

LE PANEL 2008 DES BÉNÉFICIAIRES DE CONTRATS AIDÉS ET D'UNE POPULATION TÉMOIN : UN OUTIL DÉDIÉ A L'ÉVALUATION PAR *MATCHING*

VERSION PROVISOIRE, NE PAS CITER

Isabelle BENOTEAU ()*, *Sylvie LE MINEZ (**)*, *Laurent LEQUIEN (***)* et *Véronique RÉMY(*)*

() Dares, Département des politiques de l'emploi*

*(**) Insee, Division Emploi*

*(***) Drees, Bureau « retraites »*

Introduction

Les politiques actives du marché du travail, mises en œuvre par l'ensemble des pays occidentaux, ont pour principal objectif d'aider les demandeurs d'emploi à se réinsérer sur le marché du travail. Elles comprennent les politiques d'aide à la recherche d'emploi, les politiques de formation, les politiques de subvention et les politiques de création d'emplois publics temporaires. L'importance des sommes consacrées aux politiques actives (14 Md€ en 2009) font qu'elles constituent un enjeu pour la politique économique, en particulier en France. Elles sont d'ailleurs fortement mobilisées en période de forte dégradation de la conjoncture avec l'enjeu de limiter la hausse du chômage. Ces politiques sont souvent difficiles à évaluer en raison du caractère multidimensionnel des objectifs qui leur sont assignés (aide à l'insertion, rôle contra-cyclique, soutien à l'activité, amélioration du capital humain) et de la multiplicité des situations possibles sur le marché du travail.

Le service statistique du Ministère du travail lance régulièrement depuis le début des années 1990 des travaux d'évaluation des effets du passage en contrat aidé (politique de subvention dans le secteur dit marchand et politique de création d'emplois publics temporaires dans le secteur dit non marchand). Ces travaux reposent sur la comparaison des trajectoires professionnelles de bénéficiaires et de « populations témoins », constituées d'individus éligibles aux caractéristiques proches de celles bénéficiaires, mais n'ayant pas bénéficié de la mesure. La dernière enquête auprès de demandeurs d'emploi recrutés en contrats aidés et de témoins a été menée en 2008 et en 2009 afin d'évaluer les contrats aidés issus de la loi de cohésion sociale de 2005. Cette loi a modifié les contrats aidés antérieurs en prévoyant davantage d'obligations pour les employeurs en termes d'accompagnement et de formation pendant le contrat et en créant deux contrats (le contrat d'avenir et le contrat insertion-revenu minimum d'activité) spécifiquement destinés aux allocataires de minima sociaux. Les deux autres contrats issus de la loi s'adressent quant à eux aux personnes rencontrant des difficultés particulières d'accès à l'emploi (contrat d'accompagnement dans l'emploi dans le secteur non marchand et contrat initiative emploi dans le secteur marchand). Ces contrats ont été remplacés par le contrat unique d'insertion (CUI) à partir de janvier 2010 en France métropolitaine.

En France, les derniers travaux d'évaluation de ce type de politiques actives (politiques de subvention et de création d'emplois publics temporaires) montrent, comme la plupart des études étrangères (par exemple : Sianesi, 2004), que les politiques de subvention dans le secteur marchand ont un effet positif sur l'emploi tandis que les politiques de création

d'emplois publics temporaires n'ont au mieux pas d'effets sur leurs bénéficiaires (Even et Klein (2008)). Ces derniers travaux, menés à partir du précédent panel de la Dares, ne se sont pas spécifiquement intéressés à la récurrence du passage en contrat aidé ni aux mesures d'aide à la recherche d'emploi ayant pu être dispensées aux bénéficiaires, qui peuvent jouer sur l'efficacité de ces programmes. Aussi, le nouveau panel a été élaboré de sorte à mieux appréhender les processus d'insertion professionnelle – davantage décrits dans leur durée –, la récurrence du passage en contrat aidé, et les autres mesures – aide à la recherche d'emploi, formation – dont ont pu bénéficier les demandeurs d'emploi parallèlement à leur passage en contrat aidé qui sont connues *via* les données administratives (fichier historique de l'ANPE) sur lesquelles s'appuie en partie l'enquête.

Cette présentation vise à montrer la richesse des exploitations possibles à partir du nouveau panel ainsi que les spécificités méthodologiques liées à la politique étudiée et aux données mobilisées. L'accent est mis sur la nécessaire adaptation aux données d'enquête des méthodes traditionnelles d'appariement (*matching*) pour l'utilisation desquelles a été conçu le panel. Le plan de l'article est le suivant : la première section fait un rapide tour d'horizon de la réforme des contrats aidés de 2005 ; la deuxième section présente le panel et donne les principaux résultats descriptifs sur les bénéficiaires ; la troisième section présente quelques éléments théoriques sur l'évaluation par *matching* et les choix d'implémentation retenus dans l'étude ; enfin, la dernière section illustre la méthode par quelques résultats exploratoires de l'effet moyen du passage en contrat aidé sur la trajectoire des bénéficiaires.

1 Les contrats aidés issus de la loi de cohésion sociale de 2005

Les contrats aidés sont des emplois subventionnés par l'État visant à favoriser l'insertion professionnelle de publics cibles plus ou moins éloignés du marché du travail : chômeurs de longue durée, allocataires de minima sociaux, jeunes non qualifiés, etc. La loi de programmation pour la cohésion sociale du 18 janvier 2005 a procédé à une refonte des contrats aidés en direction des personnes rencontrant des difficultés particulières d'accès à l'emploi, en vue notamment de simplifier l'accès à ces contrats. Dans le secteur non marchand, le contrat d'avenir (CAV) et le contrat d'accompagnement dans l'emploi (CAE) ont remplacé le contrat emploi solidarité (CES) et le contrat emploi consolidé (CEC) tandis que, dans le secteur marchand, le contrat initiative emploi (CIE) et le contrat d'insertion revenu minimum d'activité (CI-RMA) ont été fortement remodelés (les caractéristiques détaillées de ces contrats sont présentées dans le tableau 5 en annexe). Les critères d'éligibilité sont plus ou moins larges selon les contrats :

- les CAE et les CIE sont destinés aux « personnes sans emploi rencontrant des difficultés sociales particulières et professionnelles d'accès à l'emploi »,
- tandis que les CAV et les CI-RMA sont réservés aux bénéficiaires de minima sociaux. Les minima sociaux concernés sont : l'allocation de solidarité spécifique (ASS), l'allocation de parent isolé (API), l'allocation aux adultes handicapés (AAH) (depuis 2006) et le revenu minimum d'insertion (RMI).

Les critères pour entrer dans les nouveaux contrats issus de la loi de cohésion sociale, à l'exclusion de ceux réservés aux bénéficiaires de minima sociaux, sont volontairement moins précis que ceux de leurs prédécesseurs. En effet, pour entrer en CES ou en CEC, les demandeurs d'emploi devaient avoir des caractéristiques spécifiques : avoir une ancienneté de chômage de plus de douze mois, être âgé de plus de cinquante ans, ou encore être jeune et peu diplômé. Or, le ciblage des publics au niveau national souffrait d'un éventuel décalage avec les réalités locales. De plus, certaines personnes n'entrant pas dans les catégories requises

public secteur	Personnes sans emploi rencontrant des difficultés sociales et professionnelles d'accès à l'emploi	Allocataires de minima sociaux (RMI, ASS, API, AAH)
Marchand	CIE Contrat Initiative Emploi	CI-RMA Contrat Insertion - Revenu Minimum d'Activité
Non marchand	CAE Contrat d'Accompagnement dans l'Emploi	CAV Contrat d'Avenir

Figure 1 – Critères d'éligibilité par type de contrat

pouvaient se voir refuser l'attribution d'un contrat aidé, quand bien même un tel contrat aurait pu leur être bénéfique. À partir de 2005, pour le CAE et le nouveau CIE, l'État se borne donc à fixer l'enveloppe budgétaire annuelle et le taux maximal de prise en charge (respectivement 95% et 47%), tandis que le taux effectif de prise en charge et le ciblage des publics sont définis au niveau régional par arrêté préfectoral, après discussion entre les services déconcentrés de l'État et l'ANPE (désormais Pôle Emploi). L'aide associée au contrat peut ainsi être modulée au niveau local pour favoriser l'embauche de tel ou tel type de public en fonction du contexte. Il existe par conséquent une certaine hétérogénéité des publics ciblés selon les régions.

Pour les embauches en CIE et en CAE, l'employeur reçoit donc une subvention proportionnelle au salaire versé, sur la base d'un taux fixé localement. En revanche, le montant de l'aide associée aux contrats réservés aux bénéficiaires de minima sociaux est fixé au niveau national, et correspond au RMI versé à une personne seule. Le principe du CI-RMA et du CAV est celui d'une « activation des dépenses passives » : l'allocation qu'aurait normalement perçu le salarié en absence d'emploi est versée à l'employeur sous forme de subvention¹. Ces aides sont cumulables avec les exonérations générales de cotisations employeurs pour les contrats du secteur marchand, et complétées par une exonération étendue pour ceux du secteur non marchand. En outre, les publics recrutés dans le secteur non marchand étant généralement plus éloignés de l'emploi que ceux embauchés dans le secteur marchand, l'aide versée est plus élevée pour les premiers que pour les seconds.

De manière générale, la subvention versée dans le cadre du contrat aidé vise à baisser le coût du travail, incitant ainsi les employeurs à recruter des salariés ayant une productivité plus faible ou des caractéristiques défavorables au regard de l'insertion sur le marché du travail (chômeurs de longue durée, etc.). Le nombre d'emplois offerts est ainsi augmenté, induisant une baisse du taux de chômage, notamment du chômage de longue durée (limitant ainsi les effets pervers d'hystérésis).

Les CIE et les CI-RMA peuvent être conclus à durée déterminée ou indéterminée. En revanche, les CAE et les CAV sont systématiquement à durée déterminée contrairement aux CEC qui pouvaient être à durée indéterminée. En cas de rupture anticipée du contrat, et en l'absence de faute grave du salarié ou d'accord de sa part, l'employeur est tenu de reverser les aides perçues depuis le début du CIE ou du CI-RMA. Quel que soit le contrat, celui-ci peut être rompu avant terme à l'initiative du salarié sous réserve qu'il ait été embauché en CDI, en CDD de plus de six mois ou qu'il soit engagé dans une formation qualifiante. Les contrats signés à durée déterminée peuvent être reconduits dans la limite de la durée totale fixée par la loi. Ces reconductions sont beaucoup plus fréquentes en CAE et CAV qu'en CIE et CI-RMA, ces derniers étant plus souvent conclus à durée indéterminée. La durée hebdomadaire du travail est fixée à 26 heures en CAV, tandis que les autres contrats

1. Le détail des aides reçues par l'employeur est présenté dans le tableau 5.

peuvent être à temps plein ou à temps partiel. Dans le cas de contrats à temps plein, l'effet de *locking-in*² du contrat aidé peut être plus important : les bénéficiaires ont moins de temps pour rechercher un emploi et y sont d'ailleurs moins incités, le salaire associé au contrat aidé étant plus conséquent.

Enfin, lors de la conception de ces nouveaux contrats, l'accent a été mis sur l'accompagnement et la formation. De telles actions sont ainsi obligatoires dans le cas du CAV et prévues dans le cas du CAE tandis que les formations sont facultatives pour les CIE et les CI-RMA. Il est d'ailleurs prévu dans le cas du CAE que l'aide à l'employeur puisse être modulée en fonction des actions de formation et d'accompagnement prévues en faveur du bénéficiaire.

2 Le Panel 2008 : une enquête conçue pour une évaluation par *matching*

Le problème sous-jacent à toute évaluation de politique publique est qu'il est impossible d'observer simultanément la situation d'un bénéficiaire de la mesure et celle qu'il aurait connu en l'absence de cette dernière. L'estimateur le plus direct pour évaluer l'effet de la mesure consiste à comparer les situations respectives des bénéficiaires et des individus non bénéficiaires de la mesure. Néanmoins, les premiers présentent vraisemblablement des caractéristiques particulières susceptibles de jouer simultanément sur leur entrée en mesure et leur situation. L'estimateur naïf est donc *a priori* biaisé.

Dans le cadre de l'évaluation des politiques publiques, l'expérimentation aléatoire constitue « le paradigme des méthodes économétriques d'évaluation » (Givord, 2010) car elle fournit un estimateur sans biais de l'effet d'une politique sur la population bénéficiaire. Il s'agit cependant d'une méthode difficile à mettre en place sur le terrain et encore relativement rare, bien qu'en fort développement. Dès lors, les chercheurs sont contraints de se tourner vers des techniques d'estimation *ex post* (relatives à des politiques déjà instaurées), parmi lesquelles figurent le *matching*. Les données administratives permettent rarement de disposer de l'information nécessaire à une évaluation robuste des politiques publiques et requièrent que leur soient adossées des enquêtes conçues spécifiquement dans ce but. Toute technique d'évaluation reposant sur des hypothèses fortes, l'anticipation en amont des enjeux économétriques augmente les chances de validité des estimateurs mobilisés. À cet égard, le « Panel 2008 » constitue une originalité puisqu'il a d'emblée été conçu afin de permettre des analyses causales du passage en contrat aidé par des techniques de *matching*.

Dans sa version la plus simple, le *matching* consiste à apparier chaque bénéficiaire à un individu, dit « témoin », ayant des caractéristiques observables proches mais n'ayant pas bénéficié du « traitement » : l'effet moyen de la politique sur les bénéficiaires (*Average Treatment effect on the Treated (ATT)*) est alors estimé par la moyenne des écarts entre l'*outcome* Y_1 des bénéficiaires et l'*outcome* Y_0 des témoins utilisé pour approximer l'*outcome* contrefactuel qui aurait été celui des bénéficiaires s'ils n'avaient pas reçu le traitement :

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

L'hypothèse sous-jacente au *matching* est l'indépendance – conditionnellement à un ensemble de variables observables X – entre l'affectation au traitement et l'*outcome* : $Y_0 \perp T|X$ (hy-

2. L'intensité des recherches est réduite durant la participation à des programmes et les bénéficiaires se retrouvent « bloqués » dans leur programme.

pothèse d'indépendance conditionnelle). Autrement dit, conditionnellement aux variables X , l'*outcome* des témoins constitue un bon contrefactuel de l'*outcome* des bénéficiaires en l'absence de traitement. Cette hypothèse, qui permet de recréer les conditions d'une expérience aléatoire, est naturellement très forte et suppose qu'il n'existe pas, en dehors des X observés, de caractéristiques qui influent à la fois sur l'*outcome* et sur l'entrée en traitement.

La méthode repose en outre sur la présence, dans les données, de témoins ayant des caractéristiques identiques à celles des bénéficiaires, autrement dit sur l'existence d'un support commun : $0 < P(T = 1|X) < 1$.

La validité des estimateurs est, par conséquent, tributaire de la qualité des données disponibles, et donc en grande partie du protocole d'enquête retenu. Celui-ci se doit d'anticiper au mieux les problèmes d'identification susceptibles de se poser *ex post* pour l'évaluation.

2.1 L'échantillonnage

Afin de garantir au mieux la comparabilité des témoins et des bénéficiaires des différents contrats – indispensable à une évaluation par *matching* – l'ensemble des enquêtés du Panel 2008 ont été échantillonnés dans une même source (Crépon et Kamionka, 2006) : le fichier historique de l'ANPE (FH)³. En outre, pour faciliter la modélisation des trajectoires individuelles, le tirage a été réalisé dans le flux des inscrits à l'ANPE au deuxième trimestre 2005. Le groupe des bénéficiaires a alors été défini comme l'ensemble des individus recrutés pour la première fois entre avril 2005 et juin 2007 dans l'un des quatre contrats aidés issus de la loi de cohésion sociale. Par opposition, le groupe des témoins englobe l'ensemble des individus éligibles aux contrats aidés, mais non entrés dans l'un des quatre contrats sur la période de référence.

Afin de pouvoir identifier les bénéficiaires, il a été nécessaire d'apparier le FH au fichier des conventions individuelles d'embauche en contrat aidé de l'ASP (Agence de Services et de Paiement) qui fournit les dates d'entrée en contrat aidé : parmi les 1,1 million de demandeurs d'emploi inscrits à l'ANPE au deuxième trimestre 2005, environ 83 000 ont été recrutés pour la première fois en CAE, CAV, CIE ou CI-RMA entre avril 2005 et juin 2007.

Compte tenu du coût des enquêtes et pour limiter la taille de l'échantillon, les demandeurs d'emploi auraient pu être échantillonnés selon un taux de sondage croissant en fonction d'un score mesurant leur probabilité d'entrer dans chacun des quatre contrats aidés. Ce type de tirage aurait permis de garantir à la fois une sur-représentation des contrats aidés dans l'échantillon et une proximité suffisante entre témoins et bénéficiaires (nécessaire à l'exploitation par *matching*), tout en assurant un mode d'échantillonnage identique pour tous. Il aurait cependant laissé une incertitude sur le rapport entre la taille des populations témoins et bénéficiaires.

Une approche alternative a donc été retenue, consistant à sur-représenter de manière exogène les bénéficiaires de contrats aidés dans l'échantillon (*choice based sampling*). Pour chaque contrat, les bénéficiaires ont été tirés selon un plan de sondage équilibré⁴ parmi l'ensemble des individus identifiés comme étant entrés en contrat aidé sur la période de référence. Pour les témoins, deux populations ont été sélectionnées selon le critère d'éligibilité pertinent : une première population de témoins éligibles aux CAE et CIE (témoins 1), et une

3. Le choix de cette base de sondage présente une limite principale : l'absence de prise en compte des éligibles non inscrits sur les listes de l'ANPE. Cette catégorie de bénéficiaires représentait environ 13% des entrants en contrat aidé en 2005. Les résultats de l'estimation ne sont donc pas généralisables à l'ensemble de la population des entrants en contrat aidé.

4. Cette méthode assure le choix aléatoire d'un échantillon apte à restituer la structure de la base de sondage. L'échantillon a été équilibré notamment sur l'âge, le sexe, le diplôme, la période d'entrée en contrat aidé ou l'éventuelle perception d'un minimum social

seconde population de témoins éligibles aux CI-RMA et CAV (témoins 2), plus restreinte du fait de critères d'éligibilité plus restrictifs pour ces contrats (bénéficiaires de minima sociaux).

Afin de garantir une proximité suffisante entre témoins et bénéficiaires et d'assurer ainsi *ex ante* un support commun suffisamment large⁵, chaque population de témoins a été sélectionnée sur la base d'une probabilité d'entrée en contrat aidé la plus proche possible de celle des bénéficiaires pour lesquels ils sont censés témoigner. Une fois déterminées les caractéristiques disponibles dans le FH influençant cette probabilité (sexe, âge, formation, situation maritale, nombre d'enfants), deux stratégies ont été adoptées selon les cas :

- les témoins 1 ont été échantillonnés aléatoirement au sein de strates d'individus présentant exactement les mêmes caractéristiques que les bénéficiaires (on a alors des « jumeaux parfaits »).
- les témoins 2 ont été sélectionnés aléatoirement au sein de strates d'individus présentant la même probabilité instantanée d'entrer en contrat aidé que celle des bénéficiaires. Ces probabilités ont été estimées à l'aide d'un modèle de Cox modélisant la durée avant l'entrée en contrat aidé.

2.2 Richesse des informations recueillies

La recherche d'un support commun grâce au mode d'échantillonnage n'a d'intérêt qu'à condition de pouvoir garantir au mieux l'hypothèse d'indépendance conditionnelle, indispensable pour valider les techniques de *matching*. Cette hypothèse ne peut être crédible que si les variables disponibles pour réaliser l'appariement sont suffisamment riches. À cet égard, les données issues du fichier historique de l'ANPE sont naturellement mobilisables : parcours du demandeur d'emploi, nombre d'épisodes de chômage depuis 1997, durée passée au chômage, type d'emploi recherché, etc. Mais le Panel 2008, comme plus généralement les données d'enquête, est bien plus riche que de simples données administratives. De nombreuses questions ont été posées de manière similaire aux bénéficiaires et aux témoins en vue de capter finement l'hétérogénéité individuelle : environnement familial, origine des parents, maîtrise de la langue, état de santé, parcours professionnel avant l'inscription à l'ANPE, rapport aux études, expérience, atouts et difficultés pour accéder à l'emploi (garde d'enfants, difficultés de transport, éventuelles discriminations ressenties), motif d'inscription à l'ANPE, conditions de vie, etc. Les informations recueillies permettent donc de relativement bien appréhender l'hétérogénéité habituellement inobservée.

En théorie, les variables mobilisées pour conditionner l'appariement, si elles sont susceptibles d'être affectées par la politique, ne devraient pas être mesurées après la mise en place de celle-ci (Givord, 2010). Une telle démarche générerait en effet un problème d'endogénéité. En pratique, les questionnaires ont donc été conçus pour caractériser au mieux, de manière rétrospective, la situation qui était celle des enquêtés lors de leur inscription à l'ANPE au deuxième trimestre 2005. Près de trois ans et demi séparent la date d'inscription (point de départ du recueil de données détaillées sur les trajectoires professionnelles) de la première vague d'interrogation de l'enquête (fin 2008). Le choix de cette période résulte d'un compromis : d'une part, interroger sur un passé trop lointain peut engendrer potentiellement des biais de mémoire. D'autre part, un recul suffisant est indispensable pour que puissent être observées non seulement les entrées en contrat aidé, mais également les sorties et le devenir après le contrat. Pour compléter les trajectoires, une deuxième vague d'interrogation a eu lieu un an plus tard.

5. Des témoins ayant des caractéristiques trop différentes de celles des bénéficiaires ne pourraient être appariés par la suite et seraient automatiquement exclus de l'échantillon sur lequel mener les estimations.

Le protocole de collecte des données a également été établi afin de garantir la meilleure comparabilité possible entre les témoins et les bénéficiaires, notamment en termes de taux de réponse (Crépon et Kamionka, 2006). Ainsi, un questionnaire téléphonique commun a été mis au point pour les bénéficiaires et leurs témoins. Sur le terrain de l'enquête, les enquêteurs n'ont pas été spécialisés par population (bénéficiaires *versus* témoins), par conséquent, les efforts de recherche des coordonnées et le nombre d'appels téléphoniques ont été identiques.

2.3 Statistiques descriptives et présentation du traitement

Les publics ciblés par les différents contrats étudiés sont relativement hétérogènes : les bénéficiaires d'un contrat du secteur non marchand sont généralement plus éloignés de l'emploi que les bénéficiaires d'un contrat du secteur marchand (cf. tableau 1). De même, au sein de chaque secteur, les demandeurs d'emploi entrés dans un contrat aidé réservé aux bénéficiaires de minima sociaux présentent des caractéristiques moins favorables au regard de l'insertion dans l'emploi.

	CAE	CIE	CAV	CI-RMA	TM1	TM2
Age (%)						
moins de 26 ans	0,31	0,30	0,12	0,07	0,31	0,13
plus de 50 ans	0,10	0,17	0,09	0,09	0,13	0,13
Sexe (hommes) (%)	0,33	0,54	0,43	0,56	0,50	0,55
En couple (%)	0,50	0,60	0,35	0,48	0,54	0,36
Nombre moyen d'enfants à charge	0,85	0,81	0,99	1,02	0,94	0,89
Formation (%)						
< CAP-BEP	0,25	0,16	0,28	0,18	0,20	0,24
CAP-BEP	0,27	0,37	0,37	0,38	0,33	0,37
niveau Bac	0,21	0,22	0,19	0,23	0,20	0,18
supérieur au Bac	0,16	0,25	0,16	0,21	0,27	0,21
Parcours professionnel avant inscription (%)						
toujours ou principalement en emploi	0,38	0,58	0,29	0,49	0,57	0,44
alternance emploi, chômage, inactivité	0,33	0,26	0,39	0,30	0,21	0,31
toujours ou principalement au chômage	0,15	0,09	0,20	0,16	0,09	0,15
toujours ou principalement inactif	0,14	0,07	0,12	0,05	0,12	0,10
Santé (%)						
très bonne santé	0,35	0,33	0,30	0,35	0,38	0,28
assez bonne santé	0,38	0,38	0,39	0,36	0,38	0,33
assez mauvaise santé	0,18	0,18	0,18	0,19	0,16	0,25
très mauvaise santé	0,09	0,10	0,13	0,10	0,08	0,13

Source : Panel 2008, Dares

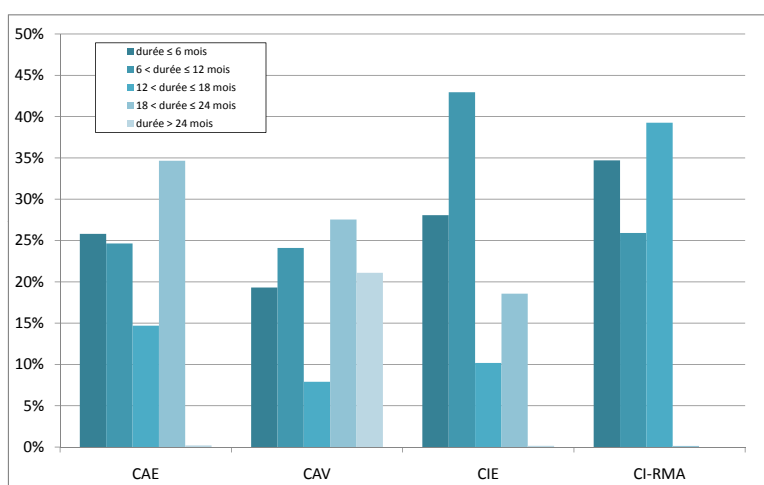
Tableau 1 – Quelques caractéristiques des bénéficiaires et des témoins

Pour tenir compte de l'hétérogénéité des publics visés, l'évaluation est menée par type de contrat. En outre, pour un contrat donné, il est nécessaire de tenir compte de l'hétérogénéité du traitement (Lechner, 2002 ; Hujer, Caliendo et Thomsen, 2003). Les principales sources d'hétérogénéité susceptibles d'influer sur l'efficacité du traitement et identifiables dans le panel sont les suivantes :

- *le délai avant l'entrée en traitement* : la fenêtre d'entrée en traitement étant relativement large (deux ans), le traitement reçu par les bénéficiaires diffère en termes de durée passée au chômage avant l'entrée en contrat aidé. Au-delà d'une tendance générale par type de contrat (entrées plus rapides pour les bénéficiaires de CAE et de CIE), le traitement reçu peut être hétérogène au sein d'une même catégorie de bénéficiaires. Or, les effets du contrat aidé sont susceptibles de varier selon le moment auquel il intervient

au cours de l'épisode de chômage et, plus largement, du parcours professionnel du bénéficiaire (Carling et Richardson, 2004 ; Sianesi, 2004). Le panel permet de traiter la question d'un éventuel « délai optimal » avant l'entrée en contrat aidé.

- *le contenu du traitement* : les données disponibles permettent de connaître la fréquence et le type d'accompagnement (interne ou externe) dont a pu bénéficier le salarié en contrat aidé ; l'existence, la durée et le type de formation éventuelle ; le type d'employeur ; le temps de travail ou encore le métier exercé durant le contrat. Ces informations sont essentielles dans la mesure où le contenu du contrat aidé est susceptible d'influencer la trajectoire ultérieure de son bénéficiaire.
- *la durée du traitement* : la durée du contrat aidé est définie comme la durée totale du contrat, reconductions éventuelles comprises (cf. graphique 2) à partir des déclarations à l'enquête⁶. Ce choix est en cohérence avec un des objectifs du contrat aidé, celui de favoriser l'accès à un emploi non aidé. La non prise en compte des reconductions conduit en effet à surestimer la fréquence des emplois aidés parmi les premiers emplois après un contrat aidé. Ceci est particulièrement vrai dans le cas des contrats aidés du secteur non marchand, pour lesquels les conventions initiales sont fréquemment reconduites. Cette discussion sur la définition retenue pour la durée du contrat est motivée par les résultats obtenus à partir d'autres sources : d'après l'enquête auprès des sortants de contrats aidés, la probabilité d'être en emploi six mois après la sortie de contrat aidé augmente avec la durée passée en contrat (Fendrich et Rémy, 2009). Celle-ci peut ainsi être le signe d'une meilleure « employabilité » du salarié ou le moyen pour lui d'acquérir des compétences valorisables lors d'une recherche d'emploi ultérieure.



Source : Panel 2008, Dares

Figure 2 – Répartition des contrats aidés selon leur durée, par type de contrat

6. Ce point devra être affiné par la suite en croisant avec les données administratives. On peut toutefois noter que la très grande majorité des salariés se souvient avoir été embauchée en contrat aidé entre le début du deuxième trimestre 2005 et la fin du deuxième trimestre 2007, et confirme l'exactitude des informations contenues dans les fichiers administratifs (dates d'entrée et de sortie du contrat, type de contrat, nom de l'employeur). La date de fin du contrat aidé est néanmoins l'élément le plus souvent infirmé par les salariés (dans un cas sur cinq). Les salariés ont en effet pu voir leur contrat renouvelé alors que la date retenue par l'enquêteur pour identifier le terme du contrat était celle de la fin du contrat initial.

3 Implémentation du *matching* sur données d'enquête

L'utilisation de données d'enquête pour évaluer la performance d'un programme s'avère souvent bénéfique dans la mesure où les informations recueillies se révèlent beaucoup plus riches que celles disponibles dans les sources administratives. La partie précédente a mis en évidence l'importance de la qualité des données pour assurer la validité de l'hypothèse identifiante du *matching*. Cependant, le recours à des données d'enquête soulève d'autres questions quant à l'implémentation du *matching*, que ce soit du fait de la taille de l'échantillon – qui conditionne en partie le choix de la méthode retenue – ou du fait de la prise en compte des pondérations – nécessaire à la généralisation des résultats à l'ensemble de la population.

3.1 Choix de la technique de *matching*

Quoique les témoins aient été échantillonnés pour ressembler aux bénéficiaires, l'appariement doit être amélioré *ex post* par la prise en compte des données collectées lors de l'enquête – plus riches que les données administratives initialement utilisées – et ce, afin de garantir au mieux l'hypothèse d'indépendance conditionnelle. Néanmoins, en présence de nombreuses variables de conditionnement, il est fort probable qu'apparier chaque bénéficiaire avec un témoin ayant exactement les mêmes caractéristiques soit impossible et que la qualité de l'estimateur s'en trouve affectée.

ENCADRÉ 1

Principales techniques de *matching* sur score de propension

Plusieurs méthodes utilisent le score de propension pour construire un contrefactuel pour chaque bénéficiaire (cf. Brodaty et al., 2007 ; Caliendo et Kopeinig, 2008). La méthode la plus simple, dite du « plus proche voisin », consiste à choisir le témoin dont le score est le plus proche de celui du bénéficiaire. L'effet moyen estimé sur l'ensemble E_1 des traités est alors la moyenne des écarts entre l'*outcome* de chaque traité et celui de son « voisin » (contrefactuel) : $\frac{1}{N_1} \sum_{E_1} (Y_{i1} - \hat{Y}_{i0})$.

Selon les données disponibles, il peut être judicieux d'associer plusieurs voisins à chaque bénéficiaire. Cette méthode est préconisée en particulier lorsque le groupe de contrôle est de taille importante puisque l'information disponible est alors exploitée de manière optimale. Diverses variantes existent :

- retenir *a priori* les n plus proches voisins : la situation contrefactuelle du bénéficiaire est alors estimée par la moyenne des *outcomes* de ces n plus proches voisins.
- définir *a priori* le rayon du voisinage (distance maximale entre les scores du bénéficiaire et des témoins) et faire la moyenne sur tous les voisins ainsi sélectionnés (radius *matching*).
- utiliser pour chaque bénéficiaire l'ensemble E_0 des témoins, et prendre la moyenne pondérée de leurs *outcomes*. Le poids donné à chaque témoin dépend négativement de la distance entre son score p_k et celui du bénéficiaire p_i :

$$\hat{Y}_{i0} = \frac{\sum_{E_0} K\left(\frac{\|p_i - p_k\|}{h}\right) \cdot Y_k}{\sum_{E_0} K\left(\frac{\|p_i - p_k\|}{h}\right)}$$

La forme fonctionnelle de cette dépendance négative (le « noyau » K et sa fenêtre h) est choisie par le chercheur et peut éventuellement donner un poids nul aux témoins les plus éloignés du bénéficiaire.

Pour l'ensemble de ces estimateurs, il n'existe généralement pas de formule donnant la variance de l'ATT, qui doit alors être estimée par *bootstrap*. Abadie et Imbens (2008) montrent cependant que le *bootstrap* n'est pas approprié pour estimer l'écart-type dans le cas d'appariements sur le plus proche voisin.

Pour palier ce problème de dimensionnalité, Rosenbaum et Rubin (1983) ont montré que l’hypothèse d’indépendance conditionnelle reste vérifiée pour toute fonction « équilibrante » des variables X (c’est-à-dire vérifiant la propriété $X \perp T | f(X)$). C’est en particulier le cas du score de propension, défini comme la probabilité d’entrer en contrat aidé : $p(X) = P(T = 1 | X)$. L’appariement ne se fait alors plus sur les X mais directement sur le score de propension $p(X)$, qui doit être estimé au préalable puisque non observé.

En pratique, on obtient $\hat{p}(X)$ par une modélisation logistique incluant l’ensemble des variables X qui influent simultanément sur la probabilité d’être traité et l’*outcome* Y :

$$p_i = \frac{\exp(f(X_i))}{1 + \exp(f(X_i))}$$

où $f(X)$ est une fonction des variables observables X pouvant faire intervenir des interactions entre les covariables, et choisie de manière à obtenir des échantillons comparables (« équilibrés » sur les X) de traités et de témoins. Il n’existe pas de règle précise pour déterminer les variables de conditionnement (Caliendo et Kopeinig, 2008). Néanmoins, les variables qui n’influent pas sur Y doivent être exclues de l’estimation quand bien même elles joueraient sur la probabilité d’être traité puisque, par définition, elles ne biaisent pas les estimations de l’effet du passage en contrat aidé sur l’*outcome*. En outre, la probabilité d’être traité ne doit pas être parfaitement expliquée sans quoi il n’existerait pas de support commun entre la distribution des scores des bénéficiaires et celles de leurs témoins ($0 < p(X) < 1$).

Le calcul des scores de propension individuels permet ensuite d’estimer les contrefactuels. La première étape consiste à restreindre l’échantillon au support commun : ont été supprimés de l’échantillon les bénéficiaires dont le score était supérieur à la valeur maximale prise par le score des témoins. Parmi les différentes techniques possibles (cf. encadré 1), nous avons privilégié un appariement avec une fonction noyau (gaussien)⁷ pour utiliser au mieux l’information disponible sur les témoins. Les écart-types ont été calculés par *bootstrap*.

3.2 Prise en compte des pondérations

Le *matching* devant être mis en œuvre sur des données d’enquête, la méthode classique doit être adaptée afin d’intégrer les pondérations, notées w_i , liées à l’échantillonnage et à la non-réponse. Pour généraliser les résultats observés sur l’échantillon à l’ensemble de la population, il est nécessaire de tenir compte du plan de sondage, de la non-réponse ainsi que de l’attrition entre les deux vagues du panel. La non-prise en compte des pondérations lors de l’estimation risquerait de conduire à des résultats biaisés (Davezies et d’Haultfœuille, 2009 ; Frölich, 2007). Ce risque de biais est particulièrement présent dans le panel 2008 puisque les bénéficiaires sont sur-représentés dans l’échantillon et que les témoins ont été sélectionnés de manière non aléatoire, conduisant ainsi à des échantillons non directement représentatifs de la population d’intérêt.

7. Le choix de la fenêtre h résulte d’un arbitrage entre biais et précision. Nous avons retenu l’approche de Silverman et fixé la fenêtre optimale selon la « règle du pouce » : $\hat{h} = 1,06 \min(\hat{\sigma}_0, (IQR/1,349)).n_0^{-1/5}$, où $\hat{\sigma}_0$ (resp. *IQR*) désigne l’écart-type (resp. écart inter-quartile) du score estimé sur l’échantillon des témoins, de taille n_0 .

Le calcul des pondérations

Afin, d'assurer la représentativité des estimations conduites sur le Panel 2008, les poids de sondage des individus ayant répondu à l'enquête ont été modifiés pour corriger de la non-réponse en vague 1, ainsi que de l'attrition entre les deux vagues. Le principe général consiste à attribuer à chaque répondant r une probabilité de réponse p_r qui, multipliée par la probabilité d'inclusion, permet de disposer d'estimateurs sans biais sur la population des répondants :

$$\hat{Y} = \sum_R \frac{1}{\pi_r} \frac{y_r}{p_r}$$

où R désigne l'ensemble des répondants, y_r la valeur de la variable y pour l'individu répondant r et π_r sa probabilité d'inclusion.

Les vraies probabilités de réponse étant inconnues, la démarche consiste à calculer les poids par un modèle de non-réponse et par calage (Caron, 2005). Pour ce faire, un modèle logistique modélisant la probabilité de réponse a été estimé séparément pour chaque population de bénéficiaires et de témoins. Le croisement des variables explicatives du comportement de réponse permet d'effectuer des regroupements d'individus, dits « groupes de réponse homogène » (GRH) car les individus qui les composent sont supposés avoir un comportement de réponse similaire. Les probabilités de réponse au sein de chaque GRH h sont calculées comme le rapport pondéré (par les probabilités d'inclusion π_k) du nombre de répondants r_h dans la sous-population h sur le nombre d'individus n_h enquêtés dans cette même sous-population :

$$\hat{p}_h = \frac{\sum_{r_h} 1/\pi_k}{\sum_{n_h} 1/\pi_k}$$

Cette méthode générale a été mise en œuvre en deux étapes : i) pour corriger de la non-réponse à la vague 1, en mobilisant les variables présentes dans la base de sondage (caractéristiques individuelles, trajectoires sur le marché du travail, caractéristiques du contrat aidé de référence) ; ii) pour corriger de l'attrition entre la première et la seconde vague, en mobilisant de surcroît les variables recueillies lors de la première vague. Il n'y a en effet aucune raison pour que les variables expliquant la non-réponse soient identiques à celles expliquant l'attrition. La prise en compte des variables d'enquête en deuxième étape permet également d'enrichir le modèle et de mieux modéliser le comportement d'attrition^a. Un calage sur marges a également été mis en œuvre à chaque étape afin de redonner à l'échantillon une structure comparable à celle de la base de sondage.

Notons que certains individus ont été repérés comme hors champ parmi les enquêtés^b : ils ont été traités comme répondants – avec dilatation de leur poids – puis supprimés de l'échantillon en fin de première étape avec leurs poids redressés et calés (ce qui permet de déterminer *a posteriori* la taille de la population de référence dans le champ). Le calage de l'échantillon de la vague 2 – qui compte uniquement les répondants à la vague 2 appartenant au champ de l'enquête – s'effectue alors sur les marges de l'échantillon des répondants à la vague 1 (hors champ exclus), pondéré par les poids calés de première étape. Cette méthode permet d'estimer le nombre d'individus dans le champ de l'enquête, et assure par conséquent une bonne représentativité de l'échantillon (Brion *et al.*, 2005 ; Brilhault et Caron, 2005).

^a. Cf. tableau 8 en annexes pour une définition détaillée des variables mobilisées pour le redressement de l'échantillon.

^b. On distingue parmi les hors champs les bénéficiaires ayant déclaré ne pas avoir eu de contrat aidé, les témoins ayant déclaré un contrat aidé sur la période de référence et les individus repérés comme hors cible par les enquêteurs.

Pour mettre en oeuvre une estimation par *matching* à partir de données d'enquête, les pondérations – notées w – doivent intervenir à différentes étapes (Frölich, 2007). Dans le cas particulier du Panel 2008, elles interviennent de la manière suivante :

1. L'estimation du score de propension doit d'abord être réalisée sur données pondérées afin que l'équation de sélection soit représentative de celle sur la population d'origine. La log-vraisemblance pondérée s'écrit :

$$\ln(L(\beta)) = \sum_{i=1}^n [w_i \cdot T_i \cdot \ln(F(X_i\beta)) + w_i \cdot (1 - T_i) \cdot \ln(1 - F(X_i\beta))]$$

où F désigne la forme fonctionnelle retenue pour estimer la probabilité d'entrer en traitement T (ici la loi logistique).

2. L'estimation par noyau du contrefactuel de chaque bénéficiaire i doit également inclure les pondérations du groupe de témoins E_0 , et s'écrit :

$$\hat{Y}_{i0} = \frac{\sum_{E_0} w_k \cdot K\left(\frac{\|p_i - p_k\|}{h}\right) \cdot Y_k}{\sum_{E_0} w_k \cdot K\left(\frac{\|p_i - p_k\|}{h}\right)}$$

avec p_i le score du bénéficiaire i , p_k celui du témoin k , w_k le poids du témoin k , h la fenêtre d'estimation, et K la fonction noyau.

3. La pondération intervient ensuite pour estimer l' ATT . L'*outcome* contrefactuel de chaque bénéficiaire étant estimé grâce aux *outcomes* des témoins, l' ATT est habituellement obtenu en faisant la moyenne sur l'ensemble des bénéficiaires des écarts entre leur *outcome* observé et leur *outcome* contrefactuel. En présence de données pondérées, l' ATT est alors calculé comme une moyenne pondérée des effets individuels pour chaque bénéficiaire :

$$\hat{\Delta}^{ATT} = \frac{\sum_{E_1} w_i \cdot (Y_{i1} - \hat{Y}_{i0})}{\sum_{E_1} w_i}$$

où E_1 désigne le groupe de bénéficiaires, Y_{i1} l'*outcome* observé du bénéficiaire i , et \hat{Y}_{i0} l'*outcome* contrefactuel de ce même bénéficiaire estimé par noyau sur les témoins. Ainsi, les poids intervenant dans l'estimation de l' ATT résultent *in fine* de la combinaison des poids d'échantillonnage – corrigés de la non-réponse et de l'attrition – et des poids $K\left(\frac{\|p_i - p_k\|}{h}\right)$ générés par l'estimateur à noyau.

Une fois obtenu l'estimateur $\hat{\Delta}^{ATT}$, il convient de mesurer sa précision. L'écart-type de l' ATT est généralement calculé par *bootstrap*. Il s'agit donc de répliquer les échantillons des bénéficiaires et des témoins. Pour les échantillons ainsi rééchantillonnés, la somme des poids initiaux n'a plus aucune raison de correspondre à la taille de la population, et la structure des données n'est donc plus statistiquement similaire à celle de l'échantillon pondéré initial. Pour calculer l'écart-type de l' ATT dans le cas de données pondérées, nous suivons la méthode de *bootstrap* proposée par Davezies et D'Haultfœuille (2009), qui suggèrent de prendre en compte dans l'algorithme la variabilité créée par le caractère aléatoire de la non-réponse dans l'échantillon.

Plutôt que de tirer avec remise des échantillons de taille fixe, les auteurs préconisent de tirer des échantillons de taille variable : pour chaque réplique, la taille n_{b1} (n_{b0}) de l'échantillon de bénéficiaires (de témoins) est déterminée par tirage dans une loi binomiale $B(N_1, n_1/N_1)$ ($B(N_0, n_0/N_0)$), où N_1 (N_0) désigne la taille de la population totale des bénéficiaires (des témoins), et n_1 (n_0) le nombre de bénéficiaires (témoins) répondants dans

l'échantillon. Il s'agit ensuite d'un tirage aléatoire simple avec remise d'un échantillon de bénéficiaires de taille n_{b1} d'une part, et d'un échantillon de témoins de taille n_{b0} d'autre part. Pour chacun de ces B tirages de bénéficiaires et de témoins, les poids peuvent être recalculés – selon une démarche similaire à celle retenue pour l'échantillon initial (modèle de non-réponse et calage en deux étapes) – en appliquant comme pondération le nombre de fois où chaque individu a été tiré. À chaque réplication, le nouveau jeu de pondérations est alors mobilisé pour le recalcul du score, ce qui permet de capter l'incertitude associée à l'estimation du score de propension (y compris l'aléa généré par la variabilité du support commun sur chaque itération).

ENCADRÉ 3

Algorithme de *bootstrap*

Le principe du *bootstrap* consiste à spécifier un processus générateur de données (DGP) en remplaçant les paramètres et distributions de probabilité inconnues par des estimations empiriques de ces derniers. La distribution *bootstrap* est approximée par des simulations. Soit un échantillon de n variables aléatoires $\mathbf{X}_n = (X_1, X_2, \dots, X_n)$ indépendantes et identiquement distribuées (*iid*). On tire dans \mathbf{X}_n , avec remise, B échantillons de taille n , qui nous donnent B pseudo-échantillons :

$$\mathbf{X}_n^{*(b)} = (X_1^{*(b)}, X_2^{*(b)}, \dots, X_n^{*(b)})$$

Cet ensemble de pseudo-échantillons permet de calculer la distribution *bootstrap* et les caractéristiques de la statistique $\hat{\Delta}(\mathbf{X}_n^*)$. En pratique, il s'agit de calculer les valeurs $\hat{\Delta}(\mathbf{X}_n^{*(b)})$ sur chacune des B réplifications. L'algorithme général est le suivant :

1. Calcul de l'estimateur $\hat{\Delta}$ du paramètre inconnu,
2. Tirage avec remise d'un échantillon de taille n et calcul de la valeur de la statistique d'intérêt sur la base des données simulées.
3. Répétition de l'étape précédente $B - 1$ fois afin d'obtenir B réalisations de la statistique *bootstrap* $\left\{ \hat{\Delta}^{*(b)} \right\}_{b=1, \dots, B}$. La fonction de distribution empirique formée par ces B réalisations est une approximation du DGP *bootstrap*.

La prise en compte des pondérations nécessite de complexifier l'étape 2, qui peut se décomposer ainsi (Davezies et D'Haultfœuille, 2009) :

1. Tirage de la taille n_{b1} (n_{b0}) de l'échantillon des bénéficiaires (témoins) dans une loi binomiale $B \sim (N_1, n_1/N_1)$ ($B \sim (N_0, n_0/N_0)$), où n_1 (n_0) la taille de l'échantillon des bénéficiaires (témoins) répondants,
2. Tirage aléatoire simple avec remise d'un échantillon de bénéficiaires (témoins) de taille n_{b1} (n_{b0}),
3. Recalcul des nouvelles pondérations selon la méthode employée sur l'échantillon original,
4. Estimation de $\hat{\Delta}^b$ selon les trois étapes présentées en section 3.2.

Les intervalles de confiance sont alors obtenus par la méthode percentile : il s'agit de remplacer les quantiles théoriques $q_{\frac{\alpha}{2}}$ et $q_{1-\frac{\alpha}{2}}$ par les quantiles correspondants de la distribution *bootstrap* empirique, obtenus en ordonnant les B valeurs $\hat{\Delta}^b$.

Notons enfin que la prise en compte des pondérations peut éventuellement compliquer l'interprétation du support commun : de manière générale, les individus sous-représentés dans l'échantillon sont davantage susceptibles de se trouver en dehors du support commun et donc d'être exclus des estimations, or ce sont ces individus qui présentent les poids de sondage les plus élevés. Le cas échéant, la proportion des individus hors du support commun

dans la population est supérieure à celle dans l'échantillon (Bryson *et al.*, 2002). Ce constat devra être gardé à l'esprit lors de l'interprétation des résultats.

4 Résultats de l'évaluation

4.1 Identification : interprétation du score

La qualité de l'estimation du score de propension détermine en grande partie la validité des résultats de l'évaluation *via* la crédibilité de l'hypothèse d'indépendance conditionnelle. Les données disponibles permettent de tenir compte de beaucoup de facteurs susceptibles d'intervenir dans le processus de sélection des bénéficiaires (Lechner, 2007) : les données sur les caractéristiques individuelles et le parcours du demandeur d'emploi ainsi que les données de contexte local (taux de chômage local, taux d'accès aux contrats aidés) permettent de tenir compte de l'*éligibilité* des demandeurs d'emploi au programme, ainsi que d'une éventuelle *sélection par le conseiller de l'ANPE* ; les variables d'enquête telles que l'objectif avant l'inscription à l'ANPE au deuxième trimestre 2005, les freins à la recherche d'emploi, le type d'emploi recherché ou encore l'état de santé permettent quant à elles de tenir compte de l'*auto-sélection du demandeur d'emploi*⁸ et de capter en partie son hétérogénéité habituellement inobservée dans les données.

Les résultats détaillés de l'estimation de la probabilité d'entrée en traitement par type de contrat sont donnés dans le tableau 6 en annexe. Seuls quelques résultats sont ici commentés. Être en assez mauvaise santé joue négativement sur l'accès aux contrats aidés réservés aux allocataires de minima sociaux. En revanche, avoir un handicap reconnu administrativement joue positivement sur la probabilité d'entrée en CAE et en CIE, contrats non réservés aux allocataires de minima sociaux. Ce public est de fait souvent ciblé par les pouvoirs publics (taux de prise en charge spécifique), ce qui peut inciter les agents de l'ANPE à proposer en priorité ces contrats aux travailleurs handicapés, et les employeurs à recruter en priorité ces travailleurs. Globalement, le taux d'accès aux contrats aidés du secteur non marchand est plus élevé lorsque les conditions de vie des personnes sont modestes : percevoir la couverture maladie universelle complémentaire (CMUC) destinée aux personnes à faibles revenus, ne pas payer l'impôt sur le revenu, percevoir un minimum social dans le cas du CAV.

Le parcours professionnel est particulièrement important pour appréhender le lien que les témoins et les bénéficiaires entretiennent avec le marché du travail et expliquer en partie leur devenir professionnel (Crépon et Kamionka, 2006). Avoir eu un parcours heurté sur le marché du travail (alternance de périodes courtes d'emploi et de chômage) ou être resté principalement au chômage ou inactif augmente les chances d'accès à un contrat du secteur non marchand par rapport au fait d'avoir été principalement en emploi. Ne pas avoir de permis de conduire ou de connexion internet joue négativement sur la probabilité d'entrer dans un contrat aidé du secteur marchand. Les employeurs du secteur marchand sélectionneraient ainsi les demandeurs d'emploi les plus mobiles et proches de l'emploi, tandis que les facteurs témoignant de conditions de vie plus défavorables joueraient positivement sur l'accès à un contrat aidé du secteur non marchand.

Deux critères importants permettent de juger de la qualité des estimations : l'obtention d'échantillons équilibrés sur les variables de contrôle X conditionnellement au score estimé (cf. tableau 7)⁹, ainsi que l'existence d'un support commun. La figure 3 représente les distri-

8. Ce dernier choisissant d'entrer ou non en programme compte tenu des avantages ou des inconvénients qu'il anticipe de sa participation.

9. Afin d'obtenir des échantillons de témoins et de bénéficiaires équilibrés, les équations présentées dans le tableau 6 ont été modifiées par l'introduction de termes d'interaction entre certaines variables.

butions des scores de propension des témoins et des bénéficiaires, par type de contrat. Deux points sont intéressants à noter : d'une part, en raison de la prise en compte des pondérations dans l'estimation de la probabilité d'entrer en contrat, les scores sont particulièrement faibles (les bénéficiaires représentent moins de 6% de la population éligible); d'autre part, en raison du mode d'échantillonnage des témoins, on se trouve dans le cas favorable d'un large support commun (peu de bénéficiaires ne peuvent être appariés).

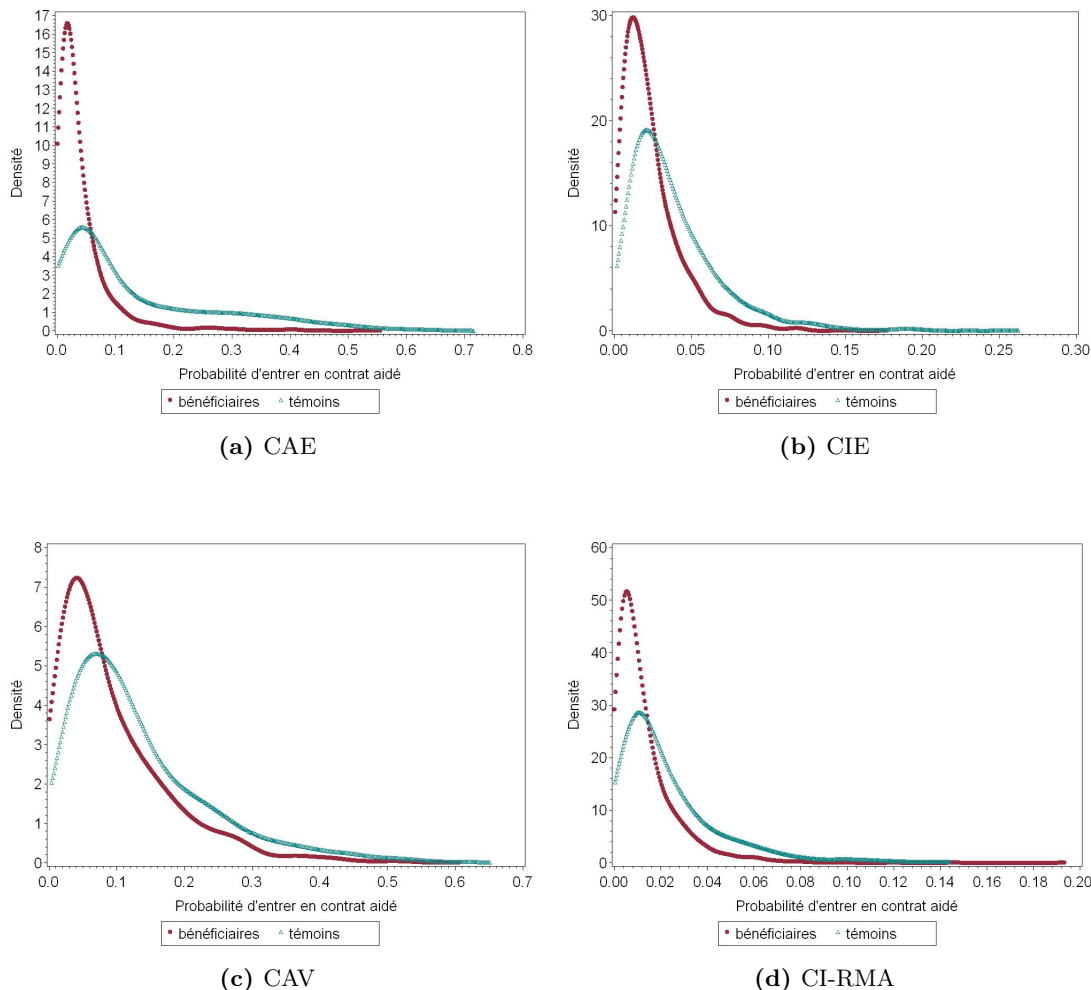


Figure 3 – Comparaisons des distributions des scores de propension

4.2 Impact du passage en contrat aidé sur les trajectoires

Une fois les scores de propension calculés et le support commun identifié, les trajectoires des bénéficiaires et de leurs témoins peuvent être comparées. Une première question qui se pose est de savoir à partir de quelle date commencer les comparaisons. En effet, les bénéficiaires des différents contrats aidés n'ont pas tous été traités au même moment ni pour une même durée. Aussi à une date donnée, les individus auront été traités selon une durée variable et seront sortis de traitement depuis plus ou moins longtemps, voire non encore sortis (et donc censurés). L'approche la plus pertinente consiste à se placer au début du traitement et à étudier le devenir des bénéficiaires à partir de cette date (Lechner et Wunsch, 2007; Hujer, Caliendo et Thomsen, 2003). Les témoins n'étant pas entrés en contrat aidé sur la

période d'éligibilité des bénéficiaires, il a été nécessaire de leur affecter une date d'entrée fictive pour comparer leur trajectoire avec celle des bénéficiaires (Lechner, 1999 ; Lechner et Wunsch, 2007). Pour déterminer cette date d'entrée, une régression a été effectuée sur la date d'entrée en contrat des bénéficiaires, et les coefficients de cette régression ainsi qu'un résidu tiré dans la distribution des résidus ont été utilisés pour prédire les dates d'entrée en contrat. Lorsque celles-ci ne se trouvaient pas dans la fenêtre de référence (deuxième trimestre 2005 - deuxième trimestre 2007), les individus ont été supprimés de l'échantillon. Les résultats présentés font référence à ces dates simulées pour les témoins.

Dans le Panel 2008, la grande majorité des contrats aidés considérés s'est terminée entre janvier 2006 et juin 2008. Compte tenu des dates d'enquêtes (fin 2008 pour la vague 1 et fin 2009 pour la vague 2), la fenêtre d'observation après le passage en contrat aidé varie donc de dix-huit à quarante-huit mois. Cette fenêtre d'observation, plus courte pour certains bénéficiaires, est la contrepartie de la large fenêtre d'entrée en traitement. Ainsi, le devenir des bénéficiaires à la sortie du contrat aidé ne pourra être observé que sur un laps de temps relativement court pour les bénéficiaires dont le contrat s'est terminé plus tardivement. Pour que le phénomène de censure ne concerne pas trop d'individus, les situations des bénéficiaires et de leurs témoins ont été comparées vingt-quatre mois après l'entrée en contrat aidé. Les tableaux 2 et 3 présentent les résultats du *matching* sur les variables d'intérêt retenues : le fait d'être en emploi à durée indéterminée vingt-quatre mois après l'entrée en contrat aidé, ainsi que la durée cumulée en CDI sur cette même période.

Les résultats présentés ici ne le sont qu'à titre illustratif dans la mesure où ils sont susceptibles de modifications ultérieures du fait d'approfondissements méthodologiques. D'autres *outcomes* pourront de plus être étudiés pour enrichir les conclusions. Globalement, et comme constaté dans de nombreux travaux, l'effet moyen d'un passage par un contrat du secteur non marchand est significativement négatif sur la probabilité d'être en CDI vingt-quatre mois après l'entrée en contrat aidé, tandis que l'effet moyen du passage par un contrat du secteur marchand est significativement positif. Les effets négatifs sur la durée cumulée en emploi à durée indéterminée sont la confirmation d'une situation de *locking-in*. Ces résultats ne sont que peu modifiés par la prise en compte des pondérations¹⁰. Ainsi, la probabilité pour une personne embauchée en CAE entre avril 2005 et le juin 2007 de se déclarer en CDI vingt-quatre mois après le début de son contrat est de 17 points plus faible qu'en l'absence de traitement dans l'estimation pondérée, contre 16 points dans l'estimation non pondérée. Quant au passage par un CIE, il augmente significativement la probabilité d'être en emploi à durée indéterminée de 9,6 points dans le modèle pondéré, et de 8,6 points dans le modèle non pondéré.

Tableau 2 – Estimation de l'effet moyen du passage en contrat aidé, non pondéré

$\hat{\Delta}^{ATT}$	CAE	CIE	CAV	CI-RMA
Être en CDI deux ans après l'entrée en contrat aidé	-0,173***	0,096***	-0,153***	0,131***
Durée cumulée en CDI sur les deux ans suivant l'entrée en contrat	-4,737***	-1,891***	-3,809***	-1,198***

Notes : écart-types calculés par *bootstrap* avec 500 réplifications.

10. Cf. Green *et al.* (2001) pour un constat similaire.

Tableau 3 – Estimation de l’effet moyen du passage en contrat aidé, pondéré

$\hat{\Delta}^{ATT}$	CAE	CIE	CAV	CI-RMA
Être en CDI deux ans après l’entrée en contrat aidé	-0,160***	0,086***	-0,148***	0,125***
Durée cumulée en CDI sur les deux ans suivant l’entrée en contrat	-4,537***	-1,785***	-3,483***	-1,042***

Notes : écart-types calculés par *bootstrap* avec 500 répliquions. Pour des raisons de lourdeur des calculs, les résultats présentés ici ne prennent pas en compte le recalcul des poids à chaque itération.

L’effet positif lié au passage par un contrat du secteur marchand peut s’expliquer par plusieurs facteurs : acquisition par le bénéficiaire d’une expérience professionnelle et de compétences valorisables lors d’une recherche d’emploi ultérieure, et ce d’autant plus que l’emploi exercé est proche d’un emploi non aidé ; maintien du salarié par l’employeur dans l’entreprise une fois la période de versement de la subvention terminée (notamment lorsque le contrat est conclu à durée indéterminée).

Dans le secteur non marchand, les effets observés peuvent être liés au fait que le métier exercé pendant le contrat aidé est trop éloigné des postes sur lesquels le demandeur d’emploi est susceptible de postuler ultérieurement, ou encore au fait que le passage par un tel contrat constitue un signal négatif pour son futur employeur potentiel. Toutefois, le recul avec lequel la situation du bénéficiaire est observée n’est pas nécessairement suffisant pour apprécier l’effet à moyen terme d’un passage en contrat aidé sur sa trajectoire. Enfin, l’accès à un CDI n’est pas le seul *outcome* à étudier pour évaluer un passage par une mesure.

Au sein de chaque secteur, marchand et non marchand, les demandeurs d’emploi passés par un contrat réservé aux bénéficiaires de minima sociaux sont globalement plus souvent en emploi à durée indéterminée vingt-quatre mois après leur entrée que ceux passés par un autre contrat. Ces résultats sont toutefois à prendre avec précaution dans la mesure où la durée du contrat est ici mesurée avec les données d’enquête qui n’ont pas encore été confrontées aux données administratives. L’effet du CIE pourrait ainsi être plus favorable en tenant compte de ces dernières tandis que celui du CAV pourrait l’être moins.

4.3 Hétérogénéité du traitement

Pour affiner l’analyse des effets du traitement, il est nécessaire de tenir compte de son hétérogénéité. Le tableau 4 présente quelques résultats pour le CAE et le CI-RMA pour illustrer sur quelques exemples les différences en termes de durée de contrat ou encore d’accompagnement. Dans le cas du CI-RMA, l’effet positif du traitement sur le taux d’accès à un CDI augmente avec sa durée. Au contraire, les chances d’accéder à un CDI étant réduites par le passage en CAE, l’effet de « *locking-in* » initial ne se trouve pas contrebalancé par une meilleure insertion à l’issue. Aussi, un contrat court diminue-t-il cet « effet pervers » du traitement en évitant que le bénéficiaire ne diminue son effort de recherche d’emploi trop longtemps pendant son passage en contrat aidé.

Une autre dimension importante du traitement est sa qualité, approchée ici par le fait d’avoir été accompagné en interne par un tuteur (ou une autre personne chez l’employeur) pendant le traitement. A titre d’exemple, bénéficiaire d’un accompagnement amoindrit l’effet négatif du traitement dans le cas du CAE et intensifie son effet positif dans le cas du CI-RMA.

Tableau 4 – Estimation de l’effet moyen du passage en contrat aidé

$\hat{\Delta}^{ATT}$	Non pondéré		pondéré	
	CAE	CI-RMA	CAE	CI-RMA
<i>Durée du contrat inférieure à 6 mois</i>				
Être en CDI deux ans après				
l’entrée en contrat aidé	-0,109***	0,076***	-0,119***	0,073**
<i>Durée du contrat comprise entre 6 et 12 mois</i>				
Être en CDI deux ans après				
l’entrée en contrat aidé	-0,156***	0,132***	-0,162***	0,119**
<i>Durée du contrat supérieure à 12 mois</i>				
Être en CDI deux ans après				
l’entrée en contrat aidé	-0,230***	0,182***	-0,224 ***	0,148***
<i>Accompagnement</i>				
Être en CDI deux ans après				
l’entrée en contrat aidé	-0,156***	0,149***	-0,161***	0,165***
<i>Sans accompagnement</i>				
Être en CDI deux ans après				
l’entrée en contrat aidé	-0,197***	0,119***	-0,186***	0,110***

Notes : écart-types calculés par *bootstrap* avec 500 réplifications. Pour des raisons de lourdeur des calculs, les résultats présentés ici ne prennent pas en compte le recalcul des poids à chaque itération.

5 Conclusion

Cette étude présente la manière dont a été construit le Panel 2008 dans l’optique d’une évaluation causale par *matching*. Il avait en effet été décidé d’évaluer l’impact des contrats aidés en comparant la situation de bénéficiaires de contrat aidé à celle de « témoins » représentatifs. Le protocole de l’enquête (échantillonnage, questionnaire, collecte des données, etc.) a ainsi été conçu avec le souci permanent de garantir autant que possible la comparabilité entre les bénéficiaires de contrat aidé et leurs témoins.

Plusieurs points méthodologiques ont été soulevés lors de l’implémentation du *matching*, certains d’entre eux restant encore à approfondir. L’accent a été mis plus particulièrement sur l’adaptation des méthodes existantes à des données d’enquête. Dans le Panel 2008, la pondération des individus résulte de l’échantillonnage, de la correction de la non-réponse et du calage sur marges. Les poids interviennent à plusieurs étapes des estimations, la plus délicate étant le calcul de la précision des estimateurs obtenus par *matching*. Les écart-types de ces estimateurs sont calculés par *bootstrap*. Pour cela, une méthode pertinente est celle proposée par Davezies et D’Haultfoeuille (2009). Les résultats des estimations – encore exploratoires à ce stade et fournis à titre illustratif – permettent ainsi de comparer la sensibilité de l’estimation et de la précision de l’*ATT* à la prise en compte des pondérations.

Par la suite, la méthode devra être améliorée pour davantage prendre en compte la nature dynamique du traitement considéré. En particulier, le score de propension pourra être estimé sur la base d’un modèle de durée modélisant la durée avant l’entrée en contrat et les *outcomes* pourront être mesurés dans la durée par des fonctions de survie. En outre, des travaux de recherche récents (Fredriksson et Johansson, 2008 ; Crépon *et al.*, 2009 ; Kastoryano et Van der Klaauw, 2010) proposent une autre manière de définir les témoins, d’apparence plus flexible, pour réaliser un *matching* dynamique. Le problème de l’approche statique présentée ici vient de ce que l’on impose une fenêtre (d’une durée de deux ans) – plutôt qu’une date fixe – d’entrée en traitement, pour comparer ensuite des *outcomes* en référence à cette date d’entrée en traitement. Or, les individus qui ont trouvé un emploi rapidement sont davantage susceptibles de se trouver dans le groupe des non-traités imposé par la fenêtre, sans qu’il

soit pour autant possible de savoir s'ils seraient entrés en traitement avant ou après la fin de la fenêtre en l'absence de la censure provoquée par leur sortie du chômage (durée avant traitement inobservée). Prendre en compte dans le groupe de témoins les individus dont la durée avant traitement est censurée par l'obtention d'un emploi, tout en excluant les individus qui ont bénéficié du traitement sur la même période, biaise l'échantillon des témoins en faveur d'épisodes plus courts de chômage (Fredriksson et Johansson, 2008) et risque de minorer l'effet du traitement. Pour éviter cet effet de sélection sur les témoins, il est nécessaire d'imposer une durée passée au chômage pour les témoins au moins égale à celle des bénéficiaires avant leur entrée en traitement. De fait, les premiers résultats obtenus en imposant cette condition, corrigent à la hausse les effets du traitement sur les traités. Dans la version dynamique, chaque individu (témoin ou bénéficiaire) non entré en contrat aidé avant la date t et encore au chômage à cette date pourrait témoigner pour un bénéficiaire entré en contrat en t . La définition du groupe de témoins est donc variable en fonction de la date d'entrée en contrat et permet un appariement de meilleure qualité. On estime alors l'effet d'être entré en traitement à la date t , soit autant d'effets que de dates d'entrée, et la moyenne de tous ces effets permet de reconstituer l'effet global de la politique.

Cette approche pourra être comparée à la méthode du « *timing of events* » proposée par Abbring et Van den Berg (2003). Cette méthode consiste en un modèle à risques concurrents à hasards proportionnels, permettant de tenir compte d'une éventuelle hétérogénéité inobservée résiduelle. Par rapport au *matching*, cette technique relâche l'hypothèse d'indépendance conditionnelle – relative uniquement aux caractéristiques observables –, mais impose en revanche une forme paramétrique au hasard (Lalive, van Ours et Zweimüller, 2008).

Enfin, d'autres *outcomes* pourront être étudiés, comme la situation vis-à-vis de l'emploi à différentes dates ou encore différentes approches de la qualité de l'emploi obtenu (type de contrat de travail, temps de travail, etc.), la récurrence du passage en contrat aidé ou celle au chômage, etc.

Annexes

	Secteur marchand		Secteur non marchand	
	CIE	CI-RMA	CAE	CAV
Publics éligibles	Les publics sont définis au niveau régional.	Bénéficiaires du revenu minimum d'insertion (RMI), de l'allocation de solidarité spécifique (ASS), de l'allocation de parent isolé (API), de l'allocation aux adultes handicapés (AAH) (depuis avril 2006).	Les publics sont définis au niveau régional.	Bénéficiaires du revenu minimum d'insertion (RMI), de l'allocation de solidarité spécifique (ASS), de l'allocation de parent isolé (API), de l'allocation aux adultes handicapés (AAH) (depuis mars 2006).
Type de contrat de travail	CDI ou CDD de 24 mois maximum.	CDI ou CDD de 6 mois minimum ou contrat de travail temporaire renouvelables deux fois dans la limite de 18 mois.	CDD de 6 mois minimum renouvelable dans la limite de 24 mois.	CDD de 24 mois, renouvelable dans la limite de 36 mois (voire 60 mois pour les plus de 50 ans et les travailleurs handicapés). Par dérogation, CDD d'une durée comprise entre 6 et 24 mois, renouvelable 2 fois dans la limite de 36 mois.
Temps de travail	Temps partiel ou temps complet ; s'il est à temps partiel, la durée hebdomadaire de travail doit être d'au moins 20 heures.	Temps partiel ou temps complet ; s'il est à temps partiel, la durée hebdomadaire de travail doit être d'au moins 20 heures.	Temps partiel ou temps complet ; s'il est à temps partiel, la durée hebdomadaire de travail doit être d'au moins 20 heures.	Durée hebdomadaire fixée à 26 heures. À partir du début 2006, elle peut être comprise entre 20 et 26 heures lorsque l'embauche est réalisée par les ateliers et chantiers d'insertion ou par une entreprise ou une association de services à la personne.
Prise en charge de l'État	Aide mensuelle de l'État fixée par arrêté du préfet de région, dans la limite de 47 % du Smic.	Aide forfaitaire égale au montant du RMI-RSA garanti à une personne isolée.	Aide mensuelle de l'État fixée par arrêté du préfet de région, dans la limite de 95 % du Smic.	Aide forfaitaire égale au montant du RMI-RSA garanti à une personne isolée. Aide dégressive complémentaire : 75 % du coût restant en charge de l'employeur la première année, 50 % les années suivantes. L'aide complémentaire de l'État a été égale à 90 % du coût restant à la charge de l'employeur les 6 premiers mois jusque fin 2007. Pour les ateliers et chantiers d'insertion, le taux de l'aide est fixé à 90 % pendant toute la durée d'exécution du contrat.
Autres avantages pour l'employeur	Cumul possible avec certains dispositifs d'allègement ou d'exonération de cotisations patronales de sécurité sociale.	Cumul possible avec certains dispositifs d'allègement ou d'exonération de cotisations sociales.	Exonérations des cotisations et contributions patronales de sécurité sociale dans la limite du Smic, de la taxe sur les salaires, de la taxe d'apprentissage, de la taxe due au titre de l'effort de construction	Exonération de cotisations et de contributions patronales de sécurité sociale. Aide supplémentaire de 1500 € en cas d'embauche en CDI avant la fin du contrat
Employeurs éligibles	Tous les employeurs entrant dans le champ de l'assurance chômage et les employeurs de pêche maritime. Les particuliers employeurs sont exclus.	Tous les employeurs entrant dans le champ de l'assurance chômage et les employeurs de pêche maritime. Les particuliers employeurs sont exclus.	Les collectivités territoriales et autres personnes morales de droit public, les personnes morales chargées de la gestion d'un service public, les autres organismes de droit privé à but non lucratif.	Les collectivités territoriales et autres personnes morales de droit public, les personnes morales chargées de la gestion d'un service public, les autres organismes de droit privé à but non lucratif.

Tableau 5 – Réglementation sur les contrats aidés

	Type de contrat			
	CAE	CIE	CAV	CI-RMA
Caractéristiques individuelles ‘				
<i>Age</i>				
moins de 26 ans	-0,166	-0,550 ***	0,575 **	0,425
50 ans et plus	-0,395 ***	-0,325 ***	0,711 ***	0,898 ***
Femme	0,725 ***	-0,063	0,424 ***	0,088
<i>Diplôme</i>				
< CAP-BEP (ref)				
CAP-BEP	0,032	0,442 ***	-0,049	0,061
niveau Bac	-0,013	0,483 ***	0,182	0,361
supérieur au Bac	-0,321 **	0,162	0,012	-0,070
<i>Santé</i>				
en très bonne santé (ref)				
en assez bonne santé	0,059	-0,058	-0,055	-0,008
en assez mauvaise santé	0,036	0,021	-0,576 ***	-0,367 *
en très mauvaise santé	-0,243	-0,107	-0,027	-0,026
<i>Handicap</i>				
pas de handicap (ref)				
handicap reconnu administrativement	0,359 **	0,696 ***	-0,502 ***	-0,752 ***
handicap non reconnu administrativement	-0,479 **	0,049	0,139	-0,074
<i>Origine</i>				
au moins un parent d'origine extra-européenne	-0,205	-0,184	-0,360 **	-0,220
Trajectoire sur le marché du travail avant le T2-2005				
<i>Parcours avant l'inscription au T2-2005</i>				
toujours ou principalement en emploi (ref)				
alternance de périodes d'emploi, de chômage et d'inactivité	0,445 ***	0,279 ***	0,346 ***	-0,229
toujours ou principalement au chômage	0,372 **	0,111	0,382 **	0,305
toujours ou principalement inactif	0,307 **	-0,030	0,380 *	-0,333
<i>Durée cumulée au chômage de 1997 au T2-2005</i>				
moins d'un mois (ref)				
de 1 à 6 mois	0,132	0,301 **	-0,406 *	-0,160
de 6 mois à un an	0,315 *	0,397 **	-0,121	0,063
d'un an à deux ans	0,361 *	0,436 **	-0,130	-0,381
plus de deux ans	0,525 **	0,044	-0,110	-0,781 *
<i>Durée cumulée au chômage indemnisé de 1997 à T2-2005</i>				
moins d'un mois (ref)				
de 1 à 6 mois	-0,494 ***	-0,110	-0,095	0,276
de 6 mois à un an	-0,190	0,076	0,007	0,713 **
d'un an à deux ans	-0,157	0,319 **	-0,139	0,829 ***
plus de deux ans	0,234	0,661 ***	0,118	1,259 ***
Nombre d'épisodes de chômage depuis 1997	0,011	-0,011	0,039 **	-0,012
Variables liées à la recherche d'emploi				
<i>Objectif lors de l'inscription à l'ANPE</i>				
travailler à tout prix (ref)				
travailler dans un domaine précis	-0,234 **	-0,393 ***	-0,267 **	-0,432 ***
autres	-0,958 ***	-1,323 ***	-0,919 ***	-1,379 ***
<i>Motif d'inscription à l'ANPE</i>				
licenciement (ref)				
démission	0,474 *	-0,166	0,762 **	-0,283
fin de contrat ou de mission d'intérim	0,551 ***	-0,404 ***	0,662 ***	-0,253
1 ^{ère} entrée ou reprise après interruption de moins de 6 mois	0,804 ***	-0,025	0,881 ***	-0,265
<i>Freins à la recherche d'emploi en 2005</i>				
aucun frein (ref)				
problèmes de santé	0,885 ***	0,516 **	0,333	0,005
problèmes de discrimination	0,201	0,332 **	0,520 ***	0,656 ***
problèmes de garde	0,366	0,055	0,165	0,287
problèmes pour trouver un emploi	0,236	-0,061	0,276	0,480
problèmes de transports	0,091	0,173	0,513 **	0,969 ***
peu d'offres d'emploi dans le domaine de compétences	0,438 ***	0,208 *	0,179	0,352 *
diplôme ou formation insuffisants	0,462 ***	0,373 **	-0,072	0,315
pas assez d'expérience professionnelle	0,173	0,162	0,738 ***	0,669 **
caractéristiques du contrat non satisfaisantes	-0,410	0,145	-0,621 *	-0,865 **
aucune aide institutionnelle (aide pour le CV, formation,...)	-0,263	0,087	0,515 *	0,325

<i>Situation lors de l'inscription à l'ANPE</i>				
immédiatement disponible et recherche d'un CDI à temps plein (ref)				
immédiatement disponible et recherche d'un CDI à temps partiel	-0,063	-0,404 ***	-0,461 **	-0,638 **
immédiatement disponible et recherche d'un CDD	-0,568 ***	-0,772 ***	-0,544 **	-0,296
non immédiatement disponible	1,419 ***	0,628 ***	0,606 ***	0,075
Permis de conduire	-0,072	0,527 ***	-0,181	0,928 ***
Accès internet	0,189 *	0,291 ***	0,076	0,301 *
Rapport aux études et littérature				
Sentiment d'avoir du arrêté trop tôt ses études	-0,065	0,104	0,127	0,147
Sentiment que les études ont aidé à réussir qqc dans la vie	0,081	0,069	0,043	0,308 **
Français (uniquement) parlé durant l'enfance	0,254 **	0,106	0,039	-0,184
Capacité à lire très facilement un journal français	-0,042	-0,223 *	-0,162	0,340
Capacité à écrire une lettre très facilement	-0,013	0,071	0,039	-0,260
Environnement familial et conditions de vie au T2-2005				
En couple	-0,043	0,186 *	0,073	0,203
<i>Nombre d'enfants à charge</i>				
aucun (ref)				
un enfant	-0,056	0,117	0,201	0,279
2 enfants	-0,030	-0,043	0,311 *	0,305
3 enfants et plus	0,032	-0,128	0,156	-0,137
Présence de personnes dépendantes à charge	-0,642 **	0,100	-0,589 *	-0,716
Propriétaire de son logement	-0,030	-0,055	0,153	0,142
Perception d'un minimum social dans la famille	-0,347 **	-0,258 *	0,454 ***	0,356 **
Paiement de l'impôt sur le revenu	-0,294 ***	-0,235 ***	-0,589 ***	-0,384 **
À l'aise financièrement	-0,020	-0,137	-0,301 **	-0,161
Présence de personnes sur qui compter	-0,135	-0,025	-0,079	-0,150
Bénéficiaire de la CMU	0,312 **	0,030	0,374 ***	-0,346 **
Données de contexte macroéconomique				
taux d'accès régional moyen aux contrats aidés sur 2005-2007	81,309 ***	182,669 ***	282,937 ***	673,069 ***
taux de chômage départemental au T2-2005	-0,007	0,002	-0,139 ***	0,008
Constante	-5,976 ***	-5,428 ***	-4,829 ***	-7,681 ***
Nombre d'obs.	4400	4280	3769	2580
Log-vraisemblance	-676,5	-411,1	-823,2	-125,8

Seuils de significativité : * 10% ; ** 5% ; *** 1%

Source : Panel 2008, Dares

Tableau 6 – Estimation de la probabilité d'entrer en contrat aidé, régressions logistiques

		CAE	CIE	CAV	CI-RMA
Sexe	non apparié	-0.165*	0.037	-0.107*	0.013
	apparié	0.008	-0.006	0.000	0.012
Formation					
< CAP-BEP	non apparié	-0.059*	0.031*	-0.031	0.066*
	apparié	-0.016	0.011	-0.001	0.002
CAP-BEP	non apparié	-0.039	-0,036	-0.013	-0.014
	apparié	0.023	0.005	-0.004	-0.002
niveau Bac	non apparié	-0.011*	-0.016	-0.002	-0.041
	apparié	-0.005	-0.003	0.006	0.007
supérieur au Bac	non apparié	0.109*	0.020	0.046*	-0.011
	apparié	-0.002	-0.013	-0.001	-0.008
Parcours professionnel avant inscription					
toujours en emploi	non apparié	0.193*	-0.004	0.146*	-0.047
	apparié	0.006	-0.007	0.006	0.005
alternance emploi, chômage inactivité	non apparié	-0.115*	-0,047*	-0.078*	0.007
	apparié	-0.008	-0,010	-0.005	-0.006
toujours ou principalement été au chômage	non apparié	-0.059*	-0.002	-0.042*	-0.002
	apparié	-0.001	-0.001	0.000	-0.002
toujours ou principalement été inactif	non apparié	-0.019	0,053*	-0.026*	0.042*
	apparié	0.003	0.007	-0.001	0.004
Santé					
très bonne santé	non apparié	0.031	0.047*	0.022	-0.065*
	apparié	0.007	-0.005	-0.023	-0.023
assez bonne santé	non apparié	-0.008	-0.001	-0.054*	-0.029
	apparié	-0.020	0.011	0.001	0.008
assez mauvaise santé	non apparié	-0.013	-0.022	0.066*	0.060*
	apparié	0.000	-0.011	0.006	0.003
très mauvaise santé	non apparié	-0.008	-0.019	-0.000	0.021
	apparié	0.011	0.007	0.002	0.001
Objectif lors de l'inscription à l'ANPE					
travailler à tout prix	non apparié	0.045*	-0.002	0.031	-0.010
	apparié	-0.005	-0.002	0.001	0.009
travailler dans un domaine précis	non apparié	-0,141*	-0,118*	-0.128*	-0.117*
	apparié	-0.007	-0.007	-0.007	-0.025
autres	non apparié	0.095*	0,120*	0.097*	0.128*
	apparié	0.002	0.009	0.006	0.016
Motif d'inscription à l'ANPE					
licenciement	non apparié	0,131*	-0.050*	0.096*	-0.050*
	apparié	0.006	-0.005	0.010	-0.009
démission	non apparié	0,024*	0.008	0.009	-0.001
	apparié	0.001	-0.000	-0.001	-0.001
fin de contrat ou de mission d'intérim	non apparié	0,092*	0,106*	0.020	0.029
	apparié	-0.001	0.004	-0.003	0.002
1 ^{ère} entrée ou reprise après interruption <6 mois	non apparié	-0,247*	-0,064*	-0.125*	0.022
	apparié	-0.006	0.001	-0.006	0.007
Durée cumulée au chômage de 1997 au T2-2005					
moins d'un mois	non apparié	0,115	0.080	0.039*	0.031
	apparié	0,005	0.004	-0.001	-0.002
de 1 à 6 mois	non apparié	0,075*	0,064*	0.039*	0.039
	apparié	0,007	0.003	0.003	0.000
de 6 mois à un an	non apparié	-0,014	-0,036*	0.017	-0.028
	apparié	0,003	-0.008	0.000	0.008
d'un an à deux ans	non apparié	-0,046*	-0,074*	0.003	-0.066*
	apparié	-0.001	-0.003	0.003	-0.001
plus de deux ans	non apparié	0,147*	-0.034*	-0.098*	0.024
	apparié	-0.014	0.004	-0.006	-0.006

Seuil de significativité : * 5%

Source : Panel 2008, Dares

Notes : Sur l'échantillon apparié, les moyennes des caractéristiques sont calculées à partir des valeurs contre-factuelles obtenues en appliquant le poids de *matching* (fonction noyau) aux caractéristiques des témoins. L'hypothèse nulle d'égalité des moyennes entre témoins et bénéficiaires sur l'échantillon apparié n'est jamais rejetée au seuil de 5%.

Tableau 7 – Test d'égalité des moyennes entre populations de bénéficiaires et de témoins

	CIE	CI-RMA	CAE	CAV	TM1	TM2
Population	22 337	1 723	44 668	13 157	1 013 465	175 796
Population avec les hors champ	22 668	1 736	45 195	13 331	1 044 417	187 051
Echantillon tiré	4 148	1 723	4 151	4 139	5 807	4 972
Vague 1						
Nb de retours en vague 1	2 308	915	2 442	2 385	2 991	2 397
Nb de répondants en vague 1	2 278	908	2 413	2 363	2 910	2 244
Variables expliquant la non-réponse en V1	État matrimonial Nationalité Formation A perçu le RMI Région Période d'entrée Sexe Age	A perçu l'ASS Formation Sexe Période d'entrée Nationalité Etat matrimonial	Sexe Age A perçu le RMI Formation Nationalité Etat matrimonial Période d'entrée Nb d'enfants	Sexe ASS en CAV Formation Nationalité Durée du CAV RMI en 2005 Région	A perçu le RMI Nationalité Etat matrimonial Sexe Formation Nb d'enfants Durée chômage	A perçu le RMI Sexe Région Formation Durée chômage Nb d'enfants Nb d'épisodes de chômage
Nb de GRH taux de réponse	16 [38 % ; 75 %]	7 [41% ; 65%]	16 [44% ; 75%]	16 [40% ; 77%]	15 [36% ;67%]	15 [30% ;65%]
Vague 2						
Nb de répondants à la vague 2	1888	746	2008	1935	2392	1834
Variables expliquant la non-réponse en vague 2	Permis en 2005 Emploi en V1 Etat matrimonial Raison du CA Age Internet Sexe	Utilité du CA A perçu le RMI Emploi en V1	Permis en 2005 Etat matrimonial Accompagnement pendant le CA Nationalité A perçu l'ASS Emploi en V1 Sexe	Emploi en V1 Permis en 2005 Internet Age Formation pendant le CA	Permis en 2005 Emploi en V1 A perçu le RMI Etat matrimonial Formation	Internet Secteur d'activité Permis en 2005 A perçu le RMI
Nb de GRH taux de réponse	13 [62% ; 92%]	5 [71% ; 91%]	11 [68% ; 93%]	12 [71% ; 96%]	12 [73% ;90%]	8 [76% ;88%]
Calage						
Variables de calage, identiques pour les deux vagues	Sexe, age, formation, RMI ANPE, ASS ANPE, RMI avant 2005, ASS avant 2005, RMI après 2005, ASS après 2005, durée du chômage (indemnisé et totale) avant 2005, catégorie et motif d'inscription à l'ANPE période d'entrée en contrat aidé, durée de la convention initiale de contrat aidé temps de travail en contrat aidé, type d'employeur en contrat aidé					
Poids de calage 1	[5,4 ; 17,1]	[1,2 ; 3,1]	[9,5 ; 35,9]	[3,4 ; 9,5]	[57 ; 1448]	[1,6 ; 1585]
Poids de calage 2	[5,6 ; 30,9]	[1 ; 5,6]	[9,2 ; 56,5]	[3,2 ; 16,1]	[65 ;1781]	[1,5 ; 1809]

Note : Le calage est réalisé par la méthode du *Raking ratio* grâce à la macro SAS CALMAR. Pour les deux populations de témoins, le rapport des poids finaux sur les poids initiaux est relativement élevé en raison du mode d'échantillonnage par *matching* qui rend plus difficile l'adéquation avec les marges sur la population totale.

Tableau 8 – Synthèse des variables intervenant pour le redressement des données

Références

- ABBRING, J. et VAN DEN BERG, G., « The Nonparametric Identification of Treatment Effects in Duration Models », 2003, *Econometrica*, Vol. 71, p. 1491-1517.
- ANGRIST, J. et IMBENS, G., « Identification and estimation of local average treatment effects », *Econometrica*, 1994, vol. 62, n°2, p. 467-475.
- ARDILLY, P., *Les techniques de sondage*, Dunod, 2006.
- BRILHAULT, G. et CARON, N., « La correction de la non-réponse totale : quelles méthodes ? », *La lettre du SSE n°58*, 2005, Dossier « Méthode et redressement de la non-réponse ».
- BRION, P., CARON, N. et PIÉTRI-BESSY, P., « Redresser la non-réponse totale dans les enquêtes auprès des entreprises : les pièges à éviter. Illustration avec l'enquête innovation », *Actes des Journées de Méthodologie Statistique*, 2005.
- BRODATY, T., CRÉPON, B. et FOUGÈRE, D., « Les méthodes micro-économétriques d'évaluation et leurs applications aux politiques actives de l'emploi », *Économie et Prévision*, 2007, Vol. 177, p. 91-118.
- BRYSON, A., DORSETT, R. et PURDON, S., « The Use of Propensity score matching in the Evaluation of Active Labour Market Policies », 2002, Working Paper.
- CALIENDO, M. et KOPEINIG S., « Some practical guidance for the implementation of propensity score matching ». *Journal of Economic Surveys*, 2008, Vol. 22, No. 1, p. 31-72
- CARLING, K. et RICHARDSON, K., « The relative efficiency of labour market programs : Swedish experience from the 1990s », *Labour Economics*, 2004, Vol. 11, p. 335-354.
- CARON, N., « La correction de la non-réponse par repondération et par imputation », Document de travail de l'INSEE, 2005, Paris.
- CRÉPON, B., FERRACCI M., JOLIVET G., et VAN DEN BERG G., « Active labor market policy effects in a dynamic setting », *Journal of the European Economic Association*, 2009, Vol. 7, No. 2-3, p. 595-605.
- CRÉPON, B. et KAMIONKA, T., « Expertise des fichiers du Panel 2000 des bénéficiaires des politiques de l'emploi de la DARES et orientations pour le futur », *Mimeo*, Dares, 2006.
- DAVEZIES, L. et D'HAULTFŒUILLE, X., « Faut-il pondérer ? ... Ou l'éternelle question de l'économètre confronté à un problème de sondage », Document de travail de l'INSEE, 2009.
- DEHEJIA, R. et WAHBA, S., « Causal Effects in Nonexperimental Studies : Reevaluating the Evaluation of Training Programs », *Journal of the American Statistical Association*, 1999, Vol. 94 (448), p. 1053-1062.
- EVEN, K. et KLEIN, T., « Les contrats et stages aidés : un profit à moyen terme pour les participants ? Les exemples du CIE, du CES et du Sife », *Économie et Statistique*, 2008, n°408-409.
- FREDRIKSSON, P. et JOHANSSON, P., « Dynamic Treatment Assignment : The Consequences for Evaluations Using Observational Data », *Journal of Business and Economic Statistics*, 2008, Vol. 26, No. 4.

- FENDRICH, Y. et RÉMY, V., « L'insertion professionnelle, six mois après la sortie de contrats aidés », *Dares Analyse*, 2009, Vol. 45, No. 1.
- FRÖLICH, M., « Propensity score matching without conditional independence assumption – with an application to the gender wage gap in the United Kingdom », *Econometrics Journal*, 2007, Vol. 10, No. 2.
- GIVORD, P., « Méthodes économétriques pour l'évaluation de politiques publiques », *Document de travail de la Direction des Études et Synthèses Économiques*, INSEE, 2010.
- GERFIN, M. et LECHNER, M., « Microeconomic evaluation of the active labour market policy in Switzerland », *Economic Journal*, 2002, vol 112, p. 854-93.
- GREEN, H., CONNOLLY, H., MARSH, A. et BRYSON, A. « The Medium-Term Effects of Voluntary Participation in ONE », Département for Work and Pension, 2001, Research Report No. 149.
- HECKMAN, J., ICHIMURA, H. et TODD, P., « Matching as an econometric evaluation estimator : Evidence from evaluating a job training programme ». *Review of Economic Studies*, 1997, Vol. 64, p. 605-54.
- HIRANO, K., IMBENS, G. et RIDDER, G., « Efficient estimation of average treatment effects using the estimated propensity score ». *Econometrica*, 2003, Vol. 71, p. 116-89.
- HUJER, R., CALIENDO, M. et THOMSEN, S., « New evidence on the effects of job creation schemes in Germany – a matching approach with threefold heterogeneity », *Research in Economics*, 2004, Vol. 58, No. 4.
- LALIVE, R., VAN OURS, J. C. et ZWEIMÜLLER, J., « The impact of active labour market programmes on the duration of unemployment in Switzerland », *The Economic Journal*, 2008, Vol. 118.
- LECHNER, M., « Earnings and employment effects of continuous off-the-job training in East Germany after unification », *Journal of Business and Economic Statistics*, 1999, vol. 17, p. 74-90.
- LECHNER, M. et WUNSCH, C., « Active Labour Market Policy in East Germany : Waiting for the Economy to Take Off », 2007, *Kiel Working Paper Collection*, No. 1.
- ROSENBAUM, P. et RUBIN, D., « The central role of the propensity score in observational studies for causal effects », *Biometrika*, 1983, Vol. 70, p. 41-55.
- SIANESI, B., « An evaluation of the Swedish system of active labour market programmes in the 1990s », *The Review of Economics and Statistics*, 2004, Vol. 86, No. 1, p. 133-155.
- SILVERMAN, B., *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, 1986.
- SMITH, J. et TODD, P., « Does matching overcome LaLonde's critique of nonexperimental estimators ? », *Journal of Econometrics*, 2005, Vol. 125, p. 305-53.