait été mis en place sans sélection aléatoire, on aurait été contraint de comparer les individus ayant choisi de participer au programme à ceux ayant refusé d'y prendre part. En utilisant des méthodes d'appariement sur score de propension comme nous l'avons fait ici, on aurait alors conclu à des conséquences négatives ou nulles de l'accompagnement renforcé, alors que l'évaluation aléatoire indique au contraire un effet positif important du programme sur la probabilité d'avoir retrouvé un emploi d'au moins six mois huit mois après l'entrée dans le programme. Notre interprétation de cet exercice est que l'hypothèse d'indépendance conditionnelle, qui sous-tend l'approche par appariement, ne possède probablement pas de validité a priori. *Tout praticien désireux d'y recourir doit examiner soigneusement, au cas par cas, si cette hypothèse a des raisons d'être vérifiée ou non. L'argument selon laquelle l'approche par appariement est la seule possible, qui est utilisé dans certains travaux, ne peut être utilisé pour justifier la validité des résultats obtenus.*

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Lorsqu'il souhaite évaluer un dispositif *a posteriori* et qu'il ne trouve ni instrument ni discontinuité permettant d'utiliser une stratégie d'identification "quasi-expérimentale", le praticien de l'évaluation est souvent contraint de s'en remettre à une hypothèse d'indépendance conditionnelle permettant de mettre en œuvre des méthodes telles que l'appariement. Cette hypothèse postule que deux individus ayant des caractéristiques observables identiques auront la même chance de choisir de bénéficier du traitement, indépendamment de leurs caractéristiques inobservables. Par construction, il est impossible de confirmer ou d'infirmer cette condition.

La simplicité des méthodes d'appariement et la possibilité technique de les appliquer en toutes circonstances les ont rendues extrêmement populaires dans quasiment tous les champs de l'économie appliquée. Cependant, elles ne constituent pas une panacée. Faute d'être vérifiée, l'hypothèse d'indépendance conditionnelle doit être au minimum discutée au cas par cas avant chaque application. Idéalement, le praticien doit être en mesure de documenter le mécanisme d'assignation des individus au programme, afin de convaincre le lecteur de la plausibilité de l'hypothèse dans le cas qui l'occupe. De manière symétrique, le lecteur d'une évaluation réalisée par appariement doit se montrer *a priori* critique vis-à-vis de la crédibilité des résultats et s'interroger sur la validité de l'hypothèse identifiante avant d'accepter les conclusions qui en sont tirées.

Dans cet article, nous utilisons les données de l'expérimentation "jeunes diplômés" (Crépon *et alii*, 2011, 2013). Cette expérimentation a été menée de 2007 à 2009 pour évaluer une prestation d'accompagnement renforcé par des opérateurs privés de placement visant les jeunes de moins de 30 ans ayant validé au moins un niveau Bac+2 et inscrits depuis plus de six mois sur les listes de Pôle emploi. Elle a été mise en place selon un dispositif d'encouragement et non un dispositif classique de sélection aléatoire des candidats souhaitant bénéficier du programme. Les jeunes éligibles ont ainsi été aléatoirement séparés en deux groupes : un groupe de traitement orienté vers l'accompagnement renforcé et un groupe de contrôle bénéficiant de l'accompagnement de droit commun offert par Pôle emploi. Le choix de participation des individus n'intervient donc qu'après leur assignation aléatoire dans le premier groupe. Pour simuler une situation non expérimentale, il est donc possible d'effacer le groupe de contrôle et de comparer le taux d'emploi des assignés, suivant qu'ils ont bénéficié ou non de l'accompagnement renforcé. Cette stratégie permet notamment de comparer les résultats expérimentaux et les résultats de l'appariement en exploitant la même source de données : celles recueillies dans le cadre de l'expérimentation. Il ne sera donc pas possible d'imputer la variation des résultats obtenus à un biais résultant de la différence entre les sources statistiques dans les deux méthodes. De plus, le caractère aléatoire de la sélection du groupe de jeunes assignés garantit *a priori* la similitude entre ce groupe et l'ensemble des jeunes faisant l'objet de l'expérimentation.

Pour mesurer ce qu'aurait été l'effet estimé selon une stratégie non expérimentale, nous avons recours à des méthodes d'appariement sur score de propension[(1)]. Nos résultats sont clairs et robustes : même lorsqu'un nombre important de variables observables est introduit dans la spécification du score de propension, les résultats non expérimentaux diffèrent largement des résultats expérimentaux. Alors que ces derniers indiquent un impact globalement positif de l'accompagnement renforcé sur l'insertion dans l'emploi (huit mois après l'entrée dans le programme), l'appariement renvoie un estimateur toujours négatif ou non-significatif de l'impact. Dans notre cas, les biais de sélection ne sont probablement pas négligeables, ce qui pourrait conduire à une conclusion opposée aux résultats expérimentaux en termes de recommandations de politique publique.

Si cet article est le premier à utiliser des données françaises, il s'inscrit dans une littérature très fournie sur la comparaison entre résultats expérimentaux et non-expérimentaux. LaLonde (1986) utilise l'expérimentation du *National Supported Work Demonstration*, une prestation d'aide à l'insertion vers l'emploi menée dans les années 1970 aux États-Unis. Afin d'évaluer le comportement des méthodes non expérimentales, il remplace artificiellement le groupe de contrôle expérimental par un groupe de contrôle obtenu dans une enquête externe au dispositif. Les résultats qu'il obtient par différence de différences sur son échantillon artificiel diffèrent singulièrement des résultats expérimentaux et varient fortement selon la spécification qu'il choisit. Dehejia et Wahba (1999) reprennent l'exemple de LaLonde et montrent que des méthodes d'appariement sur score de propension permettent de se rapprocher des résultats expérimentaux. Toutefois, ils appellent également à la prudence. Smith et Todd (2005) montrent, à partir des mêmes données, que les résultats obtenus par appariement sont très sensibles au jeu de variables de contrôle introduites. Friedlander et Robins (1995) utilisent quatre expérimentations américaines du début des années 1980 visant à insérer les inactifs dans l'emploi. Comparant les résultats expérimentaux avec ceux obtenus par appariement, ils parviennent à des résultats proches de ceux de LaLonde[(2)]. Heckman *et alii* (1998a) utilisent l'évaluation expérimentale du *Job Training Partnership Act* et comparent les résultats expérimentaux à ceux que l'on peut obtenir par des méthodes d'appariement, de différence de différences et de modèles de sélection. Ils concluent à l'existence d'un important biais de sélection.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Enfin, Arceneaux *et alii* (2006) comparent des estimateurs expérimentaux et d'appariement pour l'évaluation de l'impact d'une campagne de mobilisation électorale. Leur étude est particulièrement proche de la nôtre. Comme dans notre cas, l'assignation au traitement ne se traduit pas directement par le traitement des individus (dispositif d'encouragement).

L'expérimentation "jeunes diplômés"

Au début de l'année 2007, le ministère chargé de l'emploi a lancé un marché public de prestation d'insertion dans l'emploi à destination des jeunes diplômés demandeurs d'emploi. Dans un contexte de réflexion et de débats sur l'insertion des jeunes diplômés, ce marché visait à confier à des opérateurs privés de placement (OPP), dans dix régions[3], l'accompagnement de 10 000 jeunes. L'évaluation de cette prestation d'insertion a été menée par la direction de l'animation de la recherche, des études et des statistiques (Dares), rattachée aux ministères chargés de l'emploi et du travail, et par une équipe de chercheurs du Centre de recherche en économie et statistique (Crest) et de l'École d'économie de Paris. Elle a débouché sur la rédaction d'un rapport (Crépon *et alii*, 2011) et d'un article à vocation académique (Crépon *et alii*, 2013), auxquels le lecteur peut se référer pour plus de détails.

Menée de mi-2007 à mi-2009, la prestation d'insertion dans l'emploi des jeunes diplômés demandeurs d'emploi visait à renforcer l'offre de placement en direction de ce public, en faisant appel à des OPP pour compléter les actions déjà conduites par Pôle emploi (anciennement agence nationale pour l'emploi, ANPE) et ses co-traitants. Elle s'adressait à des jeunes de moins de 30 ans qui avaient validé un diplôme de niveau Bac+2 minimum et étaient inscrits sur les listes de Pôle emploi depuis six mois ou l'avaient été pendant 12 mois au cours des 18 mois précédents. Les jeunes qui exerçaient une activité réduite (quelle que soit sa longueur) pouvaient être éligibles à la prestation dès lors qu'ils remplissaient les conditions requises d'inscription sur les listes de Pôle emploi.

Dans le cadre de ce programme, l'OPP mettait à la disposition de chaque jeune qu'il suivait son offre de services (ateliers, simulations d'entretien, etc.) et s'engageait à suivre ce jeune de manière régulière et fréquente (un contact hebdomadaire était souhaité). Cet accompagnement vers l'emploi avait trois principaux objectifs :
1) rompre l'isolement du jeune dans sa recherche d'emploi et favoriser une meilleure adéquation entre les aspirations du jeune et les emplois disponibles sur le marché du travail ;
2) faire le bilan des compétences du jeune et établir un plan d'action ;
3) préparer le jeune aux exercices de recrutement et l'accompagner dans ses démarches.

Cette première phase, d'une durée de six mois maximum, prenait fin lorsque le demandeur d'emploi signait un contrat pour un emploi durable (contrat à durée indéterminée –CDI–, contrat à durée déterminée –CDD– d'au moins six mois, contrat de professionnalisation ou contrat d'apprentissage). La seconde phase, de six mois également, débutait dès la signature du contrat de travail. Ce suivi dans l'emploi visait à prévenir les ruptures. Concrètement, le jeune devait bénéficier de contacts réguliers avec son conseiller de l'OPP, celui-ci devant par ailleurs être disponible pour des actions supplémentaires avec le jeune ou son employeur.

Les OPP étaient rémunérés en fonction de leurs résultats. Au moment de la réponse à l'appel d'offre, chaque OPP devait indiquer un montant forfaitaire correspondant à la prise en charge intégrale d'un demandeur d'emploi (2 000 euros en moyenne, variable selon les OPP). 25 % de ce montant par bénéficiaire étaient versés à la signature du contrat d'accompagnement entre le jeune et le prestataire ; 40 % étaient versés si le jeune signait dans les six mois un contrat de travail correspondant à un emploi durable ; le reste était versé si le jeune était toujours en emploi au bout de six mois.

Dans quatre régions sur dix, les OPP retenus avaient un statut associatif ou coopératif et se définissaient eux-mêmes plutôt comme relevant de l'économie sociale (on les désignera sous la dénomination d'"opérateurs privés non marchands"). À l'inverse, dans six régions, les OPP étaient des entreprises du secteur marchand, le plus souvent des filiales de groupes spécialisés dans l'intérim et dédiées au placement ("opérateurs privés marchands"). L'orientation des jeunes diplômés vers la prestation s'est déroulée entre septembre 2007 et septembre 2008 mais les OPP pouvaient recontacter les jeunes pour leur proposer d'adhérer au programme jusqu'au 31 décembre 2008.

L'objectif de cette expérimentation était d'évaluer la plus-value de l'accompagnement renforcé dispensé par les OPP par rapport aux prestations de droit commun, dispensées par Pôle emploi ou ses co-traitants (notamment l'association pour l'emploi des cadres, APEC) en termes d'accès à l'emploi et de qualité de l'emploi retrouvé, le cas échéant. Le protocole d'orientation des jeunes vers la prestation d'accompagnement renforcé a été conçu de sorte à pouvoir comparer les trajectoires des jeunes ayant bénéficié de cette prestation à celles de jeunes de caractéristiques équivalentes n'en ayant pas bénéficié, en contrôlant au mieux les effets de sélection.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Si la sélection des jeunes à orienter vers les opérateurs privés de placement avait été laissée à la discrétion du conseiller de Pôle emploi ou du jeune lui-même, les individus bénéficiant du programme et ceux n'en bénéficiant pas auraient été *a priori* différents et leurs trajectoires n'auraient pu être précisément comparées. En effet, les jeunes orientés pourraient être sélectionnés sur la base de critères non observables (la motivation par exemple ou l'ampleur des difficultés rencontrées). Pour éviter ce problème, l'expérimentation a utilisé l'assignation aléatoire pour déterminer quels jeunes seraient ou non orientés vers la prestation. De la sorte, les individus orientés vers la prestation (désignés sous la terminologie "groupe traité") et ceux non orientés (composant le "groupe de contrôle") ont *a priori* les mêmes caractéristiques, dès lors que l'on dispose d'échantillons suffisamment grands[4].

Le protocole a également prévu de pouvoir évaluer les effets d'externalités du programme. Si le nombre de postes disponibles est en partie fixé, il est possible que le fait d'aider un demandeur d'emploi réduise les chances des autres. Afin de mesurer cet effet, les agences locales d'emploi (ALE) ont été regroupées par quintuplés homogènes au sein de chaque région. Ensuite, un tirage aléatoire a assigné à chacune des cinq ALE de tout quintuplé la proportion de jeunes éligibles qui seraient assignés au traitement, respectivement : 0 %, 25 %, 50 %, 75 % et 100 %. Les zones dans lesquelles aucun jeune n'est assigné au programme peuvent ainsi servir de référence pour évaluer les effets d'externalité. Cet aspect de l'évaluation n'est pas détaillé ici car l'exercice que nous proposons dans cet article ne concerne que la sélection aléatoire des jeunes dans le programme. Toutefois, la prise en compte de ces effets de diffusion complexifie également la comparaison d'estimateurs menée dans le cadre de cet article et nous a conduits à mener une réflexion approfondie sur les hypothèses identifiantes de chacun d'entre eux.

Données et statistiques descriptives

Description des sources

Les données mobilisées pour l'évaluation proviennent de trois sources. Tout d'abord, les fichiers de demandeurs d'emploi fournis par Pôle emploi à la Dares pour réaliser l'orientation des jeunes vers la prestation comprennent un nombre important d'informations sur le profil des jeunes demandeurs d'emploi, qu'ils aient été accompagnés par un OPP ou par Pôle emploi : âge, catégorie de demande d'emploi et ancienneté d'inscription sur les listes, niveau et spécialité des deux derniers diplômes tentés ou obtenus, ALE de rattachement, commune de résidence, code Rome[5] du métier recherché. Ensuite, les fiches de suivi des jeunes pris en charge par les OPP permettent de savoir si les jeunes ont ou non été accompagnés par un OPP.

L'information sur l'insertion dans l'emploi est fournie par une enquête réalisée pour le compte de la Dares par un institut d'enquête. Cette enquête visait à interroger les 35 000 jeunes éligibles à la prestation d'insertion, qu'ils aient ou non été aléatoirement orientés vers un OPP. Afin d'éviter les effets de montée en charge ou d'anticipation de la fin du dispositif, seuls ont été interrogés les jeunes appartenant aux dix cohortes médianes sur les 14 cohortes pour lesquelles une procédure de sélection aléatoire a été opérée (cohortes entrées dans le dispositif entre octobre 2007 et juillet 2008). L'enquête comporte quatre vagues d'interrogations : 8 mois après le tirage de l'échantillon dans la base de Pôle emploi, 12 mois après, 16 mois après et 20 mois après. Afin de favoriser un taux de réponse élevé et de limiter les coûts, le questionnaire de cette enquête est très court (moins de 10 minutes pour la première interrogation, moins de 5 minutes pour les trois suivantes) et un triple mode d'interrogation a été retenu : courrier électronique et/ou courrier postal, puis relances téléphoniques en cas de non-réponse aux modes précédents. Les premières interrogations se sont déroulées entre le mois d'août 2008 et le mois de mai 2009. La quatrième vague, correspondant à la dernière interrogation, s'est déroulée de septembre 2009 à l'été 2010. Les jeunes ont été interrogés à chaque vague, qu'ils aient ou non répondu à la vague précédente. Le taux de réponse a été globalement élevé : entre 70 % et 79 % selon les vagues. Le taux de réponse dépend peu du statut vis-à-vis de l'accompagnement renforcé. Ainsi, le taux de réponse à la première vague pour les individus orientés vers les OPP est égal 79,1 %, contre 78,2 % pour les individus du groupe de contrôle.

L'enquête a permis de collecter des données sur la situation professionnelle au moment de l'enquête (type de contrat, temps complet ou partiel, salaire, catégorie d'emploi) et, en rétrospectif, au moment du tirage dans la base de Pôle emploi. Elle comporte des questions sur le diplôme, l'expérience en accompagnement, les raisons pour lesquelles les jeunes ont accepté ou refusé un accompagnement par l'OPP, le nombre d'enfants et l'âge du dernier d'entre eux.

Statistiques descriptives

Environ 23 000 jeunes ont répondu aux enquêtes de suivi permettant d'évaluer l'impact du dispositif. Parmi eux, 4 467 se situaient dans des zones témoin où aucun jeune n'avait été assigné au traitement. Pour pouvoir comparer les estimations effectuées dans le cadre expérimental à celles obtenues par appariement, il est nécessaire de restreindre l'étude aux 18 853 jeunes restants[6]. La sélection de ces zones ayant été effectuée de manière aléatoire, cette

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

restriction de l'échantillon n'affecte pas la structure globale de la population étudiée.

Le tableau 1 résume les caractéristiques principales de cette dernière. La colonne (1) présente les caractéristiques moyennes des jeunes qui n'ont pas été orientés vers les OPP, tandis que les colonnes (2) et (3) distinguent, parmi la population orientée vers les OPP, ceux qui ont effectivement été accompagnés et ceux qui ne l'ont pas été. La colonne (4) présente les caractéristiques moyennes de l'ensemble de la population orientée vers les OPP. L'assignation aléatoire assure que les populations orientée et non-orientée, en moyenne, sont bien comparables. Les différences observées entre ces deux populations sont de fait assez faibles. La colonne (7) permet de vérifier dans quelle mesure cette hypothèse est vérifiée. Hormis les variables de région utilisées pour la stratification, assez peu de variables présentent une différence statistiquement significative à 5 %.

L'attrition est susceptible d'induire des différences entre le groupe assigné au traitement et le groupe de contrôle si elle les affecte différemment. Dans ce cas, elle constituerait une source de biais dans l'approche

Tableau 1 : statistiques descriptives

	Jeunes non orientés vers des OPP	Jeunes orientés vers des OPP			Différence	Taux d'adhésion	Différence
		Accompagnés					
		Oui	Non	Tous	(3)-(2)		(1)-(4)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Niveau de diplôme							
Doctorat	1,6 %	1,0 %	1,4 %	1,3 %	0,004*	28,1%	0,003*
Master (DEA, DESS)	13,7 %	14,2 %	11,6 %	12,5 %	-0,026***	39,6%	0,013**
École d'ingénieur ou de commerce	2,4 %	2,3 %	2,1 %	2,2 %	-0,002	36,6%	0,002
Maîtrise	7,7 %	8,1 %	7,6 %	7,7 %	-0,006	36,4%	0,000
Autre Bac+4 ou Bac+5	3,2 %	4,2 %	3,5 %	3,7 %	-0,006*	38,7%	-0,005*
Licence	18,6 %	18,1 %	18,5 %	18,4 %	0,004	34,3%	0,002
DEUG - DEUST	2,9 %	3,1 %	3,1 %	3,1 %	0,000	34,8%	-0,002
BTS - DUT	36,0 %	37,3 %	38,0 %	37,8 %	0,008	34,4%	-0,017**
Autre Bac+2 ou Bac+3	9,4 %	8,8 %	9,4 %	9,2 %	0,006	33,4%	0,002
Inférieur à Bac+2	4,5 %	3,0 %	4,8 %	4,2 %	0,018***	24,9%	0,003
Région							
Île-de-France	7,0 %	4,9 %	7,6 %	6,6 %	0,027***	25,6%	0,004
Picardie	6,6 %	8,3 %	7,2 %	7,6 %	-0,011**	38,0%	-0,010**
Haute-Normandie	4,0 %	3,5 %	5,4 %	4,8 %	0,019***	26,0%	-0,007**
Centre	10,7 %	7,0 %	8,9 %	8,2 %	0,019***	29,6%	0,024***
Nord-Pas-de-Calais	17,9 %	28,9 %	24,6 %	26,1 %	-0,043***	38,6%	-0,082***
Lorraine	7,3 %	6,5 %	6,2 %	6,3 %	-0,003	36,0%	0,010***
Pays de Loire	10,8 %	7,3 %	10,4 %	9,3 %	0,031***	27,3%	0,015***
Rhône-Alpes	19,7 %	20,8 %	16,8 %	18,2 %	-0,040***	39,9%	0,015**
Provence-Alpes-Côte d'Azur	13,6 %	10,5 %	9,7 %	10,0 %	-0,008	36,6%	0,036***
La Réunion	2,4 %	2,3 %	3,3 %	2,9 %	0,010***	27,1%	-0,006**
Homme	33,0 %	37,7 %	35,1 %	36,0 %	-0,026***	36,5%	-0,030***
En couple	47,0 %	43,0 %	48,9 %	46,8 %	0,059***	32,0%	0,001
Âge du dernier enfant							
Moins d'un an	7,1 %	4,7 %	8,5 %	7,1 %	0,038***	22,7%	-0,001
Un an ou plus	11,0 %	11,0 %	11,4 %	11,3 %	0,004	34,0%	-0,003
Situation initiale							
En emploi	44,8 %	30,8 %	50,9 %	43,9 %	0,201***	24,4%	0,009
Sans emploi	41,9 %	54,3 %	36,7 %	42,8 %	-0,176***	44,2%	-0,009
Inconnue	13,3 %	14,9 %	12,4 %	13,3 %	-0,025***	39,1%	0,000
Durée de chômage							
Dernier épisode	9,6 mois	9,1 mois	9,4 mois	9,1 mois	0,261*	-	0,363***
Cumulée sur les 18 derniers mois	11,2 mois	11,0 mois	11,0 mois	11,0 mois	-0,013	-	0,172***
Âge	25 ans 7 mois	25 ans 7 mois	25 ans 6 mois	25 ans 7 mois	-0,075*	-	0,081*
Nombre d'observations	5 812	4 541	8 500	13 041	-	34,8%	-

Lecture : * = différence significative à 10 % ; ** = différence significative à 5 % ; *** = différence significative à 1 %. DEA = diplôme d'études approfondies ; DESS = diplôme d'études supérieures spécialisées ; DEUG = diplôme d'études universitaires générales ; DEUST = diplôme d'études universitaires scientifiques et techniques ; BTS = brevet de technicien supérieur ; DUT = diplôme universitaire de technologie.

Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.

Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

expérimentale[7]. Il est toutefois possible de tester le caractère aléatoire de l'attrition au sein de ces deux groupes. Pour cela, on peut étudier un modèle probit visant à expliquer le fait de rester dans l'échantillon par l'interaction de la variable d'assignation avec les variables explicatives. Un test de nullité jointe des coefficients estimés ne permet pas de rejeter l'hypothèse de nullité (correspondant à l'attrition aléatoire), la *P-value* associée étant égale à 0,27.

Parmi la population orientée vers le programme, on constate en revanche des différences sensibles entre les individus bénéficiant du programme et ceux n'en bénéficiant pas. La colonne (5) confirme la significativité de ces écarts, établie par des tests d'égalité des moyennes. Ces différences se reflètent dans les taux de participation présentés dans la colonne (6), soit la proportion de jeunes ayant bénéficié du programme parmi les individus assignés, pour chacun des sous-groupes. Logiquement, des écarts importants dans la colonne (5) sont associés à des taux d'adhésion anormalement hauts ou bas, par rapport à la moyenne égale à 34,8 % sur l'ensemble des individus assignés à l'accompagnement renforcé.

Le tableau 1 nous renseigne d'abord sur les caractéristiques avant traitement de la population éligible. Parmi les 18 853 jeunes de l'échantillon, seulement un tiers sont des hommes. Un peu moins de la moitié vivent en couple et 15 % à 20 % des individus ont au moins un enfant. Le Nord-Pas-de-Calais et la région Rhône-Alpes sont les régions les plus représentées : elles constituent à elles seules environ 40 % de l'échantillon total. La Réunion, qui connaît une situation d'emploi tout à fait spécifique, ne représente quant à elle que 2 % à 3 % des jeunes identifiés dans le cadre de l'expérimentation. La moyenne d'âge se situe entre 25 et 26 ans. Ces jeunes ont connu en moyenne 11 mois de chômage au cours des 18 derniers mois. La durée du dernier épisode de chômage étant en moyenne de neuf mois et demi, il est probable que, dans une majorité de cas, les épisodes de chômage soient relativement longs et entrecoupés de périodes d'emploi. Une très vaste majorité d'individus de l'échantillon ne possède pas de diplôme supérieur à la licence. Il est particulièrement inquiétant de constater que près de 40 % des individus ont un brevet de technicien supérieur (BTS) ou un diplôme universitaire de technologie (DUT), des diplômes réputés professionnalisants, visant une entrée rapide sur le marché du travail.

La participation au programme dépend fortement de la région d'origine des jeunes. Alors que la Picardie, le Nord-Pas-de-Calais et la région Rhône-Alpes sont associés à des taux de participation élevés, d'autres régions (Île-de-France, Haute-Normandie, région Centre, région Pays de Loire, Réunion) ont des taux de participation particulièrement faibles. Ces disparités sont susceptibles de s'expliquer à la fois par des situations économiques contrastées et par des différences attribuables à la manière dont l'OPP a contacté et tenté de convaincre les jeunes d'entrer dans le programme.

La situation professionnelle au moment de l'assignation aléatoire au traitement est la variable jouant le rôle le plus important sur la probabilité de bénéficier du programme. En effet, il s'est avéré qu'environ 45 % des jeunes assignés au programme étaient déjà en emploi au moment du tirage au sort. Les demandeurs d'emploi éligibles au programme pouvaient par exemple être en activité réduite ou ne pas encore avoir été enregistrés comme en emploi dans la base de gestion de Pôle emploi au moment du processus de sélection. Les jeunes initialement en emploi avaient pratiquement deux fois moins de chance de participer au programme d'accompagnement renforcé que ceux se déclarant initialement sans emploi.

Les jeunes finalement accompagnés ont également tendance à être plus diplômés que les autres, à l'exception notable des titulaires de doctorat. Ces jeunes sont plus souvent des hommes, ne vivant pas en couple et n'ayant pas d'enfant de moins d'un an. La durée du dernier épisode de chômage diffère également, tandis que la durée de chômage cumulée sur les 18 derniers mois est similaire pour les deux groupes.

Identifier l'effet du traitement

Cette étude propose de comparer les résultats expérimentaux obtenus en comparant les groupes constitués aléatoirement et des résultats que l'on aurait pu obtenir si l'expérimentation n'avait pas eu lieu[8]. La constitution des deux groupes est un élément clef et dépend du contexte de l'expérimentation. Dans le cas de LaLonde (1986), l'impact du programme est obtenu directement par la comparaison des deux groupes expérimentaux. Les groupes de contrôle non-expérimentaux proviennent de deux sources externes : la *Current Population Survey* (CPS) du *Bureau of Census* et le *Panel Study of Income Dynamics* (PSID) de l'université du Michigan, deux enquêtes fréquemment utilisées aux États-Unis. Cette approche présente cependant deux inconvénients. Tout d'abord, les variables ne sont pas définies et collectées de manière homogène entre les échantillons expérimentaux et non-expérimentaux. De plus, quelques individus du groupe non expérimental sont potentiellement bénéficiaires du programme. Heckman *et alii* (1998a) utilisent, quant à eux, un groupe non expérimental dont les données sont collectées parallèlement à celles des groupes expérimentaux. Dans leur cas, les individus éligibles au programme avaient le choix de postuler ou non au programme :

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

ceux qui ont présenté une candidature et dont la candidature a été retenue par les organisateurs du programme ont été tirés au sort. Quant aux personnes éligibles ayant refusé de participer, elles ont constitué la population non expérimentale de comparaison.

Nous nous trouvons dans une situation légèrement différente. Dans l'expérimentation "jeunes diplômés", les individus sont tirés au sort avant que leur assentiment à la participation au dispositif soit connu. L'assignation est aléatoire mais pas la participation, qui relève de la décision de chaque jeune. La figure A1 en annexe présente un schéma des groupes de l'expérimentation. On appelle Z la variable qui vaut 1 pour les individus assignés au groupe de traitement et 0 pour les individus assignés au groupe de contrôle. Le groupe de gauche représente la population assignée au traitement (Z=1). Parmi ceux-ci, la partie A est traitée et la partie B ne l'est pas. Les groupes C et D ne sont pas assignés ($Z = 0$) et ne sont pas traités car, dans notre cas, aucun individu non assigné n'a pu avoir accès au programme. Le groupe C représente les individus qui seraient entrés dans le programme si on le leur avait proposé, tandis que le groupe D n'aurait pas accepté le traitement. Comme nous allons le développer ci-dessous, l'effet moyen du traitement sur les traités est la différence entre les groupes A et C. Du fait du tirage au sort, cet effet moyen du traitement peut être déduit de la différence globale entre assignés et non-assignés. Que se serait-il passé en l'absence d'expérimentation ? Les groupes A et C auraient accepté le programme et les groupes B et D l'auraient refusé. Un économètre désireux d'évaluer le dispositif aurait donc dû se contenter de comparer A et C d'une part, à B et D d'autre part. Comme le tirage au sort rend ces groupes identiques deux à deux, cela revient à comparer les groupes A et B. Dans l'approche non expérimentale, nous effacerons donc les données relatives aux groupes C et D pour ne comparer que les groupes A et B.

L'identification de l'effet du traitement repose, pour chaque méthode, sur une série d'hypothèses. Selon la méthode retenue, les hypothèses sont plus ou moins fortes. Dans cette partie, nous rappelons les hypothèses nécessaires à l'identification de l'effet du traitement dans le cas d'une expérimentation aléatoire et dans le cas d'un appariement sur caractéristiques observables. Nous noterons que les deux méthodes reposent sur certaines hypothèses, en insistant sur leurs différences. Pour les introduire, il est utile de rappeler brièvement le contexte d'évaluation selon le cadre de modélisation proposé par Rubin (1974) et développé par Imbens et Angrist (1994) (dit cadre AIR).

Dans ce cadre, on suppose pour chaque individu l'existence des variables latentes Y_i^0 et Y_i^1, qui reflètent, respectivement, le résultat auquel aboutit l'individu i dans le cas où il est traité ($T_i = 1$) et celui dans le cas où il ne l'est pas ($T_i = 0$). Notre paramètre d'intérêt est l'effet moyen du traitement sur les individus traités, ou *Average Treatment Effect on the Treated* (*ATT*), défini comme $ATT = E[Y_i^1 - Y_i^0 \mid T_i = 1]$. Cet effet ne peut jamais être directement mesuré, puisque, par définition, Y_i^0 n'est observé que lorsque l'individu i ne reçoit pas le traitement, tandis que Y_i^1 n'est observé que lorsqu'il le reçoit. L'avantage principal du cadre AIR est sa généralité. En particulier, il n'est pas nécessaire de faire l'hypothèse que l'effet du traitement est constant.

Il faut en outre distinguer le traitement T de l'assignation au traitement Z. On définit pour chaque individu i deux variables latentes de traitement, T_i^1 et T_i^0. T_i^1 vaut 1 si l'individu i est effectivement traité et 0 sinon, conditionnellement au fait d'être assigné au groupe de traitement (*i.e.* sachant $Z_i = 1$). Quant à T_i^0 il vaut également 1 si l'individu i est effectivement traité et 0 sinon, conditionnellement au fait, cette fois, d'être assigné au groupe de contrôle (*i.e.* sachant $Z_i = 0$). Dans notre cas, compte tenu de ce qui a été dit *supra*, T_i^0 est toujours nul, puisqu'aucun individu assigné au groupe de contrôle n'a été finalement traité.

Lors d'une expérimentation contrôlée, l'identification repose principalement sur le caractère exogène du tirage au sort. Cependant, il ne s'agit pas de la seule hypothèse. Nous en retenons ici quatre.

Hypothèse 1 (indépendance de l'assignation) : l'assignation Z_i est indépendante des variables latentes Y_i^1, Y_i^0, T_i^1 et T_i^0.

Hypothèse 2 (monotonie de l'effet de l'assignation) : pour tous les individus, $T_i^1 \geq T_i^0$.

Hypothèse 3 (exclusion) : l'assignation Z_i n'a d'influence sur les variables de résultat Y_i^1 et Y_i^0 que *via* le traitement T_i.

Hypothèse 4 (*stable unit treatment value assumption, SUTVA*) : le résultat potentiel Y_i^1, Y_i^0 n'est pas affecté par le statut de traitement des autres individus $j \neq i$.

L'hypothèse 1 dépend de la qualité du tirage au sort. Si le tirage au sort a été réalisé sans encombre, cette hypothèse est réalisée par construction. Il n'existe pas de test formel de cette hypothèse, puisqu'elle porte sur des quantités qui ne sont pas observables. En revanche, il est fréquent de montrer que la distribution des caractéristiques observables est indépendante de l'assignation, ce qui représente une indication de la validité du tirage au sort. L'hypothèse 2 est clairement vérifiée dans notre cas, car aucun individu ayant été assigné au groupe de

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

contrôle n'a pu bénéficier du traitement ($T_i^0 = 0, \forall i$). Les opérateurs privés de placement ne pouvaient contacter que les jeunes de la liste assignée (et n'avaient pas accès à d'autres noms). L'hypothèse 3 est plus difficile à vérifier et doit être appréhendée au cas par cas. Ici, une violation de cette hypothèse reviendrait à imaginer que le fait d'être contacté pour participer à un programme a un impact intrinsèque sur la réussite professionnelle, indépendamment de la participation effective au programme. Nous pensons raisonnable de supposer que cette hypothèse est ici vérifiée.

L'hypothèse 4 est également difficile à vérifier. En l'occurrence, une analyse approfondie menée dans Crépon *et alii* (2013) montre qu'elle est invalidée, en raison de l'existence d'effets d'équilibre général. Admettons que le nombre d'emplois vacants localement soit fixe. Si traiter un sous-ensemble des demandeurs d'emploi accélère le retour à l'emploi de ceux-ci, cela ralentit le retour à l'emploi des demandeurs d'emploi non traités. Il reste en effet à ces derniers un nombre de vacances plus faible qu'en l'absence du dispositif. Dans ce qui suit, nous remplaçons donc l'hypothèse 4 par l'hypothèse suivante :

Hypothèse 5 (externalité constante) : le résultat potentiel Y_i^0 (resp. Y_i^1) peut différer selon que l'individu *i* se trouve dans un marché où des individus sont traités ou dans un marché où aucun ne l'est. Toutefois, le résultat potentiel est constant quelle que soit la proportion d'individus traités.

Dans l'expérimentation que nous utilisons ici, la proportion d'individus assignés pouvait être égale à 0 %, 25 %, 50 %, 75 % ou 100 %. Nous supposons maintenant que le résultat potentiel de l'individu *i* dépend de cette proportion, de manière simple : $Y_i^\ell = Y_i^\ell(0)$ si *i* se trouve dans une zone assignée à 0 % et $Y_i^\ell = Y_i^\ell(1)$ dans le cas contraire. En dépit de sa simplicité, cette hypothèse a deux mérites importants. D'une part, elle représente, dans une littérature empirique dominée par l'hypothèse SUTVA, une amélioration substantielle. D'autre part, elle simplifie l'analyse : pour faire abstraction de l'externalité, il est simplement nécessaire de comparer les individus assignés et non-assignés au sein des zones traitées. Nous restreignons donc l'échantillon aux zones dans lesquelles la proportion d'individus traités est strictement positive. Pour simplifier les notations, Y_i^ℓ s'entend dans la suite de cette étude comme $Y_i^\ell(1)$, avec $\ell = 0$ ou 1. Finalement, notons que cette hypothèse n'est pas rejetée par les données que nous utilisons, comme le montre le tableau 4 de Crépon *et alii* (2013).

Sous ces quatre hypothèses 1, 2, 3 et 5, l'estimateur de Wald permet de calculer l'effet moyen du traitement sur les traités. Dans notre cas, l'estimateur de Wald *W* est égal à la différence des moyennes empiriques de la variable de résultat prises sur les groupes assigné et non assigné au traitement, rapportée au taux d'adhésion moyen du groupe assigné au traitement[9] :

$$W = \frac{\hat{E}[Y_i | Z_i = 1] - \hat{E}[Y_i | Z_i = 0]}{\hat{P}[T_i | Z_i = 1]}$$

où $\hat{E}$ désigne la moyenne empirique et $\hat{P}$ une proportion empirique, en se souvenant que $\hat{P}[T_i | Z_i = 0] = 0$ dans notre cas.

En pratique, cet estimateur de Wald peut être calculé en estimant un modèle linéaire de la variable de résultat en fonction du traitement, ce dernier étant instrumenté par l'assignation[10]. Sous l'hypothèse 1, l'assignation est également indépendante des caractéristiques observables. Inclure des variables de contrôle supplémentaires dans l'équation principale est donc asymptotiquement neutre sur la valeur du coefficient estimé et peut être intéressant pour augmenter la puissance de l'expérimentation. Nous montrons que, dans notre échantillon, l'inclusion de variables de contrôle a un effet totalement négligeable sur la valeur des effets estimés (*cf. infra*, partie "Résultats", tableaux 2 et A1).

En l'absence d'une variable d'assignation aléatoire ou quasi-aléatoire qui permettrait d'appliquer le raisonnement décrit ci-dessus, il faut adopter des méthodes reposant sur des hypothèses additionnelles, reposant sur un jeu de caractéristiques individuelles observables qui expliquent suffisamment bien la variable de résultat *Y*. En particulier, les estimateurs d'appariement reposent sur les hypothèses suivantes.

Hypothèse 6 (indépendance conditionnelle) : les variables latentes de résultat sont indépendantes (en espérance) du statut de traitement, conditionnellement aux caractéristiques observables *X* :

$E[Y^0 | T, X] = E[Y^0 | X]$ et $E[Y^1 | T, X] = E[Y^1 | X]$

Hypothèse 7 (support commun) : il n'existe pas de valeur des caractéristiques observables telle qu'un individu ayant ces caractéristiques soit certain d'être traité : $P[T_i = 1 | X_i] < 1$.

L'hypothèse 6 est la plus exigeante. Elle suppose que, pour un sous-ensemble d'individus ayant les mêmes caractéristiques, le statut de traitement est orthogonal aux variables latentes de résultat, c'est-à-dire que le traitement est aléatoire. Cette hypothèse, qui n'est pas testable, est difficile à admettre dans tous les cas où des variables inobservables ayant un impact sur les variables de

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

résultat sont susceptibles d'avoir influencé la décision d'être traité. La motivation, par exemple, est souvent citée comme un facteur inobservable corrélé à la fois au fait de bénéficier d'un programme et à la réussite professionnelle.

L'hypothèse 7 est plus technique. Elle suppose que les caractéristiques ne sont pas des prédicteurs parfaits du fait d'être traité. Cette hypothèse, qui sert à garantir que l'on puisse disposer d'un groupe de contrôle non vide pour chaque individu traité, se vérifie empiriquement.

Sous les hypothèses 5, 6 et 7, on peut montrer que le paramètre d'intérêt s'exprime comme une différence de deux paramètres identifiables :

$$ATT = E[Y_i | T_i = 1] - E\{E[Y_i | T_i = 0, X_i] | T_i = 1\}$$

Le premier terme de cette différence est l'espérance de la variable de résultat pour le groupe traité. Le second terme est une moyenne théorique pondérée de la variable de résultat du groupe non traité, la pondération garantissant que la distribution pondérée des observables du groupe non traité se retrouve égale à celle du groupe traité.

Nous privilégions les méthodes d'appariement par rapport aux régressions linéaires pour deux raisons : leur flexibilité fonctionnelle et leur capacité à mettre en évidence les problèmes éventuels de support des caractéristiques observables. Comme l'analyse inclut un nombre important de caractéristiques observables, l'appariement sur score de propension est de plus préféré à l'appariement exact[11]. Il s'agit alors d'estimer :

$$ATT = E[Y_i | T_i = 1, P_i] - E\{E[Y_i | T_i = 0, P_i] | T_i = 1\}$$

où le score de propension $P_i = P_i(X_i) = P[T_i = 1 | X_i]$ exprime la probabilité pour qu'un individu choisisse de participer au programme étant données ses caractéristiques *ex ante*. Smith et Todd (2005)[12] montrent que les estimateurs d'appariement peuvent alors tous s'écrire sous la forme :

$$att = \frac{1}{n_1} \sum_{i \in I_1 \cap S_p} \left[Y_i^{T=1} - \sum_{j \in I_0} w(i,j) Y_j^{T=0} \right]$$

où I_1 représente l'ensemble des individus ayant reçu le traitement, I_0 l'ensemble des individus ne l'ayant pas reçu et S_p le support du score de propension estimé p. n_1 est le nombre d'individus de $I_1 \cap S_p$. La pondération $w(i, j)$ dépend de la distance entre p_i et p_j (les valeurs estimées du score de propension pour les individus i et j), et de la méthode retenue.

L'appariement sur score de propension repose donc sur une estimation en deux temps. On estime tout d'abord un score de propension, fonction des caractéristiques X. Dans un second temps, on compare les situations des individus traités et non traités *ex post* en utilisant une pondération particulière des individus non traités. Dans cet article, nous avons estimé les scores de propension selon des régressions de type probit. Nos principaux résultats sont par ailleurs obtenus en utilisant une méthode de régression Ridge permettant notamment de réduire l'impact des observations pour lesquelles le score estimé est associé à une faible densité (Seifert et Gasser, 1996). Frölich (2004) montre en effet que cette méthode permet d'obtenir des estimations plus fiables à distance finie que d'autres méthodes plus usuelles d'estimation. Toutefois, nous montrons que les résultats obtenus selon cette méthode d'estimation sont extrêmement proches de ceux obtenus par régression linéaire locale (Fan, 1992 ; Heckman *et alii*, 1998b) et selon la méthode de Nadaraya-Watson (*cf.* partie "Résultats", tableau A4)[13].

Résultats

Résultats expérimentaux

Le tableau 2 présente le résultat de l'estimation expérimentale. Chaque ligne du tableau correspond à l'une des quatre variables expliquées : le taux d'emploi, le taux d'emploi durable et les deux composantes de ce dernier : les CDD d'au moins six mois et les CDI. Chaque colonne du tableau correspond à un champ différent. Dans les colonnes (1) et (2), nous gardons la totalité des individus. Dans les deux suivantes, nous nous restreignons à la sous-population des individus qui n'avaient pas d'emploi à la date de l'assignation aléatoire. Enfin, nous segmentons cette dernière population par sexe : les hommes (colonnes (5) et (6)) et les femmes (colonnes (7) et (8)). Les colonnes impaires correspondent à la totalité des régions, tandis que le champ des colonnes paires est restreint aux six régions dans lesquelles l'OPP était marchand.

Selon les résultats obtenus, l'impact du dispositif sur l'emploi est en général non significativement différent de zéro. La probabilité d'être en emploi durable huit mois après l'assignation aléatoire n'est pas significativement augmentée par le programme lorsque l'on considère l'ensemble des individus. Si l'on se restreint à la sous-population des individus qui n'avaient pas d'emploi au moment de l'assignation, on obtient cependant un effet de 5 points de pourcentage (pp), significatif à 5 %. Cet effet est deux fois plus important, de l'ordre de 9,5 pp si l'on se restreint aux régions dans lesquelles les OPP étaient marchands. Ce résultat est économiquement important, puisque, dans les régions dont l'OPP est marchand, l'accompagnement renforcé ferait en moyenne passer le taux d'emploi des individus sans emploi initialement de 47,0 % à 56,5 %. Lorsque l'on

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Tableau 2 : effet de l'accompagnement sur l'accès à l'emploi - résultats expérimentaux

Individus	Tous		Sans emploi initialement		Sans emploi initialement - hommes		Sans emploi initialement - femmes	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Variable de résultat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Emploi	0,022	0,002	0,039*	0,008	0,047	0,085	0,041	-0,012
	(0,024)	(0,042)	(0,027)	(0,052)	(0,037)	(0,070)	(0,035)	(0,060)
Emploi durable	0,013	0,028	0,050**	0,095**	0,076**	0,176**	0,039	0,074
	(0,024)	(0,043)	(0,029)	(0,056)	(0,043)	(0,079)	(0,034)	(0,063)
CDD de six mois ou plus	0,025*	0,058**	0,058***	0,109***	0,116***	0,257***	0,028	0,029
	(0,019)	(0,035)	(0,022)	(0,041)	(0,035)	(0,079)	(0,027)	(0,045)
CDI	-0,012	-0,029	-0,008	-0,015	-0,040	-0,081	0,011	0,045
	(0,020)	(0,038)	(0,025)	(0,047)	(0,041)	(0,081)	(0,031)	(0,062)
Nombre d'observations	18 853	7 109	8 018	2 982	3 001	1 112	5 017	1 870

Lecture : l'impact est le coefficient relatif à l'indicatrice d'avoir été accompagné(e) par un OPP. Cette variable est instrumentée par le fait d'avoir été assigné(e) à l'accompagnement. Le modèle estimé comprend également le sexe, l'ancienneté au chômage et son carré, le diplôme et des indicatrices relatives aux quintuplés de l'agence locale d'emploi (ALE). Les écarts-types (entre parenthèses) sont corrigés de la corrélation entre les observations d'une même ALE. * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.

Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi, calculs des auteurs.

Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés".

décline ce résultat par sexe, l'impact semble plus significatif pour les hommes que pour les femmes (pour lesquelles le traitement ne paraît pas avoir d'impact).

L'impact positif du dispositif, lorsqu'il est observé, est-il dû à une augmentation des CDD de six mois ou plus ou à une augmentation des CDI ? Le tableau 2 montre que la totalité de l'impact sur l'emploi durable passe par une augmentation des CDD de six mois ou plus. L'impact sur les CDI n'est quant à lui jamais significativement différent de zéro. L'impact sur le fait d'être en CDD de six mois ou plus, huit mois après l'assignation, est proche de celui sur l'emploi durable, en plus important et plus significatif statistiquement. L'effet est cette fois détectable sur les individus dans leur ensemble et plus net sur ceux qui n'ont pas d'emploi initialement, essentiellement sur les hommes. Les régions avec OPP marchands semblent conduire la totalité de cet effet.

Des interprétations plus précises et une analyse des externalités sont disponibles dans Crépon *et alii* (2011) et Crépon *et alii* (2013).

Estimation par appariement

Nous nous tournons maintenant vers l'exercice non expérimental, en restreignant l'échantillon à la population assignée au programme. L'objectif est de comparer les résultats obtenus par appariement aux résultats expérimentaux.

La première étape consiste à estimer le score de propension qui sera utilisé pour l'appariement. Sa construction doit répondre à plusieurs critères. Pour que l'estimation de l'effet du programme soit fiable, il faut tout d'abord rendre compte efficacement de la décision de participation. Pour cela, il importe d'inclure de nombreuses variables pouvant expliquer les choix individuels et de vérifier la qualité explicative du modèle retenu.

Le tableau 3 présente les différentes spécifications de score que nous avons envisagées et leurs principales caractéristiques. Certaines variables ont été incluses de manière systématique et retracent les éléments les plus essentiels de la situation objective des individus : le sexe, le nombre de mois passés au chômage au cours des 18 derniers mois, la situation initiale d'emploi, l'âge, le fait d'être en couple et le niveau de diplôme. Nous avons également inclus des indicatrices par cohorte pour contrôler d'éventuelles variations conjoncturelles. La spécification (1) représente la spécification la plus simple que nous avons testée. Elle ne fait intervenir aucune interaction entre variables et ne contrôle les effets fixes spatiaux que par le biais d'indicatrices régionales. Lorsque nous intégrons les quintuplés d'ALE et non plus les régions pour rendre compte des variations spatiales (spécification (2)), le modèle semble gagner en qualité. Les spécifications (3), (4) et (5) intègrent de plus des variables d'interaction, notamment avec le sexe. En outre, la spécification (5) prend en compte le secteur dans lequel les individus recherchent un emploi, identifié par deux positions du code Rome. Ces éléments n'ont pas tous le même impact sur la qualité du score estimé. Nous obtenons ainsi des résultats très proches dans les spécifications (3) et (4), bien que cette dernière comprenne une différenciation du niveau de diplôme selon le sexe. La spécification (5) permet, quant à elle, d'augmenter un peu le pseudo-R^2, mais seulement au prix de l'ajout d'un grand nombre de variables, puisqu' y sont introduits à la fois l'interaction du niveau de diplôme avec la situation initiale d'emploi et le domaine d'emploi recherché.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Tableau 3 : différentes spécifications du score de propension et qualité de l'ajustement

	(1)	(2)	(3)	(4)	(5)
Nombre de mois au chômage sur les 18 derniers mois	Oui	Oui	Oui	Oui	Oui
Au Carré	Oui	Oui	Oui	Oui	Oui
Situation initiale d'emploi	Oui	Oui	Oui	Oui	Oui
Âge	Oui	Oui	Oui	Oui	Oui
Au carré	Oui	Oui	Oui	Oui	Oui
Niveau de diplôme (10 positions)	Oui	Oui	Oui	Oui	Oui
Décomposition hommes/femmes	Non	Non	Non	Oui	Oui
Interaction avec la situation initiale	Non	Non	Non	Non	Oui
Secteur de l'emploi recherché	Non	Non	Non	Non	Oui
Perception d'indemnités	Oui	Oui	Oui	Oui	Oui
Homme	Oui	Oui	Oui	Oui	Oui
En couple	Oui	Oui	Oui	Oui	Oui
Décomposition hommes/femmes	Non	Non	Oui	Oui	Oui
Nombre d'enfants	Non	Non	Oui	Oui	Oui
Décomposition hommes/femmes	-	-	Oui	Oui	Oui
Âge du dernier enfant	Non	Non	Oui	Oui	Oui
Décomposition hommes/femmes	-	-	Oui	Oui	Oui
Régions	Oui	Non	Non	Non	Non
Quintuplés d'ALE	Non	Oui	Oui	Oui	Oui
Cohorte	Oui	Oui	Oui	Oui	Oui
Nombre d'observations	13 041	13 041	13 041	13 041	13 041
Log-vraisemblance	-8 018	-7 967	-7 920	-7 914	-7 861
Pseudo-R^2	0,049	0,055	0,060	0,061	0,067

Lecture : les estimations ont été réalisées selon une régression de type *probit*. Lecture détaillée du tableau : *cf.* corps du texte.
Champ : individus appartenant à une zone traitée, assignés au programme et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Il faut en outre que le score retenu permette de vérifier l'hypothèse 6 d'indépendance conditionnelle. S'il est formellement impossible de s'assurer de la validité de cette hypothèse, deux tests constituent des moyens de vérifier que le score choisi est compatible avec celle-ci. Ces tests peuvent conduire à rejeter *a priori* la compatibilité du jeu de variables observables ou de la spécification choisie. Le premier, proposé par Dehejia et Wahba (2002), repose sur un raisonnement local, pour des valeurs du score $p(X) = \hat{P}(T=1|X) = \hat{E}(T|X)$ proches. La réalisation de ce test comprend deux étapes :

1) les observations sont stratifiées de manière à avoir dans chaque strate des individus bénéficiaires et des individus non bénéficiaires dont les scores ont des valeurs proches. Cela revient à vérifier que les valeurs du score prises sur les individus traités $(p(X)|T=1)$ et non traités $(p(X)|T=0)$ ne sont pas statistiquement différentes dans chacune des strates ;

2) une fois les strates établies, on effectue dans chaque strate et pour chaque variable de contrôle X^k un test d'égalité des moyennes consistant à vérifier que l'on obtient des moyennes pour la variable de contrôle X^k non statistiquement différentes sur le groupe des individus traités et sur le groupe des individus non traités de la strate $\left(\hat{E}[X^k|T=1] \approx \hat{E}[X^k|T=0]\right)$.

Un score est considéré comme non équilibré si, pour un niveau de test α %, on rejette l'hypothèse nulle d'égalité des moyennes des variables de contrôle dans plus de α % des cas lors de cette seconde étape. Nous avons privilégié une approche plus conservatrice pour la première étape du test que pour la seconde, en imposant des tests de niveau 1 % pour la première étape, car la définition des strates est déterminante pour la seconde étape du test. Il est donc central de s'assurer que les valeurs du score sont égales une fois conditionnées par la variable de traitement. Si, suite à un découpage en S strates, il s'avère que les valeurs du score ne sont pas équilibrées dans au moins l'une d'entre elles, on effectue un découpage en $S+1$ strates et on refait le test d'égalité. On ne passe à la seconde étape du test que lorsque celui-ci est validé dans toutes les strates.

Le résultat de ce premier test est reporté dans la partie haute du tableau 4. Parmi les spécifications que nous avons envisagées, la première et la dernière ne peuvent être acceptées selon ce critère. La première spécification n'est pas équilibrée au seuil classique de 5 % lors de la seconde étape du test, tandis que l'algorithme ne parvient pas à identifier des strates de valeurs du score satisfaisantes lors de la première étape pour la cinquième spécification. En effet, ce score est trop polarisant et, au terme des itérations et de l'affinement des strates, nous avons fini par obtenir des strates ne contenant que des individus

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Tableau 4 : tests de l'hypothèse 6 (indépendance conditionnelle)

Spécification du score	(1)	(2)	(3)	(4)	(5)
Test de Dehejia et Wahba (2002)					
Nombre de strates	11	9	11	12	12
Proportion de tests d'égalité des moyennes (seconde étape) aboutissant à un rejet au seuil de 5 %	5,2 %	4,1 %	3,9 %	4,1 %	-
Test de Shaikh *et alii* (2009)					
P-value	0,194	0,387	0,723	0,945	0,992

Champ : individus appartenant à une zone traitée, assignés au programme et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Tableau 5 : effet de l'accompagnement sur l'accès à l'emploi : résultats non expérimentaux

Individus	Tous		Sans emploi initialement		Sans emploi initialement - hommes		Sans emploi initialement - femmes	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Variable de résultat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Emploi	-0,051***	-0,019*	-0,026**	0,014	-0,033*	-0,007	-0,019	0,021
	(0,008)	(0,014)	(0,013)	(0,025)	(0,025)	(0,042)	(0,019)	(0,034)
Emploi durable	-0,091***	-0,055***	-0,071***	-0,026	-0,071***	-0,014	-0,063***	-0,045*
	(0,008)	(0,014)	(0,014)	(0,027)	(0,025)	(0,045)	(0,018)	(0,034)
CDD de six mois ou plus	-0,014**	-0,009	-0,014	-0,007	-0,025	0,002	-0,003	-0,032
	(0,007)	(0,012)	(0,013)	(0,024)	(0,020)	(0,037)	(0,015)	(0,030)
CDI	-0,077***	-0,046***	-0,057***	-0,019	-0,046**	-0,016	-0,060***	-0,012
	(0,007)	(0,014)	(0,014)	(0,024)	(0,021)	(0,039)	(0,017)	(0,026)
Nombre d'observations	13 041	4 891	5 582	2 102	2 152	808	3 430	1 294

Lecture : l'impact est le coefficient relatif à l'indicatrice d'avoir été accompagné(e) par un OPP. Les écarts-types sont reportés entre parenthèses. Ils ont été calculés par *bootstrap* (99 réplications). Ces estimations ont été réalisées par régression Ridge en utilisant le score de propension (3) du tableau 3 et une fenêtre de longueur optimale. * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.
Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

traités ou que des individus non traités avant que la condition d'égalité des valeurs du score conditionnellement à la variable de traitement ne puisse être atteinte.

Ce premier test a fait l'objet de plusieurs critiques, notamment parce qu'il a tendance à être trop conservateur et parce qu'il se heurte au problème bien connu de la réalisation de tests multiples sur un même échantillon. C'est pourquoi nous avons également mis en œuvre le test de spécification introduit par Shaikh *et alii* (2009). Celui-ci repose sur une comparaison plus globale du score estimé pour les individus du groupe traité et du groupe de comparaison. La statistique de ce test est asymptotiquement normale sous l'hypothèse nulle selon laquelle le score a été correctement identifié. Elle diverge dans le cas contraire. Les résultats présentés dans le tableau 4 ne permettent de rejeter aucune des spécifications que nous avons retenues aux niveaux de significativité conventionnels. Toutefois, il est important de noter que la *P-value* obtenue n'est pas uniforme selon les spécifications. Ainsi, elle est maximale pour les quatrième et cinquième spécifications. Elle est également très élevée pour la troisième spécification, qui passe le mieux le test de Dehejia et Wahba (2002). À la lumière de l'ensemble de ces résultats, nous avons donc retenu la spécification (3) comme référence. Toutefois, nos tests de robustesse font également intervenir les autres spécifications (annexe, tableau A3).

Une fois le score estimé, on peut procéder à l'appariement. Les résultats obtenus sont présentés dans les tableaux 5 et A2, qui reprennent la même structure que le tableau 2. Le tableau 5 présente les estimations de l'effet de l'accompagnement sur le taux d'emploi, le taux d'emploi durable, le taux de CDD de six mois ou plus et le taux de CDI. Le tableau A2 en annexe présente quant à lui les écarts entre les estimations tirées des analyses expérimentale et non expérimentale, ainsi que la significativité de ces écarts.

Les résultats obtenus par méthode d'appariement sont globalement éloignés de ceux obtenus en utilisant le groupe de contrôle expérimental. Ainsi, dans le tableau 5, tous les coefficients estimés sont négatifs ou non significativement différents de zéro. Au total, si l'on devait interpréter ces résultats de manière causale, on conclurait que le programme n'a au mieux pas eu d'impact et qu'il a même eu un impact significativement négatif sur certaines sous-populations et certaines variables de résultat. En particulier, l'impact moyen sur l'ensemble de la

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

population et des régions (colonne (1)) est négatif, quelle que soit la mesure d'insertion dans l'emploi utilisée. Le tableau A2 confirme que les écarts entre ces deux jeux d'estimations sont non seulement quantitativement importants, mais aussi statistiquement très significatifs dans une très grande majorité des cas.

Les résultats expérimentaux suggèrent que l'impact positif du programme se situe essentiellement au niveau des chances des individus accompagnés de trouver un CDD de six mois ou plus, quel que soit le sous-groupe étudié (hormis celui des femmes initialement sans emploi, pour lesquelles aucun effet significatif du programme n'est observé). L'analyse par appariement conduit à une estimation globalement négative de l'impact du dispositif sur les chances de trouver un CDD de six mois ou plus (colonne (1)) avec des effets par sous-groupe non significatifs. Dans le tableau 5, l'impact négatif mesuré semble de moindre ampleur et moins souvent significatif pour les régions ayant des OPP marchands que pour l'ensemble de la population. Si ce résultat va dans le même sens que les résultats expérimentaux, les conclusions qualitatives issues de ces estimations demeurent dans l'ensemble radicalement différentes.

Pour nous assurer de la robustesse de ces résultats, nous avons réalisé plusieurs estimations supplémentaires, faisant varier trois critères importants dans la littérature sur les méthodes d'appariement : le score de propension, la méthode d'estimation et le support du score estimé. Dehejia et Wahba (1999) préviennent l'économètre : le résultat de l'appariement dépend du jeu de variables utilisées pour le conditionnement. En théorie, il doit y avoir suffisamment de variables de conditionnement pour que l'hypothèse d'indépendance conditionnelle soit raisonnable. Afin de tester la robustesse de nos résultats d'appariement à l'ajout ou à l'omission de certaines caractéristiques observables, nous faisons varier la spécification de notre score de propension. Le tableau A3 en annexe présente les estimations de l'impact sur la probabilité d'obtenir un CDD d'au moins six mois, pour chacune des cinq spécifications envisagées. Les résultats obtenus par la méthode d'appariement sont globalement robustes à la spécification du score. Cette conclusion est valide pour toutes les sous-populations étudiées.

Enfin, le tableau A4 en annexe présente les estimations pour la seule spécification (3) du score, mais en faisant varier la méthode et le support d'estimation. En effet, il est courant pour les méthodes d'appariement de limiter l'estimation à une partie seulement de l'échantillon pour laquelle la densité du score conditionnellement au traitement est non nulle (Dehejia et Wahba, 2002) ou supérieure à un certain seuil (Smith et Todd, 2005). Cependant, comme le montrent les graphiques A1.1 à A1.5 en annexe, les différentes spécifications que nous avons retenues ont toutes un support commun large. Nous avons toutefois mis en œuvre une procédure dite de réduction (*trimming*) consistant à écarter de l'échantillon les 2 % des observations associées aux valeurs de score ayant la plus faible densité. Ces différentes altérations méthodologiques ne modifient pas nos conclusions, les effets estimés étant quantitativement très proches[14].

Discussion et conclusion

L'approche expérimentale repose classiquement sur les hypothèses 1 (indépendance de l'assignation), 2 (monotonie de l'effet de l'assignation), 3 (exclusion) et 4 (SUTVA), tandis que l'approche par appariement repose sur les hypothèses 4, 6 (indépendance conditionnelle) et 7 (support commun). Nous substituons ici à l'hypothèse 4 l'hypothèse 5 (externalité constante), plus crédible dans le contexte étudié. Cette hypothèse n'est pas rejetée empiriquement et justifie la légitimité des comparaisons présentées dans cet article. Par ailleurs, nous avons vérifié empiriquement la validité de l'hypothèse 7, en examinant le support du score de propension.

Nous tirons parti d'une expérimentation aléatoire pour comparer ses résultats avec ceux qui découleraient de l'approche par appariement. Nous obtenons une divergence nette des résultats obtenus par les deux méthodes. Ceci signifie que certaines des hypothèses citées dans le paragraphe précédent ne sont pas vérifiées. En toute rigueur, on ne peut considérer aucune des deux méthodes comme une référence incontestable dont les résultats refléteraient à coup sûr la réalité, des hypothèses sous-jacentes à chaque méthode pouvant ne pas être vérifiées. Cependant, compte tenu des analyses préexistantes, il nous semble très probable que la méthode par appariement a failli en raison de la difficulté à trouver un jeu de variables de contrôle suffisamment riche pour que la condition 6 d'indépendance conditionnelle soit vérifiée[15]. Les différences entre les résultats expérimentaux et non expérimentaux mises en évidence dans le tableau A2 reflètent vraisemblablement l'existence d'un biais de sélection important dans le cas de ce dispositif. Ce biais serait négatif : en l'absence de programme, les individus qui y participent auraient eu un taux d'emploi nettement plus faible que les individus qui n'y participent pas. Ceci peut traduire le fait que les individus non traités ont refusé d'entrer dans le programme parce qu'ils avaient déjà un emploi en vue ou encore parce qu'ils savent, à juste titre, que leur famille ou leur entourage pourront les aider à trouver un emploi en cas de besoin. Il serait difficile de pouvoir mesurer de telles variables dans une enquête, alors qu'elles peuvent avoir un impact

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

déterminant sur la situation d'entrer dans le programme (et sur l'insertion dans l'emploi).

Au-delà de cet exemple, l'exercice mené est l'occasion de rappeler que, lorsque le mécanisme d'assignation au traitement est mal compris ou qu'il fait intervenir des variables qu'il est impossible de mesurer, les estimations obtenues suivant les méthodes d'appariement risquent de souffrir d'un biais. Ce biais de sélection est potentiellement important et peut changer radicalement le sens et la magnitude des estimations de l'effet moyen des programmes. Cela ne veut pas dire que l'approche par appariement doit être abandonnée mais que le choix du groupe de contrôle et du jeu de variables de contrôle introduites dans le score de propension doit être discuté.

Notes

(1) NDLR : cette méthode est décrite dans Givord (2014), dans ce numéro.

(2) Les différences de différences et l'appariement ne sont pas les seules méthodes à avoir fait l'objet de telles comparaisons. Black *et alii* (2007) et Cook et Wong (2008) utilisent des expérimentations dans lesquelles une partie du traitement a été assignée aléatoirement et une autre selon un critère strict, tel que seuls les individus dont le score était inférieur à un certain seuil étaient éligibles. Ces auteurs peuvent ainsi comparer les estimateurs obtenus par régression sur discontinuité aux résultats expérimentaux.

(3) Les dix régions participantes étaient les suivantes : Île-de-France, Picardie, Haute-Normandie, Centre, Nord-Pas-de-Calais, Lorraine, Pays de Loire, Rhône-Alpes, Provence-Alpes-Côte d'Azur (PACA) et La Réunion.

(4) NDLR : les principes de l'assignation aléatoire sont détaillés et discutés dans Bouguen et Seban (2014), dans ce numéro.

(5) Le code Rome fait référence au Répertoire Opérationnel des Métiers et des Emplois (Rome) de Pôle emploi. Ce répertoire sert à identifier aussi précisément que possible toute offre et toute demande d'emploi afin d'être capable de les rapprocher. Un peu plus de 10 000 appellations de métiers et emplois sont traitées à travers plus de 500 fiches métiers. Source : Pôle emploi.

(6) Cette restriction résulte de l'hypothèse 5, détaillée dans la troisième partie de cet article intitulée "Identifier l'effet du traitement".

(7) NDLR : *cf*. Bouguen et Seban (2014).

(8) NDLR : le vocabulaire inhérent aux techniques d'évaluation est défini dans Givord (2014) et Bouguen et Seban (2014), dans ce numéro.

(9) La preuve de ce résultat peut être trouvée dans Imbens et Angrist (1994).

(10) NDLR : *cf*. Bouguen et Seban (2014), encadré 2.

(11) Le taux d'appariement exact en utilisant toutes les variables explicatives de la spécification 3 du tableau 4 est de 1,01%.

(12) NDLR : pour plus de détails sur l'appariement par le score de propension, voir Givord (2014), dans ce numéro.

(13) Les techniques évoquées dans ce paragraphe sont brièvement décrites en annexe, dans la note de lecture sous le tableau A4.

(14) Nous avons également effectué ces analyses pour les autres sous-échantillons (sous-échantillons par genre, notamment), sans obtenir de conclusion différente.

(15) NDLR : se reporter au commentaire de Sylvain Chabé-Ferret, dans ce numéro, pour une discussion de ce point.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Bibliographie

Arceneaux K., Gerber A.S. et Green D.P. (2006). "Comparing Experimental and Matching Methods Using a Large-Scale Voter Mobilization Experiment", *Political Analysis*, vol. 14, n° 1, pp. 37-62.

Black D., Galdo J. et Smith J. A. (2007). "Evaluating the Bias of the Regression Discontinuity Design Using Experimental Data", Manuscript, University of Michigan.

Bouguen A. et Seban J. (2014). "L'assignation aléatoire comme méthode d'évaluation des politiques publiques", *Économie et Prévision*, n° 204-205, pp. 119-143, ce numéro.

Cook T.D. et Wong V.C. (2008). "Empirical Tests of the Validity of the Regression Discontinuity Design: Implications for its Theory and its Use in Research Practice", *Annales d'Économie et de Statistique*, n° 91-92, pp. 127-150.

Crépon B., Duflo E., Gurgand M., Rathelot R. et Zamora P. (2011). "Accompagnement des jeunes diplômés demandeurs d'emploi par des opérateurs privés de placement : les enseignements d'une évaluation", *Dares Analyses* 2011-094, 14 pages.

Crépon B., Duflo E., Gurgand M., Rathelot R. et Zamora P. (2013). "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment", *Quarterly Journal of Economics*, vol. 128, n° 2, pp. 531-580.

Dehejia R.H. et Wahba S. (1999). "Causal Effects in Non Experimental Studies: Reevaluating the Evaluation of Training Programs", *Journal of Statistical American Association*, vol. 94, n° 448, pp. 1053-1062.

Dehejia R.H. et Wahba S. (2002). "Propensity Score-Matching Methods for Nonexperimental Causal Studies", *Review of Economics and Statistics*, vol. 84, n° 1, pp. 151-161.

Fan J. (1992). "Design-Adapted Nonparametric Regression", *Journal of the American Statistical Association*, vol. 87, n° 420, pp. 998-1004.

Friedlander D. et Robins P.K. (1995). "Evaluating Program Evaluations: New Evidence on Commonly Used Nonexperimental Methods", *American Economic Review*, vol. 85, n° 4, pp. 923-937.

Frölich M. (2004). "Finite-Sample Properties of Propensity-Score Matching and Weighting Estimators", *Review of Economics and Statistics*, vol. 86, n° 1, pp. 77-90.

Givord P. (2014). "Méthodes économétriques pour l'évaluation de politiques publiques", *Économie et Prévision*, n° 204-205, pp. 1-28, ce numéro.

Heckman J.J., Ichimura H., Smith J.A. et Todd P.E. (1998a). "Characterizing Selection Bias Using Experimental Data", *Econometrica*, vol. 66, n° 5, pp. 1017-1098.

Heckman J.J., Ichimura H. et Todd P.E. (1998b). "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, vol. 65, n° 2, pp. 261-294.

Imbens G.W. et Angrist J.D. (1994). "Identification and Estimation of Local Average Treatment Effects", *Econometrica*, vol. 62, n° 2, pp. 467-475.

LaLonde R.J. (1986). "Evaluating the Econometric Evaluations of Training Programs with Experimental Data", *American Economic Review*, vol. 76, n° 4, pp. 604-620.

Rubin D. (1974). "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies", *Journal of Educational Psychology*, vol. 66, n° 5, pp. 688-701.

Seifert B. et Gasser T. (1996). "Finite-Sample Variance of Local Polynomials: Analysis and Solutions", *Journal of the American Statistical Association*, vol. 91, n° 433, pp. 267-275.

Shaikh A.M., Simonsen M., Vytlacil E.J. et Yildiz N. (2009). "A Specification Test for the Propensity Score Using its Distribution Conditional on Participation", *Journal of Econometrics*, vol. 151, n° 1, pp. 33-46.

Smith J.A. et Todd P.E. (2005). "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?", *Journal of Econometrics*, vol. 125, n° 1-2, pp. 305-353.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Annexe : illustrations

Tableau A1 : effet de l'accompagnement sur l'accès à l'emploi (résultats expérimentaux) - robustesse à l'introduction de variables de contrôle

Individus	Tous		Sans emploi initialement		Sans emploi initialement - hommes		Sans emploi initialement - femmes	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Variable de résultat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Emploi	0,016	-0,013	0,050**	0,025	0,044	0,054	0,053*	0,010
	(0,025)	(0,045)	(0,028)	(0,055)	(0,040)	(0,080)	(0,036)	(0,062)
Emploi durable	0,006	0,018	0,057**	0,103**	0,071*	0,137*	0,049*	0,085*
	(0,025)	(0,044)	(0,029)	(0,056)	(0,045)	(0,087)	(0,034)	(0,062)
CDD de six mois ou plus	0,020	0,055*	0,062***	0,116***	0,112***	0,238***	0,041*	0,055
	(0,019)	(0,037)	(0,022)	(0,042)	(0,035)	(0,078)	(0,030)	(0,047)
CDI	-0,014	-0,037	-0,005	-0,013	-0,041	-0,101*	0,009	0,030
	(0,021)	(0,040)	(0,025)	(0,045)	(0,040)	(0,078)	(0,033)	(0,062)
Nombre d'observations	18 853	7 109	8 018	2 982	3 001	1 112	5 017	1 870

Lecture : l'impact est le coefficient relatif à l'indicatrice d'avoir été accompagné(e) par un OPP. Cette variable est instrumentée par le fait d'avoir été assigné(e) à l'accompagnement. Les écarts-types (entre parenthèses) sont corrigés de la corrélation entre les observations d'une même ALE. * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.
Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Tableau A2 : effet de l'accompagnement sur l'accès à l'emploi - différence entre résultats expérimentaux et par appariement

Individus	Tous		Sans emploi initialement		Sans emploi initialement - hommes		Sans emploi initialement - femmes	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Variable de résultat	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Emploi	0,073***	0,021*	0,065***	-0,006	0,080***	0,092**	0,060***	-0,032
	(0,009)	(0,014)	(0,014)	(0,024)	(0,024)	(0,040)	(0,019)	(0,034)
Emploi durable	0,105***	0,083***	0,122***	0,120***	0,147***	0,190***	0,102***	0,119***
	(0,009)	(0,015)	(0,014)	(0,025)	(0,026)	(0,043)	(0,018)	(0,033)
CDD de six mois ou plus	0,039***	0,067***	0,072***	0,116***	0,141***	0,255***	0,031**	0,061**
	(0,007)	(0,013)	(0,013)	(0,022)	(0,022)	(0,037)	(0,017)	(0,030)
CDI	0,066***	0,017*	0,050***	0,004	0,006	-0,065*	0,071***	0,057**
	(0,008)	(0,013)	(0,012)	(0,021)	(0,021)	(0,042)	(0,017)	(0,032)
Nombre d'observations	18 853	7 109	8 018	2 982	3 001	1 112	5 017	1 870

Lecture : les écarts-types sont reportés entre parenthèses. Ils ont été calculés par *bootstrap* (99 réplications) et sont corrigés de la corrélation entre les observations d'une même ALE. Le modèle expérimental estimé comprend également le sexe, l'ancienneté au chômage et son carré, le diplôme et des indicatrices relatives aux quintuplés de l'ALE. Les estimations par appariement ont été obtenues par régression Ridge en utilisant le score de propension (3) du tableau 3 et une fenêtre de longueur optimale. * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.
Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Tableau A3 : effet estimé par appariement - robustesse à la spécification du score de propension

Individus	Tous		Sans emploi initialement		Sans emploi initialement - hommes		Sans emploi initialement - femmes	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Spécification du score	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Spécification (1)	-0,012**	-0,007	-0,009	-0,004	-0,014	0,004	0,000	-0,010
	(0,007)	(0,014)	(0,012)	(0,020)	(0,020)	(0,034)	(0,016)	(0,029)
Spécification (2)	-0,012**	-0,007	-0,011	-0,007	-0,030*	0,025	0,000	-0,023
	(0,007)	(0,014)	(0,013)	(0,020)	(0,020)	(0,039)	(0,016)	(0,029)
Spécification (3)	-0,014**	-0,009	-0,014	-0,007	-0,025	0,002	-0,003	-0,032
	(0,007)	(0,014)	(0,013)	(0,019)	(0,020)	(0,039)	(0,016)	(0,031)
Spécification (4)	-0,014**	-0,008	-0,009	0,000	-0,025	0,002	-0,003	-0,032
	(0,007)	(0,014)	(0,013)	(0,020)	(0,020)	(0,039)	(0,016)	(0,031)
Spécification (5)	-0,015**	-0,008	-0,015	0,000	-0,028	0,012	-0,010	-0,024
	(0,007)	(0,015)	(0,013)	(0,022)	(0,023)	(0,042)	(0,016)	(0,033)
Nombre d'observations	13 041	4 891	5 582	2 102	2 152	808	3 430	1 294

Lecture : l'impact est le coefficient relatif à l'indicatrice d'avoir été accompagné(e) par un OPP, la variable de résultat étant l'obtention d'un CDD d'au moins six mois. Les écarts-types sont reportés entre parenthèses. Ils ont été calculés par *bootstrap* (99 réplications). Les estimations ont été obtenues par régression Ridge en utilisant une fenêtre de longueur optimale. * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.
Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Tableau A4 : robustesse de l'effet estimé par appariement à la méthode d'estimation

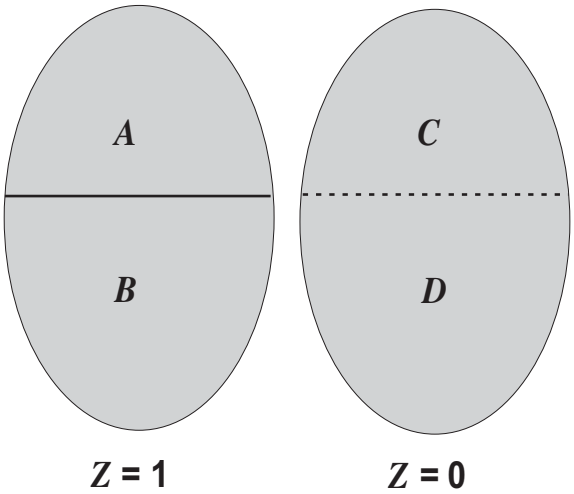
Support	Tous				Support commun			
Individus	Tous		Sans emploi initialement		Tous		Sans emploi initialement	
OPP	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands	Tous OPP	OPP marchands
Méthode d'estimation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Méthode de Nadaraya-Watson[a]	-0,014*	-0,009	-0,014	-0,007	-0,013**	-0,010	-0,009	-0,009
	(0,009)	(0,014)	(0,012)	(0,020)	(0,007)	(0,014)	(0,012)	(0,017)
Régression linéaire locale[b]	-0,014*	-0,009	-0,014	-0,003	-0,013**	-0,010	-0,009	-0,009
	(0,009)	(0,014)	(0,012)	(0,021)	(0,007)	(0,013)	(0,012)	(0,019)
Régression Ridge (fenêtre optimale)[c]	-0,014*	-0,009	-0,014	-0,007	-0,013*	-0,011	-0,009	-0,010
	(0,009)	(0,014)	(0,012)	(0,020)	(0,008)	(0,013)	(0,013)	(0,020)
Régression Ridge (fenêtre optimale + 50 %)[d]	-0,014*	-0,008	-0,014	-0,005	-0,013*	-0,011	-0,009	-0,008
	(0,009)	(0,014)	(0,012)	(0,020)	(0,008)	(0,014)	(0,012)	(0,018)
Régression Ridge (fenêtre optimale - 50 %)[e]	-0,014*	-0,010	-0,014	-0,007	-0,012*	-0,012	-0,009	-0,011
	(0,009)	(0,015)	(0,012)	(0,022)	(0,008)	(0,014)	(0,013)	(0,019)
Nombre d'observations	13 041	4 891	5 582	2 102	12 781	4 794	5 471	2 060

Lecture : l'impact est le coefficient relatif à l'indicatrice d'avoir été accompagné(e) par un OPP, la variable de résultat étant l'obtention d'un CDD d'au moins six mois. Les écarts-types sont reportés entre parenthèses. Ils ont été calculés par *bootstrap* (99 réplications). Ces estimations ont été réalisées par appariement sur le score (3) du tableau 3. Le support commun correspond aux individus sélectionnés selon une procédure de réduction (*trimming*) excluant les 2 % des individus pour lesquels la densité empirique du score est la plus faible. (a) La méthode de Nadaraya-Watson est une méthode classique d'estimation par régression par noyau[1]. (b) La régression linéaire locale constitue une généralisation des méthodes de régression non paramétriques par noyau qui se ramène à une estimation par moindres carrés pondérés localement. L'estimateur obtenu est connu pour ses bonnes propriétés asymptotiques. Toutefois, contrairement à la méthode de Nadaraya-Watson, la régression linéaire locale a l'inconvénient à distance finie de déboucher sur des variances conditionnelles non bornées lorsqu'un noyau à support compact est utilisé. C'est pourquoi Seifert et Gasser (1996) ont proposé d'appliquer l'idée de la régression Ridge à la régression linéaire locale[2]. (c), (d) et (e) présentent les résultats de la régression linéaire Ridge, lorsque la fenêtre choisie pour le noyau est la fenêtre optimale (c), la fenêtre optimale élargie de 50 % (d) ou réduite de 50 % (e). * = impact significatif à 10 % ; ** = impact significatif à 5 % ; *** = impact significatif à 1 %.

Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.

Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Figure A1 : schéma des populations de l'expérimentation



Lecture : la variable muette *Z* vaut 1 pour les individus assignés au groupe de traitement et 0 pour les individus assignés au groupe de contrôle. Le groupe de traitement se décompose en deux sous-groupes : le sous-groupe *A* (les individus qui, assignés au groupe de traitement, ont accepté ce dernier) et le sous-groupe *B* (les individus qui, assignés au groupe de traitement, ont refusé ce dernier). Le groupe de contrôle se décompose également en deux sous-groupes : le sous-groupe *C* (les individus qui, non assignés au groupe de traitement, auraient accepté ce dernier si on le leur avait proposé) et le sous-groupe *D* (les individus qui, non assignés au groupe de traitement, auraient refusé ce dernier si on le leur avait proposé).

Source : auteurs.

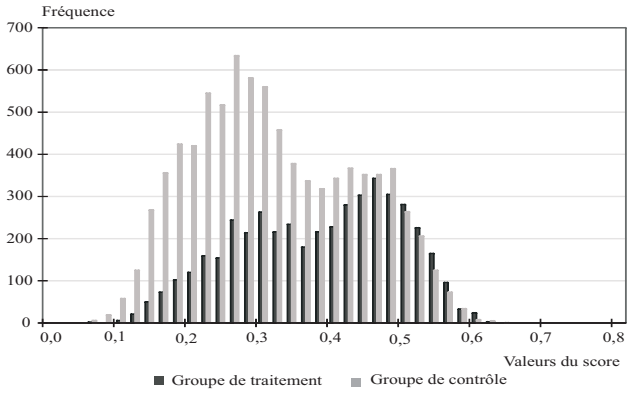
(1) NDLR : un noyau est une fonction de pondération utilisée dans les techniques d'estimation non-paramétrique, ou dans les régressions paramétriques pour estimer des espérances conditionnelles. Un noyau est fonction d'un paramètre nommé "fenêtre", qui régit le degré de lissage de l'estimation. Plus la fenêtre est large, plus l'estimation est lissée ; plus elle est étroite, plus l'estimation reproduit finement les évolutions locales.

(2) NDLR : dans le cadre d'un modèle linéaire ordinaire $y = Xb + u$, la régression Ridge consiste à modifier l'estimateur des MCO $(X'X)^{-1}X'y$ de b par l'introduction d'un terme correcteur dans la matrice $(X'X)^{-1}$ de façon à en réduire les termes de variance les plus élevés, au prix de l'apparition d'un biais qui devra être suffisamment limité pour que l'erreur quadratique moyenne de l'estimateur Ridge ainsi obtenu soit sensiblement inférieure à celle de l'estimateur des MCO. L'estimateur Ridge se présente, sous sa forme la plus simple, comme : $(X'X + \lambda I)^{-1}X'y$, où I est la matrice identité et λ un paramètre de rétrécissement, déterminé par optimisation. Les régressions Ridge effectuées dans cet article consistent à appliquer le même type de correction à la régression linéaire locale pour obtenir des matrices de variances conditionnelles bornées à distance finie.

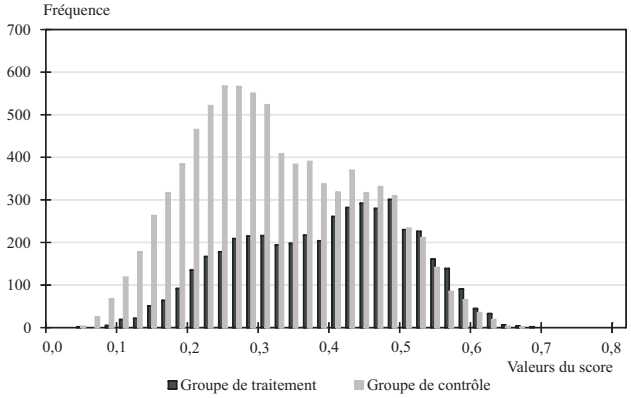
Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)

Graphiques A1 : support des différentes spécifications du score selon le statut de participation

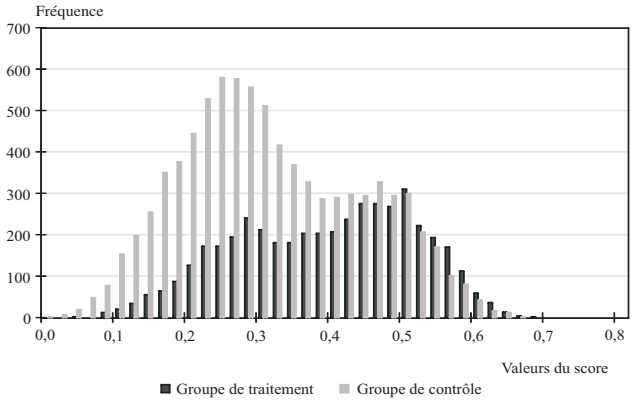
Graphique A1.1 : histogramme du score (1)



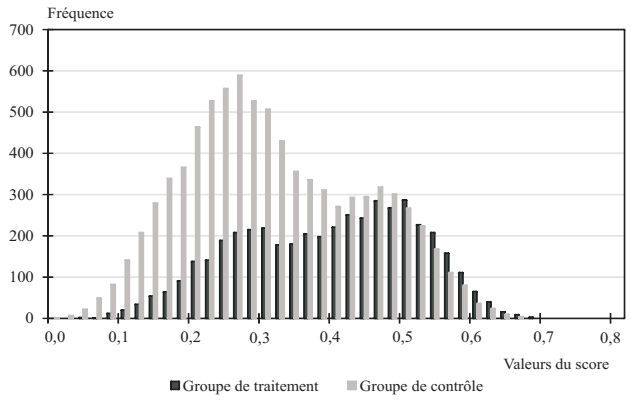
Graphique A1.2 : histogramme du score (2)



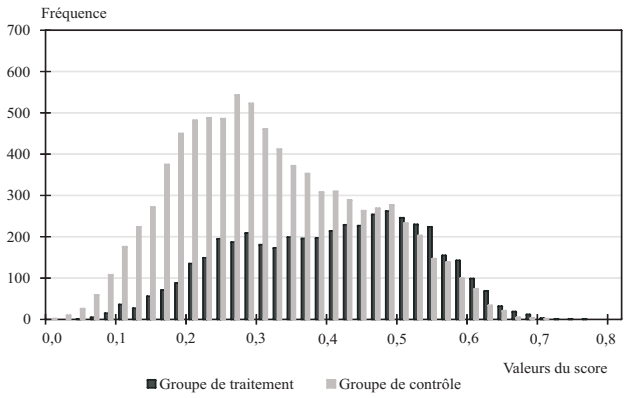
Graphique A1.3 : histogramme du score (3)



Graphique A1.4 : histogramme du score (4)



Graphique A1.5 : histogramme du score (5)



Lecture : le numéro de chaque score renvoie à la spécification correspondante définie dans le tableau 3.
Champ : individus appartenant à une zone traitée et ayant répondu à la première interrogation de l'enquête de suivi.
Sources : Dares, enquête auprès des jeunes éligibles à la prestation d'insertion "jeunes diplômés", calculs des auteurs.

Direction générale du Trésor | Téléchargé le 27/10/2026 sur https://shs.cairn.info par Courtier Théophile (IP: 81.1.6.66)