

Evaluation d'impact de l'accompagnement des demandeurs d'emploi par les Opérateurs Privés de Placement et le programme Cap Vers l'Entreprise

Rapport Intermédiaire *

Luc Behaghel[†], Bruno Crépon[‡], Julien Guitard[§] et Marc Gurgand[¶]

30 juin 2008

*Nous remercions l'ensemble des partenaires de cette évaluation : l'ANPE, l'Unédic et la DARES, et Claude Seibel qui en a présidé le comité de pilotage. Nous remercions également Lucie Gadenne pour son excellent travail d'assistant de recherche.

[†]Ecole d'Economie de Paris (Inra), Crest et J-PAL

[‡]Crest et J-PAL

[§]Crest

[¶]Ecole d'Economie de Paris, Crest et J-PAL

Résumé

Ce rapport présente les premiers résultats de l'évaluation d'impact de l'effet de l'accompagnement renforcé assuré par des opérateurs privés de placement (OPP), d'une part, et par la prestation Cap Vers l'Entreprise (CVE), d'autre part. Ces deux prestations ont été fournies, par l'Unédic et par l'ANPE respectivement, aux demandeurs d'emploi présentant des risques de chômage de longue durée. Similaires par leur contenu, ces deux prestations ont été proposées à des demandeurs d'emploi récemment inscrits à l'ANPE et présentant des droits d'indemnisation d'au moins 365 jours ("flux indemnisable"), parfois dans les mêmes régions (en particulier Ile-de-France, Rhône-Alpes, Nord-Pas de Calais), parfois dans des régions distinctes. La prestation CVE a également été proposée à des demandeurs d'emploi dont les droits étaient plus courts ("flux non indemnisable") ou la demande plus ancienne ("stock"). Ces prestations ont pour but de concentrer des moyens humains importants sur l'accompagnement de ce public à risque de chômage de longue durée. L'ANPE comme l'Unédic ont mobilisé les moyens suffisants pour assurer chacun l'accompagnement de 40000 demandeurs d'emploi.

Cette évaluation d'impact a été réalisée en utilisant une méthode inédite en France jusque-là. Cette méthode consiste à ventiler au hasard les bénéficiaires éligibles potentiels en plusieurs groupes, l'un se voyant proposé l'accompagnement classique de l'ANPE, l'autre l'accompagnement dans le cadre du programme CVE et le troisième l'accompagnement OPP proposé par l'Unédic. L'intérêt de cette méthode est de partir de trois populations statistiquement identiques. De ce fait, les différences observées dans le devenir de ces cohortes ne peuvent que provenir de la participation à des programmes d'accompagnement différents. Cette méthode permet donc de mesurer sans ambiguïté la plus-value apportée par les deux prestations d'accompagnement renforcé par rapport à l'accompagnement classique. Nous nous intéressons pour l'instant dans le cadre de ce rapport à l'impact sur les sorties vers l'emploi à 3 mois, 6 mois et 9 mois. Ultérieurement nous examinerons l'impact sur une gamme de variables de résultats plus large, incluant par exemple les versements d'indemnités de sorte qu'il soit possible de procéder à de réelles analyses coût-bénéfices.

L'évaluation dont les premiers résultats sont présentés ici résulte de la coordination d'efforts importants de la part de multiples intervenants, en particulier de l'ANPE qui a développé au sein des 393 Agences Locales de l'Emploi impliquées dans cette opération, un outil

permettant d'assurer et de pérenniser la ventilation des bénéficiaires potentiels. Au cours de l'année 2007 plus de 200000 demandeurs d'emploi ont été orientés aléatoirement vers CVE et OPP (groupes de traitement) ou vers l'accompagnement classique de l'ANPE (groupe de contrôle).

Cette orientation aléatoire a effectivement permis de constituer des groupes de traitement et un groupe de contrôle aux caractéristiques observables statistiquement identiques. Par exemple, la proportion de femmes affectées est la même dans le groupe bénéficiant de l'accompagnement classique que dans chacun des deux autres groupes, bénéficiant de l'accompagnement CVE et de l'accompagnement OPP. On vérifie qu'il en est de même pour tout un ensemble de variables observables, telles que le diplôme, la qualification, l'âge, etc. Cela atteste que la ventilation aléatoire en trois groupes a bien fonctionné et que s'il en est ainsi pour des caractéristiques observables, il en est de même pour toute autre caractéristique inobservable, en particulier toutes celles ayant un impact sur le retour à l'emploi : les comparaisons que nous effectuons relèvent bien des seules différences de participation aux programmes.

Les demandeurs d'emploi auxquels étaient proposés les programmes CVE ou Unédic étaient libres d'y entrer ou non. La proportion de demandeurs d'emploi entrant dans le dispositif d'accompagnement renforcé vers lequel ils ont été orientés varie, selon les groupes considérés, entre un peu plus de 30% et un peu moins de 50%. Nous constatons que le taux d'entrée dans le programme CVE comme dans le programme OPP varie d'un groupe social à l'autre. Il est plus élevé pour des demandeurs plus jeunes et plus diplômés ; il s'agit plus souvent de femmes et de techniciens ou de professions intermédiaires. Cette observation souligne bien que les populations entrant dans les dispositifs sont particulières, ce qui rend difficile de faire des évaluations d'impact si on ne met pas en place un cadre de ventilation aléatoire tel que celui qui a été développé ici.

On présente ensuite de premiers résultats amenés à être complétés lorsque de nouvelles données seront disponibles. Dans ce résumé, on se focalise sur le taux de sortie à 6 mois vers l'emploi (horizon raisonnable pour que l'accompagnement produise des effets, et pour lequel on dispose d'échantillons suffisants). L'effet du programme CVE pour les demandeurs d'emploi du flux indemnisable est positif, quantitativement important et statistiquement significatif. Toutes zones confon-

dues, on estime une augmentation du taux de sortie vers l'emploi qui se situe dans une fourchette de 3,7 à 13,1 points de pourcentage (avec un degré de confiance de 95%). Dans le flux indemnisable également, mais pour des zones en partie différentes, la plus-value de l'accompagnement OPP se situe entre -1,8 et 7,9 points de pourcentage. Il n'est donc statistiquement pas possible de conclure, à ce stade, si l'impact est positif. Les effets du programme CVE sur des demandeurs d'emploi présentant des droits d'indemnisation plus courts ou dont la demande est plus ancienne sont également favorables, mais plus modérés : entre 1,5 et 10,6 points de pourcentage dans le flux indemnisable, et entre 1 et 8,4 points de pourcentage dans le stock. Comme l'expriment les fourchettes encore larges, ces résultats présentent une marge d'incertitude non négligeable, amenée à se réduire au fur et à mesure que les données s'accumulent.

Les demandeurs qui sont effectivement entrés dans un accompagnement, après y avoir été orientés, sont *a priori* différents de ceux qui l'ont refusé ou, pour d'autres raisons, n'ont pas été pris en charge. Ce processus de *sélection* peut en outre être différent pour l'entrée dans le programme OPP et dans le programme CVE. Or, il est possible de reconstituer ce qu'auraient été les taux de retour en emploi des bénéficiaires de l'accompagnement s'ils n'avaient pas bénéficié de cet accompagnement (situation contrefactuelle). En les comparant au taux de retour en emploi des non-bénéficiaires, on met en évidence ce processus de sélection.

Ainsi, dans le flux indemnisable, les demandeurs d'emploi qui entrent en accompagnement CVE ou OPP auraient eu, *en l'absence d'accompagnement*, des chances de sortie à 6 mois vers l'emploi sensiblement plus faibles que ceux qui ont été orientés mais ne sont pas entrés dans les dispositifs. Ainsi, dans les zones où CVE et OPP ont été développés conjointement, le taux de sortie des personnes orientées vers CVE et qui ont été accompagnées aurait été de 17%, tandis que celui des personnes orientées vers CVE mais qui n'ont finalement pas été accompagnées a été de 25,5%. *Ainsi, le processus de sélection à conduit à la prise en charge de personnes relativement plus démunies.* La sélection dans le programme OPP est de même nature, mais avec un contraste moins fort, donc un processus de sélection moins marqué : les personnes accompagnées auraient eu un taux de sortie de 21,5%, contre 24,6% pour les personnes non accompagnées.

L'accompagnement renforcé a néanmoins permis à ces demandeurs de "rattraper" leur retard. Ceci est particulièrement net pour le pro-

gramme CVE, puisqu'il a permis de faire passer leur taux de sortie d'un potentiel de 17% à un taux effectif de 26,2%. Dans les mêmes zones, le programme OPP n'a quant à lui fait gagner que 1,2 points en moyenne (de 21,5% à 22,7%). Mais il faut souligner que le taux d'emploi potentiel de la population prise en charge par les OPP partait de moins bas.

Dans les zones où les deux programmes ont été développés indépendamment, on observe des processus de sélection similaires pour les deux programmes. Les entrants dans les programmes ont des potentiels de sortie vers l'emploi plus faible que les non entrants, et les programmes leur permettent de rejoindre le niveau de ces non entrants. *A contrario*, les demandeurs d'emploi du flux non indemnisable et du stock pris en charge par le programme CVE avaient des chances de sortie vers l'emploi en l'absence d'accompagnement similaires à celles des demandeurs orientés mais n'entrant pas dans CVE (31% dans le flux non indemnisable, 21,9% dans le stock). La plus-value de l'accompagnement CVE leur permet alors d'obtenir des taux de sortie vers l'emploi supérieurs (37,1% dans le flux non indemnisable, 26,6% dans le stock). L'ensemble de ces estimations contient une forte incertitude statistique; elles suggèrent néanmoins des effets importants, notamment pour le programme CVE (augmentation de 20% à 25% des chances de sortie).

Enfin, des résultats désagrégés exploratoires suggèrent une forte hétérogénéité de l'impact des accompagnements renforcés selon la date d'entrée dans le dispositif et selon les caractéristiques sociodémographiques du demandeur. Cependant, la précision statistique n'est pas suffisante pour conclure à ce stade.

Le rapport met aussi l'accent sur les aspects méthodologiques associés à la production d'évaluations d'impact rigoureuses. Deux points sont particulièrement mis en avant : la qualité des données utilisées et la détermination des estimateurs utilisés ainsi que leur interprétation.

L'évaluation porte sur la sortie des listes de l'ANPE vers l'emploi à différents horizons (3, 6 et 9 mois). Cette sortie vers l'emploi est mal renseignée par les données administratives. Une enquête complémentaire téléphonique a permis cependant de limiter le pourcentage de situations inconnues à environ 10% (pour un horizon de sortie de 6 mois). Ce taux ne varie que peu, et pas de façon systématique, entre les groupes de traitement et de contrôle.

Une attention particulière est portée à l'interprétation des estimateurs utilisés, sachant que tous les demandeurs d'emploi n'entrent pas

dans l'accompagnement prévu et que les effets de l'accompagnement peuvent différer fortement d'un individu à l'autre. On montre comment il est possible d'estimer un effet moyen sur le groupe dit des *compliers*, c'est-à-dire ceux qui ont effectivement bénéficié d'un accompagnement renforcé parce qu'ils ont été orientés aléatoirement vers cet accompagnement. L'annexe méthodologique détaille sous quelles conditions cet estimateur s'étend au cas des zones où coexistaient les deux programmes, en dépit des difficultés rencontrées par le protocole dans ces zones (demandeurs d'emploi du groupe de contrôle entrant dans l'accompagnement OPP).

Ce rapport intermédiaire a pour objet l'évaluation de l'effet de l'accompagnement renforcé assuré par des opérateurs privés de placement (OPP), d'une part, et par la prestation Cap vers l'entreprise (CVE), d'autre part. Ces deux prestations ont été fournies, par l'Unédic et par l'ANPE respectivement, aux demandeurs d'emploi présentant des risques de chômage de longue durée. Les OPP sont proposés uniquement aux demandeurs d'emploi qui ont un droit ouvert pour au moins 365 jours, tandis que CVE est accessible à tous les demandeurs appartenant au public cible. Les prestations consistent en un suivi hebdomadaire personnalisé par un conseiller unique, chargé d'un nombre limité de demandeurs. Elles durent 6 mois (prolongeables dans certains cas) et visent la reprise d'un emploi stable. L'accompagnement se prolonge par un suivi dans l'emploi. Les OPP sont rémunérés forfaitairement à la prise en charge, puis à la mise en emploi, puis après 6 mois de maintien dans l'emploi.

Les deux programmes ont été déployés simultanément dans les régions Ile-de-France, Lorraine, Midi-Pyrénées, Nord-Pas de Calais et Rhône-Alpes. En Alsace, dans le Centre, en Haute-Normandie et en Provence-Alpes-Côte d'Azur, seuls les OPP sont intervenus. Enfin, seul le dispositif CVE a été mis en place en Aquitaine. Au total, chacun des deux programmes visait environ 40 000 bénéficiaires sur l'année 2007.

Plusieurs études françaises et étrangères indiquent que ce type d'intervention favorise le plus souvent l'accès à l'emploi des bénéficiaires, mais elles sont rarement menées dans un cadre méthodologique parfaitement adapté, parce qu'il est très rare que les bénéficiaires soient rigoureusement comparables aux non-bénéficiaires qui fournissent une situation de référence. Le protocole que nous avons conçu avec l'ANPE et l'Unédic et qui a été mis en oeuvre durant toute l'année 2007 et le premier trimestre 2008 permet de corriger très rigoureusement les biais de sélection. Il repose en effet sur un dispositif de tirage au sort qui constitue des groupes d'individus parfaitement semblables, ce qui permet d'isoler avec certitude l'impact des accompagnements. Il s'agit d'une des toutes premières opérations de cette nature menées en France, et elle permet d'avoir des résultats robustes et convaincants.

Ce rapport présente de premières estimations de la plus-value apportée par l'accompagnement renforcé des demandeurs d'emploi dans les dispositifs CVE et OPP. L'évaluation porte ici sur la sortie vers l'emploi : dans quelle mesure le fait de bénéficier d'un accompagnement renforcé accroît-il les chances qu'un demandeur d'emploi sorte des listes de l'ANPE pour re-

prendre un emploi ? On donne de premiers résultats à différents horizons – 3, 6 et 9 mois – car les effets de l’accompagnement peuvent mettre du temps à se manifester. Ces résultats sont provisoires car il est encore trop tôt, dix-huit mois après le début de l’expérimentation, pour que tous les effets des dispositifs aient pu se produire et être mesurés. Il faudra encore attendre un an afin que toutes les cohortes de demandeurs d’emploi entrées dans l’expérience puissent être suivies pendant douze mois. Les résultats porteront alors sur des échantillons plus importants, ce qui les rendra plus précis. Ils couvriront les effets de l’accompagnement renforcé sur douze mois et intégreront la question de la récurrence au chômage : dans quelle mesure le fait de bénéficier d’un accompagnement renforcé réduit-il les chances qu’un demandeur d’emploi retourne au chômage après sa sortie vers l’emploi ? Même s’ils sont provisoires et incomplets, ces résultats marquent une étape importante du processus d’évaluation. Ce rapport consacre une attention particulière à expliquer leur construction, et à présenter la marge d’incertitude statistique liée à toute estimation. Une telle attention est nécessaire pour que le protocole suivi, une fois compris de tous, favorise l’émergence d’un consensus sur les faits. A dessein, l’interprétation des résultats et *a fortiori* la recommandation de politiques publiques sont absents de ce rapport.

La section qui suit présente le protocole expérimental et sa mise en oeuvre. On détaille ensuite les données utilisées et les mesures d’impact utilisées. Les résultats sont présentés dans une dernière section. Les détails techniques sur les estimateurs statistiques sont rassemblés en annexe.

1 Le protocole expérimental

Le but de l’évaluation menée est de mesurer **l’impact** des dispositifs sur le retour à l’emploi des bénéficiaires. La principale question à laquelle nous cherchons à répondre est : *quelle est la plus-value apportée par chacun des deux dispositifs ?* La plus-value du programme est définie comme la différence entre la situation d’emploi des individus bénéficiant du dispositif avec ce qu’aurait été leur situation s’ils n’en avaient pas bénéficié. Cette définition présente une difficulté méthodologique certaine : elle fait intervenir la situation d’un bénéficiaire s’il n’avait pas bénéficié du dispositif, or cette situation n’est pas observée puisqu’elle n’a pas eu lieu. Toute les méthodes d’évaluation cherchant à mesurer l’impact d’un dispositif doivent reconstituer cette situation alternative non observée. En général ceci est effectué en

référence à un *groupe de contrôle*.

La constitution du groupe de contrôle approprié est difficile. Il est fréquent que l'on utilise par exemple les individus ayant refusé le dispositif, ou appartenant à des zones géographiques voisines. Cela peut conduire à des erreurs d'appréciation très fortes car les comparaisons effectuées entre le groupe des bénéficiaires et le groupe de contrôle mesurent simultanément deux choses qui sont indiscernables : l'effet du programme, qui est ce que l'on souhaite mesurer, mais aussi un effet *population*. Cet effet population provient du fait que décider de participer à un programme, ou être sélectionné comme un bénéficiaire potentiel, est sans doute lié aux caractéristiques ou situations qui déterminent la capacité d'un individu à retrouver un emploi. Dans ce cadre, une partie des différences dans les taux de retour à l'emploi entre groupe de bénéficiaires et groupe de contrôle correspond donc à l'existence de différences préexistantes au programme et qu'on ne saurait lui attribuer.

Pour que la situation du groupe de contrôle représente parfaitement la situation qu'auraient connu les bénéficiaires en l'absence du dispositif, il faut que l'un et l'autre soient tirés au sort. C'est la manière la plus simple et la plus transparente de constituer des groupes comparables, quelle que soit la diversité des situations individuelles dans l'ensemble du public concerné. Le protocole mis en place dans le cadre de cette expérimentation vise à constituer de tels groupes.

1.1 Principes du protocole

La méthode que nous avons utilisée consiste à tirer au hasard trois groupes dans la population des bénéficiaires potentiels. C'est ce qui a été fait grâce à l'Outil de constitution des cohortes (OCC), mis en place par l'ANPE. Il s'agit d'une application développée à l'automne 2007 par l'ANPE et mise en place sur les postes informatiques des conseillers. Il tire au sort la prestation qui doit être proposée au demandeur, avec des probabilités spécifiques à différentes situations et à chaque agence.

Le cas le plus général concerne les demandeurs d'emploi qui viennent s'inscrire à l'ANPE, dans des zones où à la fois CVE et OPP sont expérimentés. Les personnes sont reçues par un conseiller qui, dans le cadre de l'entretien dit PPAE (Projet personnalisé d'accès à l'emploi), détermine le parcours le plus approprié à la situation du demandeur. Les personnes orientées vers un parcours 3 (P3), celles qui présentent un risque important de chômage de longue durée, sont éligibles à la prestation CVE (sous réserve qu'ils ne

soient pas co-traités ¹et que leur projet professionnel ne demande pas à être redéfini); elles sont également éligibles aux OPP lorsqu'elles ont un droit à l'assurance chômage ouvert pour au moins 365 jours. Dans la suite, nous utiliserons *par commodité* le terme *indemnisable* pour cette population : ce terme exclut par convention les demandeurs indemnisables sur de courtes durées.

A l'occasion de cet entretien PPAE, une fois le diagnostic réalisé, l'OCC ventile aléatoirement les demandeurs indemnisables en trois groupes : ceux auquel on propose l'accompagnement par les OPP, ceux auxquels on propose l'accompagnement CVE et enfin ceux bénéficiant de l'accompagnement classique des demandeurs d'emplois en parcours P3. Pour les demandeurs relevant du parcours P3 mais non indemnisables (au sens large défini plus haut), ou pour tous les demandeurs P3 dans les zones où OPP n'est pas expérimenté, la ventilation se fait en deux groupes seulement : ceux auxquels on propose CVE et ceux auxquels on propose l'accompagnement classique. Enfin, dans les zones où CVE n'est pas expérimenté, les individus P3 indemnisables se voient proposer soit OPP, soit le parcours classique. Dans certains cas, par exemple lorsque le statut d'indemnisation n'est pas disponible à l'ANPE le jour du PPAE, le tirage au sort est réalisé non pas en situation d'entretien, mais sur des listes. Les bénéficiaires sont alors avertis par téléphone ou par courrier. Les personnes relevant du parcours P3 mais déjà inscrites à l'ANPE depuis plus de trois mois, sont également concernées par la prestation CVE (mais jamais par les OPP). Dans ce cas, les tirages sont également réalisés sur liste. Enfin, lorsqu'un demandeur d'emploi auquel on a proposé le programme CVE ou l'accompagnement OPP le refuse, il bénéficie alors systématiquement de l'accompagnement classique.

Pour récapituler, il existe trois grandes populations à l'intérieur de demandeurs orientés vers un parcours P3 :

- (1) Les personnes dont la demande est inférieure à trois mois (pour la plupart des nouvelles inscriptions) et qui sont indemnisables : on parlera du *flux indemnisable* ;
- (2) Les personnes dont la demande est inférieure à trois mois (pour la plupart des nouvelles inscriptions) et qui ne sont pas indemnisables : on parlera du *flux non indemnisable* ;
- (3) Les personnes dont la demande est antérieure à trois mois : on parlera

¹C'est-à-dire ce qui sont accompagnés conjointement par d'autres organismes, comme l'Association pour l'emploi des cadres, l'Apec

du *stock*.

Les évaluations seront toujours menées séparément pour ces trois catégories de populations, car elles présentent des caractéristiques très différentes. Gardons à l'esprit que les OPP interviennent uniquement auprès du flux indemnisable.

Les probabilités d'envoi ont été fixées ALE par ALE et séparément pour chacune des 3 catégories de population. Ces probabilités ont été calculées en tenant compte du potentiel de l'ALE et des objectifs d'envoi vers les OPP et CVE, et elles ont été ajustées mois par mois en fonction de l'évolution de la situation. Au bout de quelques mois, les probabilités d'envoi vers les OPP ont été fortement rehaussées pour faire face aux engagements d'envois pris par l'Unédic. Parallèlement, des notes de cadrage opérationnelles émises par les deux directions générales ont ouvert la possibilité d'élargir le recrutement vers les OPP. Dans ce cas, l'orientation s'est faite hors OCC. Pour le flux indemnisable et dans les zones où intervenaient les OPP, les probabilités d'envoi vers les OPP étaient élevées, ce qui limitait fortement les effectifs qui pouvaient être orientés vers CVE ou le parcours classique. Cela implique que dans ces zones, la précision statistique tend à être plus faible.

Ainsi, la probabilité de tirage vers OPP était de 0.85 dans 90% des ALE ; la probabilité de tirage vers CVE était alors de 0.06 ou 0.08, le complément étant pour le parcours classique. Dans les zones où les OPP n'étaient pas présents, et sur le flux non indemnisable et sur le stock, la probabilité la plus fréquente était 0.5 vers CVE et 0.5 vers classique, mais dans de nombreux cas, notamment dans le stock, la probabilité d'envoi vers CVE était inférieure, en raison d'un déséquilibre entre le potentiel et les capacités d'accueil de CVE.

Au total, la mise en oeuvre de ce processus de ventilation aléatoire en différentes sous-populations permet, lorsque les effectifs sont suffisamment nombreux, de disposer de sous-populations statistiquement identiques, dans le sens où n'importe quelle caractéristique sera identiquement répartie dans chacune d'entre elles. Il y aura ainsi la même proportion de femmes, la même proportion de demandeurs d'emploi titulaires d'un diplôme donné, mais aussi la même proportion de demandeurs d'emploi fortement motivés pour retrouver au plus vite un emploi, dans chacune des sous-populations déterminées au hasard. Il est à noter que seul le recours à un outil comme OCC peut garantir que les demandeurs d'emplois orientés vers CVE, vers les OPP ou vers le parcours classique soient parfaitement identiques sur le plan statistique. Les différences de trajectoire observées ultérieurement sont donc nécessairement dues au fait que les programmes d'accompagnement qui ont été proposés

sont différents. Ainsi les différence entre la cohorte des demandeurs d'emplois du flux indemnisable auxquels on a proposé le programme OPP et celle des demandeurs du flux indemnisable auxquels on a proposé le programme d'accompagnement classique ne peut provenir que de la différence d'effets entre l'accompagnement classique et l'accompagnement OPP.

1.2 La montée en charge du dispositif

L'OCC a été développé dans 393 agences dans 16 régions au total. Parmi elles, 67 n'ont prescrit que des accompagnements classiques ou CVE, 91 n'ont prescrit que des accompagnements classiques ou OPP et les 235 restantes ont disposé des trois formules d'accompagnement. La montée en charge de l'OCC a été très rapide. Comme l'indique le tableau 1, dès le premier trimestre 2007, l'outil a orienté plus de 60 000 demandeurs d'emploi. Mais il a été d'abord massivement utilisé dans le stock et en requêtes par listes, tandis que la montée en charge dans le flux, indemnisable et non indemnisable, a été progressive. A fin décembre, 219 208 demandeurs d'emploi sont passé par OCC, dont 84 666 dans le stock, 60 269 dans le flux non indemnisable et 74 214 dans le flux indemnisable. Au total, 50 563 on été orientés vers les OPP, 67 067 vers CVE et 101 519 vers l'accompagnement classique.

Mois d'entrée dans OCC	Nombre de DE entrés par trimestre																				
	FI			FNI			Stock			Inconnue			Total			Classique			Destination		
	FI	FNI	Stock	FI	FNI	Stock	FI	FNI	Stock	FI	FNI	Stock	FI	FNI	Stock	FI	FNI	Stock	CVE	OPP	Inconnue
Trimestre 1	11740	11891	36888	11740	11891	36888	2			2			60521	39357	13365	7797	2				
Trimestre 2	24528	17542	24156	24528	17542	24156	7			7			66233	26857	22113	17256	7				
Trimestre 3	20767	15341	13700	20767	15341	13700	9			9			49817	19451	16403	13954	9				
Trimestre 4	17179	15495	9922	17179	15495	9922	41			41			42637	15854	15186	11556	41				
Total	74214	60269	84666	74214	60269	84666	59			59			219208	101519	67067	50563	59				

TAB. 1 – Montée en charge de OCC

Il est évident qu'un grand nombre de demandeurs éligibles sont entrés dans les dispositifs sans être orientés par l'OCC, surtout en début d'année et dans le flux indemnisable, en raison de la nécessité d'alimenter rapidement les OPP. Ces personnes sont hors expérimentation et cela constitue une perte d'effectif, donc de précision statistique. Mais l'évaluation est robuste à cette circonstance : parmi les individus qui ont été orientés avec l'OCC, et qui sont donc dans le périmètre de l'évaluation, la distribution entre les différents groupes (classique, CVE ou OPP) est parfaitement homogène. Les trois premières colonnes du tableau 2 en font la démonstration². La première colonne présente d'abord la moyenne d'un ensemble de caractéristiques individuelles décrites dans le Fichier historique de l'ANPE, sur les individus orientés vers l'accompagnement classique, en distinguant les zones OPP dans le flux indemnisé, les zones CVE dans le flux indemnisé, le flux non-indemnisé et le stock. La colonne suivante indique les moyennes des mêmes caractéristiques dans les populations orientées OPP ou CVE selon les cas. Les populations de ces deux premières colonnes ayant été déterminées au hasard, on vérifie que leurs caractéristiques sont semblables. La troisième colonne contient une étoile lorsque les différences sont statistiquement significatives, ce qui peut arriver en raison des aléas d'échantillonnage, mais de façon exceptionnelle (en moyenne pas plus d'une fois sur 20). On peut donc en conclure que l'évaluation qui suit compare des demandeurs parfaitement comparables. La comparaison des populations du flux indemnisable orientées OPP ou CVE dans les zones où les deux prestations coexistent ferait apparaître la même ressemblance.

1.3 Orientation et prise en charge effective

L'OCC ne définit qu'une orientation : la prise en charge est proposée au demandeur, qui est libre de la refuser. Les demandeurs ont également pu trouver un emploi avant d'entrer en accompagnement. Par ailleurs, il semble que, pour des raisons mal comprises, les listes de personnes à contacter issues de certaines requêtes OCC n'ont pas toujours été parfaitement transmises aux plates-formes OPP ou CVE. Dans certains cas encore, les demandeurs dont l'identifiant a été saisi dans l'OCC pour réaliser le tirage au sort se sont révélés inéligibles, par exemple parce que leur droit à indemnisation n'atteignait pas 365 jours ou parce qu'ils ne relevaient pas du parcours P3. Dans tous les cas, le

²Les autres colonnes de ce tableau seront analysées plus loin.

Flux indemnisable, zones OPP	Orienté			Entrant	
	Classique	OPP	(1) - (2)	OPP	(2) - (3)
	(1)	(2)		(3)	
Diplôme supérieur	0.24	0.25		0.25	
Bac	0.18	0.18		0.19	* ^a
CAP/BEP, autres secondaires	0.39	0.39		0.40	*
BEPC ou non-diplômés	0.18	0.19		0.17	*
âge	36.86	36.77		37.16	*
femme	0.51	0.50		0.51	
Ouvrier	0.16	0.15	*	0.15	
Employé	0.61	0.62		0.61	*
Technicien	0.10	0.10		0.11	*
Cadre	0.12	0.12		0.12	
Flux indemnisable, zones CVE	Orienté			Entrant	
	Classique	CVE	(1) - (2)	CVE	(2) - (3)
	(1)	(2)		(3)	
Diplôme supérieur	0.24	0.24		0.25	
Bac	0.18	0.18		0.20	*
CAP/BEP, autres secondaires	0.41	0.42		0.42	
BEPC ou non-diplômés	0.16	0.16		0.13	*
âge	36.11	35.96		36.16	
femme	0.53	0.53		0.55	*
Ouvrier	0.18	0.18		0.18	
Employé	0.60	0.61		0.60	
Technicien	0.11	0.12		0.14	*
Cadre	0.10	0.09	*	0.09	
Flux non indemnisable	Orienté			Entrant	
	Classique	CVE	(1) - (2)	CVE	(2) - (3)
	(1)	(2)		(3)	
Diplôme supérieur	0.26	0.26		0.30	*
Bac	0.17	0.17		0.18	*
CAP/BEP, autres secondaires	0.37	0.37		0.36	
BEPC ou non-diplômés	0.20	0.20		0.15	*
âge	33.44	33.49		32.88	*
femme	0.55	0.55		0.58	*
Ouvrier	0.17	0.17		0.16	*
Employé	0.67	0.67		0.65	*
Technicien	0.09	0.09		0.11	*
Cadre	0.07	0.07		0.08	
Stock	Orienté			Entrant	
	Classique	CVE	(1) - (2)	CVE	(2) - (3)
	(1)	(2)		(3)	
Diplôme supérieur	0.21	0.22		0.24	*
Bac	0.16	0.17	*	0.18	*
CAP/BEP, autres secondaires	0.40	0.40	*	0.40	
BEPC ou non-diplômés	0.22	0.22		0.18	*
âge	37.84	37.72		37.46	*
femme	0.54	0.53		0.53	
Ouvrier	0.17	0.17		0.17	
Employé	0.65	0.65		0.63	*
Technicien	0.10	0.10		0.12	*
Cadre	0.08	0.08		0.08	

^aUne * indique une différence significative au seuil de 5%

TAB. 2 – Comparaison des populations suivant leur assignation et leur entrée dans les programmes

OPP	Effectif assigné 50 563	Signataire % 47	%Non signalé 33	Signataire inc. 20	en CVE 0.3
CVE	Effectif Assigné	% ayant accepté	% refus ou inc.		% OPP
Flux indemnisable	9 938	42.8	56.3		0.9
Flux non indemnisable	26 399	34.0	65.9		0.2
Stock	30 730	35.5	64.2		0.3
Classique	Effectif Assigné	Classique	% CVE		OPP
Flux indemnisable	13 713	96.3	0.5		3.2
Flux non indemnisable	33870	99.2	0.4		0.3
Stock	53 936	99	0.5		0.5

TAB. 3 – Pourcentage d’entrant et écart d’affectation

résultat visible est le même : seule une partie des personnes orientées vers une prestation ont été effectivement prises en charge. Le tableau 3 illustre ce fait pour les OPP. Sur les 50563 demandeurs orientés OPP, seulement 47% ont finalement signé. Une partie de la déperdition tient au fait que ces personnes n’ont pas été signalées aux OPP (20%) : le chiffre de 47% sous-estime peut-être la force de conviction des OPP. Mais du point de vue de l’évaluation, seul ce ratio compte. Le tableau 3 permet de faire le même constat pour les demandeurs qui ont été orientés vers CVE : le taux de prise en charge effective sur l’année 2007 est un peu plus bas dans le flux indemnisable (43%), et plus encore dans le flux non indemnisable et dans le stock (34% et 36% respectivement).

Les conséquences de cette situation pour l’évaluation sont doubles. D’une part, les effectifs passés par l’OCC et qui contiennent de l’information sur l’effet des dispositifs sont sensiblement plus faibles que ce que le déploiement d’OCC pouvait faire espérer. Ils restent cependant élevés : 23 500 pour les OPP, 10 000 pour CVE dans le flux indemnisable et 26 000 et 31 000 dans le flux non indemnisable et dans le stock. D’autre part, et surtout, les demandeurs concernés ne sont plus comparables à la population témoin, car ils sont maintenant sélectionnés selon tout un ensemble de critères qu’il nous

est impossible de décrire. L'effet de cette sélection est apparent dans le tableau 1. Il traite séparément les demandeurs du flux indemnisable orientés classique ou OPP dans les zones OPP, les demandeurs du flux indemnisable orientés classique ou CVE dans les zones CVE, les demandeurs du flux non indemnisable et enfin les demandeurs du stock, orientés classique ou CVE dans les zones CVE. Les trois premières colonnes ont été présentées plus haut : elles contiennent les moyennes des caractéristiques individuelles dans les deux populations tirées au sort, donc semblables. La quatrième colonne présente ensuite les caractéristiques de la petite moitié de demandeurs qui, parmi ceux orientés vers OPP ou CVE, selon les cas, ont effectivement bénéficié de la prestation. Ces populations sont particulières et diffèrent par conséquent des populations de départ et notamment de l'échantillon témoin orienté classique. Dans l'ensemble, les personnes prises en charge sont plus souvent des femmes, sont plus éduquées et, plus souvent des techniciens que des ouvriers. Bien que les différences de moyenne entre la population de départ et cette population sélectionnée paraissent faibles, elles sont souvent statistiquement significatives. Ainsi, on ne saurait comparer simplement ces demandeurs pris en charge avec d'autres demandeurs de l'échantillon non pris en charge, sans risquer de confondre l'effet du programme avec celui de leurs différences d'éducation ou d'autres caractéristiques observées ou non. Nous présentons plus loin (section 4) la manière dont on peut procéder pour réaliser l'évaluation en tirant parti du tirage au sort initial, malgré ce processus de sélection dans les dispositifs.

Un autre phénomène apparaît dans le tableau 3 qui décrit le devenir des personnes orientées vers l'accompagnement classique³ : une proportion non-négligeable, de l'ordre de 3% en moyenne, a été prise en charge par un OPP. D'autres mouvements de cette nature (orientés CVE qui entrent en OPP, orientés classique qui entrent en CVE ou encore orientés CVE qui entrent en OPP) existent mais restent très rares. Les entrées en OPP de demandeurs orientés classique résultent de la forte pression exercée sur les réseaux par la nécessité d'alimenter les opérateurs privés. Il est par exemple arrivé que des listes constituées sans avoir recours à l'OCC aient été transmises aux OPP, alors qu'au même moment les demandeurs étaient orientés classique par l'OCC au cours de leur entretien de PPAE. Cette situation complique encore l'évaluation car le groupe témoin (les personnes orientées classique) contient

³Dans ce tableau, les "situations inconnues" correspondent pour partie aux entrées dans les dispositifs classiques qui ne sont pas documentées dans nos données pour l'instant.

des personnes prises en charge par les OPP qui sont, à leur tour, particulières. Le tableau 4 montre que ces situations ont été très fréquentes en Alsace et, dans une moindre mesure, dans le Centre, en Languedoc-Roussillon et en Provence-Alpes Côte d'Azur, surtout en milieu d'année, lorsque les tensions sur les OPP se sont fait le plus fortement sentir. A l'inverse, l'entrée en OPP des demandeurs orientés classique est restée exceptionnelle en Rhône-Alpes. Ces situations sont potentiellement problématiques pour l'évaluation. Leur traitement est détaillé dans l'annexe technique en fin de ce rapport. Elles sont davantage problématiques pour le programme CVE que pour les OPP. En effet le principe de base de l'évaluation est de dire que pour chaque programme, il y a deux populations : le groupe des individus potentiellement éligibles et groupe de contrôle. S'il n'y avait qu'un seul programme, le fait que des individus affectés au groupe de contrôle entrent dans le programme ne serait pas un problème. Cela réduirait au plus un peu la capacité de détection de l'évaluation et la portée des résultats. On aurait deux populations identiques à l'origine et deux programmes : le programme classique et le programme étudié, auquel chacune des deux populations participe avec des taux différents. Il y a là assez d'information pour mesurer la plus-value apportée par le programme. C'est ce qu'explique en détail la section consacrée à la méthodologie. Cette section explique aussi que ce que l'on mesure alors est la plus-value apportée par le programme sur une sous-population particulière que l'on définit à cette occasion et qui est fonction des taux d'entrée dans le programme dans les deux populations.

Le problème provient ici du fait que l'on a trois programmes, ce qui rend nécessaire des schémas d'entrée plus contraints : pas d'entrées des personnes orientées vers le programme classique dans les programme OPP et CVE, et, pour les groupes orientés CVE et OPP, entrées seulement dans le programme vers lequel on est orienté ou, à défaut, le programme classique. On peut néanmoins tirer parti du fait que les écarts ont toujours été des entrées chez les OPP. En effet, si on compare le groupe affecté au parcours classique et le groupe affecté aux OPP, le groupe affecté aux OPP entre soit dans le programme classique soit dans le programme OPP. Le groupe affecté au parcours classique entre soit dans le programme classique, soit (pour 3% d'entre eux) dans le programme OPP. Conformément à ce que l'on vient d'expliquer, les problèmes occasionnés sont limités : pour ces deux populations on a en effet deux programmes, OPP ou classique. On veut néanmoins se prémunir d'une trop grande hétérogénéité de la population sur laquelle le programme est évalué. C'est pourquoi il a été décidé de ne pas inclure

l'Alsace dans l'évaluation. Pour l'Alsace on a en effet un taux d'envoi des demandeurs d'emploi affectés au parcours classique chez les OPP de 33%, ce qui est considérable.

Pour CVE les problèmes sont potentiellement plus importants. En effet, si on considère les populations affectées à CVE et affectées au parcours classique, on a bien deux populations, mais ces deux populations sont susceptibles d'entrer dans trois programmes : le programme CVE pour les individus affecté au parcours CVE, le programme classique pour les individus affectés au programme classique ou les individus affectés au programme CVE mais n'y entrant pas, mais aussi le programme OPP pour les individus orientés vers le programme classique ou le programme CVE, mais réorientés vers le programme OPP et y entrant. On sent donc bien qu'il y a là un manque d'information. Il est néanmoins possible de le résoudre, à condition de faire l'évaluation des programmes OPP et CVE dans le flux indemnisable de façon conjointe et de supposer que l'effet du programme OPP est en moyenne le même dans les populations entrant dans ce programme. Pour limiter l'impact de ce problème, nous avons éliminé, parmi les régions dans lesquelles les programmes OPP et CVE ont été développés conjointement, la région Midi-Pyrénées pour laquelle 4% des demandeurs d'emplois affectés au programme classique s'étaient retrouvés chez les OPP.

2 Les données

2.1 Sources utilisées

Le fichier de données utilisées dans ce rapport provient de l'appariement de différentes sources qui permettent de rassembler l'information nécessaire à l'évaluation dans quatre domaines :

1. Les caractéristiques des demandeurs d'emploi (DE) entrés dans l'expérience (caractéristiques socio-démographiques) sont fournies par le Fichier historique (FH) de l'ANPE.
2. Le statut d'assignation par OCC aux groupes de contrôle et de traitement est historisé dans le fichier res_occ de l'ANPE.
3. L'accompagnement effectivement suivi est renseigné à partir de deux sources :
 - L'information sur le suivi d'un accompagnement CVE est issu de l'outil CVE développé par l'ANPE et historisé dans le fichier res_cve_opp ;

	Effectif			%		
	assigné Classique	Entrant OPP	CVE	OPP	CVE	Classique
Région OPP						
Alsace	127	42	0	33.1	0.0	66.9
Centre	78	6	0	7.7	0.0	92.3
Haute Normandie	230	44	0	19.1	0.0	80.9
Languedoc Roussillon	391	57	0	14.6	0.0	85.4
Alpes Provence Côte d'Azur	752	93	0	12.4	0.	87.6
Total	1 578	242	0	15.3	0	84.7
Total hors Alsace	1 451	200	0	13.8	0	86.2
Région CVE						
Aquitaine	1 503	0	25	0	1.7	98.3
Région CVE OPP						
Ile de France	4 032	113	6	2.8	0.1	97.0
Lorraine	863	9	5	1.0	0.6	98.4
Midi Pyrénées	277	11	1	4.0	0.4	95.7
Nord Pas de Calais	2 497	55	10	2.2	0.4	97.4
Rhone Alpes	2 963	15	17	0.5	0.6	98.9
Total	10 632	203	39	1.90	0.37	97.72
Total hors Midi Pyrénées	10 355	192	38	1.85	0.37	97.78

TAB. 4 – Ecart d'affectation par région

- L’information sur le suivi d’un accompagnement OPP est issu du fichier compilé par l’Unédic à partir des données remontées par les OPP (fichier Orientation_DES_Unédic).
4. Le taux de sortie vers l’emploi des DE à partir de leur entrée dans l’expérience provient de deux sources :
- Pour les DE restés inscrits à l’ANPE ou dont le motif de sortie a été bien renseigné, l’information sur l’emploi provient du FH ;
 - Pour les DE sortis du FH et dont le motif de sortie est inconnu, l’information sur l’emploi provient d’une enquête téléphonique. Cette enquête est effectuée sur un sous-échantillon seulement.

Une source de données complémentaire n’a pu être mobilisée dans le cadre de ce rapport : le Fichier national des allocataires (FNA) de l’Unédic. Cette source devrait permettre à terme d’approfondir le bilan économique des accompagnements, en intégrant dans l’analyse l’effet sur les indemnités chômage versées.

Dans la mesure où la détermination des taux de sortie vers l’emploi est cruciale pour l’évaluation, la sous-section suivante détaille la procédure mise en oeuvre pour les mesurer.

2.2 Information sur la sortie vers l’emploi : rôle de l’enquête téléphonique

Le FH sert de base à la mesure de la sortie vers l’emploi dans la mesure où il constitue une première source exhaustive de suivi des centaines de milliers de DE de l’expérience. Néanmoins, il ne suffit pas à mesurer la sortie vers l’emploi de tous, dans la mesure où de nombreuses demandes d’emploi sont annulées sans que le motif soit connu : il peut s’agir d’une sortie vers l’emploi aussi bien que d’une sortie vers l’inactivité. Sur le modèle de ”l’Enquête Sortants” mise en oeuvre chaque trimestre par l’ANPE, on a donc procédé à une enquête téléphonique auprès d’un sous-échantillon de demandeurs d’emploi sortis du FH sans qu’on sache s’il s’agissait ou non d’une reprise d’emploi. Cette enquête a été réalisée par un prestataire extérieur, l’institut de sondage Louis Harris 2 (LH2), déjà en charge de l’Enquête Sortants. Chaque mois, une nouvelle vague d’enquête a été effectuée auprès de demandeurs d’emploi récemment sortis du FH. Le pourcentage de demandeurs d’emploi interrogé diffère selon les groupes considérés (DE issus du stock, du flux, et tirés au

sort CVE, OPP ou classique) : il a été choisi de façon à maximiser la précision de l'évaluation.

Le tableau 5 récapitule les grandes données de l'enquête pour les six premières cohortes d'OCC (de janvier à juin 2007) et pour un horizon de sortie vers l'emploi de six mois. La partie gauche du tableau donne différents effectifs : le nombre total de demandeurs passés par OCC, le nombre de ceux qui sont sortis du FH dans les six mois qui ont suivi, le nombre de ceux dont le motif de sortie n'a pas été renseigné, le nombre de ceux qui ont été échantillonnés pour être interrogés par téléphone, et le nombre de ceux qui ont pu être contactés et qui ont répondu à l'enquête téléphonique. La partie droite du tableau permet d'analyser ces effectifs à partir de quelques ratios : taux de sortie, proportion de sorties dont le FH ne permet de connaître le motif, proportion de ces demandeurs pour lesquels on met en place une enquête téléphonique, taux de réponse à cette enquête et, finalement, taux de demandeurs d'emploi sortis du FH dont on connaît le statut d'emploi en cumulant information administrative et résultats de l'enquête téléphonique.

On constate tout d'abord que la part des sorties dont les motifs ne sont pas renseignés dans le FH est extrêmement important : 65% en moyenne. Cette part varie beaucoup d'une population à l'autre. Elle est la plus élevée dans le flux non indemnisable (trois demandeurs sur quatre). Vraisemblablement, ces demandeurs d'emploi ne se sentent pas tenus d'informer l'ANPE de leur motif de sortie dans la mesure où ils ne sont pas indemnisés. Elle est de l'ordre de 50% en moyenne dans le flux indemnisable. Elle prend une valeur intermédiaire dans le stock.

Les taux de tirage pour l'enquête ont été choisis de façon à maximiser la précision de l'évaluation. Ils sont donc d'autant plus importants que la part de sorties non renseignées est élevée, et d'autant plus faibles que l'échantillon initial est important. Par exemple, le taux de tirage pour l'enquête est de 100% pour les groupes CVE et classique du flux indemnisable. Ces groupes avaient en effet une faible probabilité de tirage dans OCC et leur effectif limité réduit la précision de l'analyse : on s'efforce d'améliorer cette précision par un suivi renforcé des sorties.

Comparés à la part de motifs renseignés dans le FH, le taux de réponse à l'enquête téléphonique est beaucoup plus homogène d'une population à l'autre. Il avoisine en général les 55%.

La dernière colonne du tableau résume les résultats de l'enquête. Il est important pour l'évaluation que les taux de réponse diffèrent le moins possible entre groupes de traitement et groupe de contrôle. La crainte serait en effet

	Entré dans OCC a	Effectif		Sortie		Enq. f	Rép. g	Part				DE statut connu ^a	
		FH à 6 mois Prés. b	Sorti c	connue d	Inc. e			sortie c/a	Inc. e/c	enq. f/e	rép. enquête g/f		
Assignation OCC	A1. Flux indemnisable, zones avec OPP et CVE												
OPP	19579	12746	6833	3426	3407	1086	530	35	50	32	49	91	
CVE	2122	1297	825	431	394	394	180	39	48	100	46	90	
CLA	3422	2178	1244	582	662	662	372	36	53	100	56	92	
	A2. Flux indemnisable, zones avec CVE seulement												
CVE	1993	1022	971	484	487	487	280	49	50	100	57	90	
CLA	2241	1288	953	438	515	515	306	43	54	100	59	91	
	A3. Flux indemnisable, zones avec OPP seulement												
OPP	3860	2494	1366	720	646	184	90	35	47	28	49	91	
CLA	692	473	219	118	101	101	55	32	46	100	54	93	
	B. Flux non indemnisable												
CVE	11436	4565	6871	1690	5181	3281	1829	60	75	63	56	80	
CLA	16542	7042	9500	2289	7211	3402	1844	57	76	47	54	80	
	C. Stock												
CVE	18250	10581	7669	2710	4959	2009	1118	42	65	41	56	88	
CLA	40222	23932	16290	5429	10861	2625	1421	41	67	24	54	88	
	D. Total												
Tous	120359	67618	52741	18317	34424	14746	8025	44	65	43	54	87	

$$^a(b + d + e \times g/f)/a$$

TAB. 5 – Statut à 6 mois des demandeurs d'emploi passés par OCC

que ceux qui ne bénéficient pas d'un accompagnement renforcé répondent moins volontiers à l'enquête. Cela biaiserait les résultats de l'évaluation si on ne distinguait pas (dans les comparaisons entre groupes de traitement et groupe de contrôle) ce qui relève de différences de reprise d'emploi et ce qui relève simplement de différences de comportement de réponse. Ici, ces craintes semblent relativement peu fondées. Si on prend l'exemple du flux indemnisable dans les zones où ont coexisté les dispositifs CVE et OPP, on constate en effet que les taux de réponse obtenus sont très proches. Dans les deux groupes de traitement (CVE et OPP), ils sont même virtuellement identiques (91 et 90%). Autrement dit, l'utilisation jointe de l'information administrative du FH et de l'enquête téléphonique permet de connaître le statut d'emploi des demandeurs sortis (parmi ceux qui sortent des fichiers de l'ANPE) dans les mêmes proportions que le demandeur d'emploi ait été assigné CVE ou OPP. Il est donc peu vraisemblable que la non réponse soit une source de biais dans la comparaison entre les accompagnements CVE et OPP. Le taux de réponse dans le groupe de contrôle (assigné classique) est très proche (92%). S'il y a une différence et si elle est due à un comportement de réponse différent, c'est, contrairement à ce qu'on pouvait attendre, que le groupe de contrôle a un meilleur taux de réponse à l'enquête téléphonique.

Si on regarde les autres populations, on constate de la même façon que les groupes de traitement et de contrôle ont des taux de réponse similaires. Les écarts sont minimes, et non systématiques : tantôt, le groupe de traitement répond davantage, tantôt, c'est le groupe de contrôle. Dans aucun cas les différences de comportement de réponse ne semblent susceptibles d'affecter massivement les résultats.

En résumé, l'effort considérable mené avec l'enquête téléphonique a permis d'obtenir des taux de réponse très élevés (de l'ordre de 90%), et comparables dans les groupes de traitement et de contrôle. Ces taux de réponse élevés sont indispensables pour tirer tout le parti de l'orientation aléatoire par OCC. Pris ensemble, orientation aléatoire et taux de réponse élevés justifient l'hypothèse que les échantillons de répondants dans les groupes de contrôle et de traitement sont comparables en tout, excepté le traitement. On traite alors la non-réponse résiduelle en la considérant comme aléatoire (*missing at random*) : les non-répondants de chaque groupe (contrôle et traitement) sont représentés par les répondants du même groupe, en utilisant les poids appropriés (voir l'annexe). Comme cela se fait souvent pour redresser la non-réponse (dans des enquêtes où la non-réponse est en général beaucoup plus élevée), il est également possible de calculer des poids spécifiques par

groupes sociodémographiques. Cela a été fait pour vérifier la robustesse des résultats ; il s'avère que ces derniers sont pratiquement inchangés lorsqu'on procède de la sorte.

L'enquête téléphonique a une conséquence sur les données disponibles dans le cadre de ce rapport. Dans la mesure où la procédure est assez lourde (repérage des sorties non suivies de nouvelle demande dans les trente jours, échantillonnage, envoi d'un courrier, puis enquête téléphonique), il n'a pas été possible de connaître le statut d'emploi des demandeurs d'emploi sortis du FH après le 31 décembre 2007. Dans tous les cas, l'analyse des taux de sortie vers l'emploi s'arrête donc fin 2007.

2.3 Constitution du fichier d'analyse et cohortes étudiées

Les principes de constitution du fichier d'analyse sont les suivants. Tout d'abord, les fichiers issus des différentes sources, anonymisés, sont appariés à partir d'un identifiant banalisé. Le champ d'analyse est restreint ensuite aux cohortes entrées dans OCC après le 1er janvier 2007, en excluant les régions Alsace et Midi-Pyrénées où le protocole n'a pas été suivi de façon satisfaisante.

Sur les douze cohortes de 2007 étudiées dans ce rapport, le fichier comporte 211 360 demandeurs d'emploi. Cependant, les douze cohortes sont rarement toutes mobilisées dans l'analyse. Comme on vient de le préciser, on ne dispose d'un suivi des transitions vers l'emploi que jusqu'à fin décembre 2007. L'analyse des sorties vers l'emploi doit donc porter sur un sous-ensemble de cohortes. Ce sous-ensemble est d'autant plus limité qu'on étudie la sortie vers l'emploi à un horizon long. Par exemple, si on veut étudier le taux de sortie vers l'emploi à 3 mois, on ne peut analyser que les 9 premières cohortes, passées par OCC entre janvier et septembre : à partir de la cohorte d'octobre, en effet, certaines reprises d'emploi à trois mois interviennent en 2008 et ne sont pas observées. De façon similaire, les taux de retour à 6 mois ne peuvent être étudiés à ce stade que pour les six premières cohortes, passées par OCC entre janvier et juin 2007. Les données mobilisées varient donc selon l'horizon de sortie vers l'emploi étudié.

3 Méthodologie

L'évaluation de l'impact de l'accompagnement renforcé repose sur la comparaison des taux de sortie vers l'emploi, à différents horizons, de demandeurs d'emploi qui ont été orientés aléatoirement soit vers un accompagnement renforcé (OPP ou CVE) soit vers le parcours classique de l'ANPE. Les deux premiers groupes constituent les groupes de traitement (ou groupes test), le troisième constitue le groupe de contrôle (ou groupe témoin). Ces trois groupes, puisqu'ils sont issus d'une seule et même population par tirage au sort, sont au départ statistiquement identiques. Si leurs taux de sortie vers l'emploi diffèrent, on peut donc attribuer cette différence à l'accompagnement suivi et à lui seul. Autrement dit, dans le cas le plus simple, l'écart de taux de sortie vers l'emploi entre le groupe de traitement CVE (respectivement OPP) et le groupe de contrôle mesure la plus-value de l'accompagnement CVE (respectivement OPP) par rapport à l'accompagnement classique de l'ANPE. Les estimateurs mis en oeuvre dans ce rapport sont légèrement plus élaborés mais reposent sur le même principe en l'adaptant aux circonstances spécifiques de l'expérimentation. Cette section présente de manière intuitive les estimateurs utilisés ; l'annexe méthodologique les décrit de façon plus précise, dans un cadre plus formel.

3.1 Des évaluations séparées dans trois populations

L'accompagnement CVE a été proposé à trois populations distinctes de demandeurs d'emploi : ceux nouvellement inscrits à l'ANPE et disposant de droit à l'indemnisation pour un an au moins (le "flux indemnisable"), les nouveaux inscrits disposant d'une durée d'indemnisation inférieure (le "flux non indemnisable") et les demandeurs d'emploi inscrits depuis plus de trois mois (le "stock"). L'accompagnement OPP n'a été proposé qu'aux demandeurs d'emploi du flux indemnisable. Ces trois populations présentent des dynamiques de retour à l'emploi différentes et il est probable que la plus-value d'un accompagnement renforcé n'est pas la même pour chacune. Il est donc nécessaire d'évaluer séparément la plus-value de l'accompagnement CVE sur chacun des trois groupes.

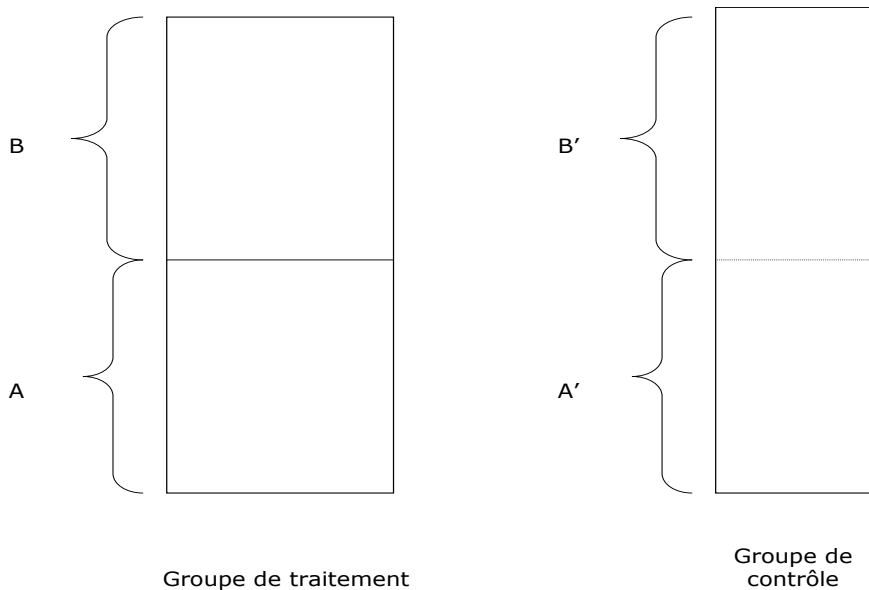
3.2 Prise en compte des écarts entre l'orientation aléatoire et l'entrée effective en accompagnement

L'entrée en accompagnement CVE ou OPP repose sur l'adhésion volontaire du demandeur d'emploi. Toutes les personnes orientées vers CVE (respectivement OPP) ne sont donc finalement pas entrées en accompagnement CVE (respectivement OPP). Symétriquement certaines personnes orientées vers le parcours classique ont finalement bénéficié d'un accompagnement OPP ou CVE (voir section 1.3).

Prenons l'exemple de l'évaluation du programme CVE dans le flux indemnisable. On constate qu'en moyenne environ 50% des personnes orientées aléatoirement vers CVE entrent effectivement en accompagnement CVE. En revanche, la quasi-totalité de ceux qui sont orientés vers le parcours classique suivent ce parcours. Il peut sembler au premier abord que cette situation remette en cause le principe même de l'évaluation. Si on compare le groupe de traitement dans son ensemble (groupes A et B sur la figure) avec le groupe de contrôle, on ne mesure pas l'effet de CVE sur A et B puisque seul A a suivi l'accompagnement. Si on compare le groupe A (seul) avec le groupe de contrôle, on n'obtient pas non plus l'effet de l'accompagnement sur A. En effet, le groupe A n'est pas comparable avec le groupe de contrôle : il s'est par exemple auto-sélectionné en acceptant d'entrer dans le dispositif CVE et on peut penser que cette auto-sélection est liée à des caractéristiques propres, inobservées, potentiellement liées au taux de sortie vers l'emploi. La comparaison entre A et le groupe de contrôle mêle alors deux choses : l'effet d'accompagnement CVE et l'effet des caractéristiques particulières de ceux qui entrent en CVE. Autrement dit, on retombe dans le problème du biais de sélection (ou de la comparaison de groupes non comparables) que l'affectation aléatoire vise précisément à éviter.

L'impasse n'est qu'apparente. Il faut voir que le groupe de contrôle est en réalité composé, même si on ne les observe pas, de deux sous-groupes A' et B' exactement comparables à A et B. A' est ainsi le groupe de ceux qui seraient effectivement entrés en accompagnement CVE si le tirage au sort les avait orientés vers cet accompagnement. On peut alors décomposer la différence entre le groupe de traitement et le groupe de contrôle en deux :

- D'une part, la différence entre A et A' mesure l'effet de l'accompagnement CVE sur A, puisque les deux groupes ne diffèrent que par l'accompagnement reçu ;
- D'autre part, on sait *a priori* que la différence entre B et B' est nulle :



les deux groupes sont statistiquement identiques et ont suivi le même accompagnement (classique).

L'écart entre les taux de retour à l'emploi des groupes de traitement et de contrôle peut donc être attribué *uniquement* à l'effet de l'accompagnement CVE sur le groupe A. Supposons que cet écart soit de 10 points de pourcentage. On en déduit que le taux de retour à l'emploi de A a augmenté de 20 points de pourcentage. En effet, un écart de 20 points sur la moitié des groupes de contrôle et de traitement (A et A') et de 0 point sur l'autre moitié (B et B') implique un écart moyen de 10 points ($20 \times 0.5 + 0 \times 0.5 = 10$). Autrement dit, l'évaluation repose bien toujours sur une différence entre le groupe de traitement et le groupe de contrôle tels qu'ils ont été déterminés par le tirage au sort. Mais on pondère cette différence par l'inverse de la proportion de ceux qui sont effectivement entrés dans l'accompagnement CVE. On mesure ainsi l'effet de CVE sur ceux qui sont effectivement entrés dans le programme (le groupe A, par rapport au groupe A').

Plus généralement, on peut étendre le raisonnement au cas où des demandeurs d'emploi du groupe de contrôle entrent néanmoins en accompagnement CVE. Dans ce cas, l'estimateur est la différence entre le groupe de traitement et le groupe de contrôle pondérée par l'inverse de la proportion de *compliers*— le terme désignant ceux qui n'auraient pas suivi l'accompagne-

ment CVE s'ils n'avaient pas été orientés aléatoirement vers celui-ci, mais qui le suivent lorsque le tirage au sort les y incite.

Cet estimateur s'écrit :

$$\hat{\beta} = \frac{\bar{Y}_t - \bar{Y}_c}{\bar{T}_t - \bar{T}_c},$$

où \bar{Y}_t et \bar{Y}_c mesurent les taux de retour à l'emploi dans les groupes de traitement et de contrôle, et \bar{T}_t et \bar{T}_c mesurent les taux d'entrée en accompagnement CVE dans ces deux groupes. Il peut s'obtenir par une estimation à variables instrumentales de Y sur T en instrumentant T par Z , variable indicatrice du fait que le demandeur d'emploi a été orienté aléatoirement vers CVE.

Il est connu sous le nom d'estimateur de Wald. Au-delà des considérations techniques, l'important est de bien l'interpréter comme une *mesure de l'effet de l'impact du dispositif sur un groupe particulier, celui des compliers (le groupe A)*. On ne sait rien de ce qu'aurait été l'effet de CVE sur le groupe B. En d'autres termes, l'évaluation renseigne uniquement sur l'effet moyen de l'accompagnement renforcé lorsque l'adhésion est volontaire et que seulement 50% de ceux auxquels le programme est proposé y adhèrent ; l'effet moyen sur l'ensemble des demandeurs d'emploi si l'adhésion était obligatoire et si 100% des demandeurs d'emploi adhérerait serait potentiellement différent.

Dans tout ce qui précède pour simplifier l'exposition, on a décrit l'entrée dans le dispositif comme résultant seulement du choix du demandeur d'emploi. Il se peut que les équipes CVE ou OPP aient également procédé à une sélection plus ou moins explicite, décourageant ou réorientant certains demandeurs d'emploi (pratique de l'*écrémage*). L'estimateur de Wald reste valide pour évaluer l'effet de l'accompagnement renforcé sur ceux qui ont suivi cet accompagnement. Mais il faut garder à l'esprit qu'il s'agit de l'effet du programme sur ce groupe bien précis : les demandeurs d'emploi qui ont choisi d'adhérer au dispositif et que les agents de l'ANPE (ou des OPP) n'ont pas dissuadés.

3.3 Prise en compte des variations des probabilités d'orientation vers les différents dispositifs

Pour permettre la montée en charge des dispositifs d'accompagnement renforcé et éviter leur saturation, on a fait varier les probabilités d'orientation vers les différents accompagnements, d'une ALE à une autre, et d'une période à l'autre.

Il en résulte des effets de composition au sein des groupes de traitement et de contrôle. Le groupe de traitement CVE sous-pondère les zones et périodes pour lesquelles peu de DE étaient orientés vers CVE (et surpondère les autres). Quand peu de DE étaient orientés vers CVE, c'est en général que beaucoup l'étaient vers OPP : par conséquent, le groupe de traitement OPP surpondère les zones où CVE est sous-pondéré (et sous-pondère les autres). Pour que ces effets de composition ne viennent pas biaiser les résultats, trois méthodes d'estimation alternatives sont possibles :

- La première contrôle les effets de composition en introduisant autant de variables indicatrices qu'il y a de zones et périodes où les probabilités d'orientation vers les trois dispositifs sont restées constantes. Cet estimateur *within* repose alors sur des comparaisons au sein de chacune de ces zones et périodes. Les effets de composition sont ainsi neutralisés. L'endogénéité des variables de participation au programme T est neutralisée en utilisant comme variables instrumentales les variables d'affectation au groupe de programme potentiel Z .
- La deuxième méthode contrôle comme la première les effets de composition en introduisant autant de variables indicatrices qu'il y a de zones et périodes où les probabilités d'orientation vers les trois dispositifs sont restées constantes. Mais l'estimateur utilise comme variables instrumentales pour les variables de participation au programme T , les variables d'affectation au groupe de programme potentiel Z ainsi que leur interaction avec les indicatrice de zone de probabilité constante.
- La dernière méthode consiste à repondérer les échantillons des groupes de contrôle et de traitement. Avec des poids appropriés, on s'assure que chaque groupe représente bien la population initiale moyenne, toutes ALE et périodes confondues.

Les implications de ces trois méthodes sont développées dans l'annexe méthodologique et des résultats de sensibilité sont présentés dans le tableaux 10. En pratique, les trois ont été mises en oeuvre et les résultats sont robustes au choix effectué. C'est la première qui est retenue pour la plupart des tableaux.

4 Résultats

Le tableau 6 présente les résultats globaux de l'évaluation. Il porte sur l'impact de la participation au programme sur les **sorties des listes ANPE**

vers l'emploi à différents horizons suivant la date de passage par OCC. On examine l'effet sur les sorties à trois mois, sur les sorties à 6 mois et sur les sorties à 9 mois. Ces résultats sont obtenus en utilisant toute l'information disponible hors celle correspondant aux régions Alsace et Midi-Pyrénées dans lesquelles le protocole expérimental n'a pas été respecté. Les résultats à 3 mois mobilisent donc les informations sur les 9 premières cohortes d'envoi, c'est-à-dire de janvier à septembre. Les résultats à 6 mois mobilisent les informations des 6 premières cohortes, de janvier à juin et enfin les résultats à 9 mois mobilisent les informations des trois premières cohortes de janvier à mars. Les effectifs introduits dans ces analyses sont donc différents en fonction de l'horizon. Cela a une conséquence directement visible sur la précision des estimations qui est très élevée pour les résultats à 3 mois et se dégrade fortement pour les résultats à 9 mois. De même les informations utilisées concernent toutes les zones dans lesquelles les programmes CVE ou OPP ont été développés. On n'opère pas de distinction entre les zones dans lesquelles seul le programme CVE a été développé et celles où il a été développé conjointement avec le programme OPP, et réciproquement.

La partie haute du tableau donne les résultats pour le flux indemnisable. La partie droite concerne CVE et la partie gauche les OPP. La partie inférieure du tableau concerne seulement le programme CVE. La partie droite donne les résultats pour le stock et la partie gauche les résultats pour le flux non indemnisable. Pour chacune de ces populations et chacun de ces programmes, on donne d'abord l'estimation de l'impact sur le taux de sortie, puis l'écart-type. Enfin on donne la borne inférieure de l'intervalle de confiance à 95%, la borne supérieure de l'intervalle de confiance et les effectifs impliqués dans la production du résultat. Si la borne inférieure est positive, cela signifie que le programme accroît le retour à l'emploi de façon statistiquement significative. Symétriquement, si la borne supérieure est négative, cela signifie que le programme accroît la rétention en chômage. Lorsque la borne inférieure est négative et la borne supérieure positive, on ne détecte pas d'effet statistiquement significatif sur la sortie en emploi.

Evaluation CVE et OPP dans le flux indemnisable									
	Effet de l'accompagnement OPP			Effet de l'accompagnement CVE			Borne		
	Est	Std	Effectif	Est	Std	Effectif	Inf.	Sup.	Effectif
Emploi à 3 mois	1.7	(1.5)	53261	5.5	(1.4)	53261	2.7	8.4	53261
Emploi à 6 mois	3.0	(2.4)	33909	8.4	(2.3)	33909	3.7	13.1	33909
Emploi à 9 mois	5.7	(4.6)	10793	16.9	(4.8)	10792	7.3	26.4	10792

Evaluation CVE dans le flux non indemnisable et le stock									
	Flux non indemnisable			Stock			Borne		
	Est	Std	Effectif	Est	Std	Effectif	Inf.	Sup.	Effectif
Emploi à 3 mois	4.1	(1.5)	42722	2.3	(1.3)	42722	-0.3	4.9	71626
Emploi à 6 mois	6.1	(2.3)	27978	4.7	(1.8)	27978	1.0	8.4	58472
Emploi à 9 mois	1.9	(3.8)	11345	6.2	(2.9)	11345	0.5	11.9	35324

TAB. 6 – Evaluation des programmes OPP et CVE - Toutes zones confondues - Toutes cohortes confondues

Les résultats montrent que le programme CVE a un effet positif sur la sortie en emploi à chacun des horizons considérés. L'effet sur la sortie en emploi est croissant avec l'horizon considéré : 5.5 points à 3 mois, 8.4 points à 6 mois et 16.9 points à 9 mois. On ne détecte pas en revanche d'effet du programme OPP sur les sorties vers l'emploi. On est dans la situation dans laquelle la borne inférieure de l'effet est négative et la borne supérieure positive. On note aussi que bien que le coefficient ne soit pas significativement différent de 0, l'effet estimé tend à croître avec le temps : 1.7 points à 3 mois, 3.0 points à 6 mois et 5.7 points à 9 mois. On note en outre que la précision des estimations est similaires entre les deux programmes.

Le programme CVE dans le flux non indemnisable et dans le stock a un effet moindre. On note qu'il n'est pas toujours significativement différent de 0. Dans le stock, l'effet estimé est croissant avec le temps et est significativement différent de zéro dès l'horizon de 6 mois. L'impact est de 2.3 points à 3 mois (pas significativement différent de zéro), de 4.7 points à 6 mois et enfin de 6.2 points à 9 mois. Dans le flux non indemnisable, les effets à 3 et 6 mois sont significatifs et de bonne taille mais l'effet à 9 mois se réduit fortement et n'est plus significativement différent de 0.

Les tableaux 7 et 8 montrent les résultats respectivement dans le flux indemnisable, le flux non indemnisable et le stock, par cohorte (c'est-à-dire selon la date d'orientation par l'OCC). Ces tableaux présentent d'abord les résultats obtenus pour les cohortes de janvier à mars. Dans ce cas sont présentés les résultats à 3 mois, 6 mois et 9 mois. Ils présentent ensuite les résultats obtenus sur les cohortes d'avril à juin, dans ce cas seuls les résultats à 3 et 6 mois sont présentés. Enfin ils présentent les résultats obtenus pour les cohortes de juillet à septembre. Dans ce cas, seuls sont disponibles les résultats à 3 mois.

Ces tableaux montrent que les résultats dépendent fortement de la cohorte considérée. Dans le flux indemnisable, les résultats sont bien plus élevés pour les trois premières cohortes que pour les trois suivantes. Ceci concerne aussi bien le programme CVE que le programme OPP. La réduction est très importante pour le programme CVE. L'effet sur la sortie à 3 mois est estimé à +12 points sur les trois premières cohortes. Il n'est que de 3.1 points sur les trois suivantes, soit une réduction de 8.9 points. En outre l'effet n'est plus significativement différent de 0. Pour les 3 cohortes suivantes, l'effet estimé est un peu plus élevé et significativement différent de 0 : 5.5 points, soit une réduction de 6.5 points par rapport au niveau des trois premières cohortes. Pour l'effet à 6 mois, on note aussi une forte décroissance de l'effet estimé,

de 15.6 pour les trois premières cohortes à 5.3 pour les trois suivantes, en outre à la limite de la significativité. La réduction est aussi sensible pour le programme OPP, bien que de moindre ampleur. On note que les effets sur les trois premières cohortes, bien que largement non significatifs, sont élevés : 5.8 points à 3 mois et 7.0 points à 6 mois. Les effets sur les trois cohortes suivantes sont très faibles : 0.6 point à trois mois et un effet nul à 6 mois. Sur les trois dernières cohortes, la performance se dégrade encore un peu et l'effet devient négatif, bien que le paramètre ne soit pas significativement différent de zéro.

Evaluation CVE et OPP dans le flux indemnisable										
	Effet de l'accompagnement OPP			Effet de l'accompagnement CVE			Borne			
	Est	Std	Effectif	Est.	Std.	Effectif	Inf.	Sup.	Effectif	
Cohorte Janvier à Mars										
Emploi à 3 mois	5.8	(3.3)	-0.7	12.4	10793	12.0	(3.8)	4.4	19.6	10793
Emploi à 6 mois	7.0	(4.3)	-1.5	15.5	10793	15.6	(4.6)	6.5	24.7	10793
Emploi à 9 mois	5.7	(4.6)	-3.5	14.8	10793	16.9	(4.8)	7.3	26.4	10793
Cohorte Avril à Juin										
Emploi à 3 mois	0.6	(2.2)	-3.9	5.1	23116	3.1	(2.2)	-1.3	7.6	23116
Emploi à 6 mois	0.0	(3)	-5.9	5.9	23116	5.3	(2.7)	-0.1	10.7	23116
Cohorte Juillet à Septembre										
Emploi à 3 mois	-0.6	(2.3)	-5.3	4.0	19352	5.5	(2.1)	1.3	9.7	19352

TAB. 7 – Evaluation des programmes OPP et CVE dans le flux indemnisable - Toutes zones confondues -
Par cohortes

On observe la même variabilité des résultats pour l'évaluation de CVE dans les autres publics, stock et flux non indemnisable. Les résultats sont présentés dans le tableau 8. L'évolution est néanmoins dans ce cas en sens inverse : on observe une amélioration des performances du programme entre les trois premières cohortes et les trois cohortes suivantes. C'est particulièrement net pour le flux non indemnisable. L'effet à 6 mois progresse de 2.4 points (non significatif), à 8.6 (largement significatif). Dans le stock, on observe aussi la même instabilité, mais les évolutions sont en outre erratiques. Ainsi, pour l'effet à 3 mois, on observe d'abord un résultat positif, mais faible et non significatif, puis un effet plus important et significatif et enfin un effet négatif et non significatif.

Evaluation de CVE										
Flux non indemnisable					Stock					
Borne					Borne					
	Est.	Std.	Inf.	Sup.	Effectif	Est.	Std.	Inf.	Sup.	Effectif
Cohortes Janvier à Mars										
Emploi à 3 mois	0.9	(3.3)	-5.6	7.4	11345	1.8	(2.1)	-2.4	6.0	35324
Emploi à 6 mois	2.4	(3.7)	-4.9	9.8	11345	4.2	(2.6)	-0.9	9.4	35324
Emploi à 9 mois	1.9	(3.8)	-5.7	9.5	11345	6.2	(2.9)	0.5	11.9	35324
Cohortes Avril à Juin										
Emploi à 3 mois	7.2	(2.6)	2.1	12.3	16633	4.5	(2.1)	0.2	8.8	23148
Emploi à 6 mois	8.3	(2.9)	2.5	14.2	16633	5.1	(2.6)	-0.2	10.4	23148
Cohortes Juillet à Septembre										
Emploi à 3 mois	3.1	(2.4)	-1.8	7.9	14744	-0.2	(2.6)	-5.4	4.9	13154

TAB. 8 – Evaluation des programmes CVE dans le flux non indemnisable et le stock- Toutes zones confondues
- Par cohortes

Une autre dimension d'hétérogénéité intéressante concernant l'évaluation dans le flux indemnisable est la zone sur laquelle le programme a été développé. On peut opérer en effet une distinction entre les zones sur lesquelles les programmes OPP et CVE ont été développés conjointement et les zones sur lesquelles ils ont été développés seuls. Les résultats sont présentés dans le tableau 9. La partie supérieure du tableau présente les résultats sur les zones communes, et la partie inférieure les résultats sur les autres zones. On note d'abord la forte imprécision des résultats. Pour les résultats à 9 mois l'écart-type estimé est ainsi de l'ordre de 7 pour le programme CVE et entre 6 et 10 pour le programme OPP. On observe de meilleures performances du programme CVE sur les zone commune mais des résultats moins nets pour le programme OPP. On observe en effet une amélioration pour l'effet à 6 mois mais une détérioration pour l'effet à 9 mois. Au total, compte tenu de l'importance des écart-types, on ne trouve pas de résultats significativement différents suivant le type de zone dans lequel le programme a été développé.

Il est utile de reconstituer ce qu’auraient été les taux de sortie vers l’emploi des bénéficiaires de l’accompagnement s’ils n’avaient pas bénéficié de cet accompagnement (situation contrefactuelle). En d’autres termes, on calcule le taux de sortie vers l’emploi des différents groupes de la figure de la section 3, selon qu’ils bénéficient ou pas du traitement : le groupe A ou A’ (les *compliers* selon qu’ils bénéficient effectivement ou non de l’accompagnement), le groupe B ou B’ (ceux qui n’entrent pas dans l’accompagnement renforcé, pour les diverses raisons détaillées à la section 2). Ces taux sont directement observés pour A et B, celui pour A’ doit être reconstitué.⁴ Dans le cas où les demandeurs d’emploi affectés au groupe de contrôle n’entrent pas dans le programme (cas par exemple de CVE là où il est développé seul), on l’obtient simplement en retranchant l’impact estimé de la situation moyenne observée pour les bénéficiaires. Dans le cas où des demandeurs d’emploi du groupe de contrôle entrent dans le programme (cas des zones dans lesquelles le programme OPP a été développé, surtout dans celles où il a été développé seul), la formule est un peu plus compliquée mais repose sur un principe similaire. Elle est donnée en annexe. Dans le cas des zones où CVE et OPP ont été développés conjointement, cette reconstitution n’est qu’une approximation, qu’il faut donc prendre avec prudence.

Les figures 1 à 6 montrent les taux de sortie ainsi obtenus pour des horizons de 3 et 6 mois. A chaque fois, la première barre donne le taux de sortie du groupe B, la seconde présente celui des *compliers* sans accompagnement (en clair) et la plus-value apportée par l’accompagnement (en grisé). Enfin, la troisième barre donne le taux d’emploi moyen qui en résulte, lorsqu’on fait la moyenne des groupes A et B, pondérée par leurs parts respectives dans chaque population. A nouveau, ce taux est présenté sans accompagnement (en clair), et la plus-value de l’accompagnement, qui ne provient que des *compliers*, est indiquée en grisé.

L’écart entre la barre de gauche et celle du milieu (hors effet de l’accompagnement, c’est-à-dire parties claires seulement) traduit la sélection à l’entrée dans chaque dispositif. Le fait que ceux qui n’entrent pas dans le dispositif aient un taux de sortie vers l’emploi plus élevé traduit le fait que ces demandeurs ont des caractéristiques, observables ou non, qui leur permettent de trouver plus rapidement un emploi que ceux qui entrent en accompa-

⁴Comme on l’a vu section 3, celui de B’ est égal à celui de B, puisqu’il s’agit de demandeurs d’emploi qui ne bénéficient jamais de l’accompagnement renforcé, quelle que soit leur orientation initiale.

ment (et ce indépendamment de l'existence d'un accompagnement renforcé ou non). Il convient de rappeler que cette *sélection* peut venir de nombreuses sources : elle peut être liée au choix du demandeur d'accepter ou non l'accompagnement, aux pratiques d'orientation des conseillers ANPE ou des équipes CVE et OPP, ou tout simplement au délai entre orientation et entrée effective en accompagnement. En effet, si certains demandeurs retrouvent très vite un emploi, ils n'ont tout simplement pas le temps d'entrer effectivement en accompagnement, alors qu'il l'auraient fait pour certains s'ils étaient restés plus longtemps en recherche d'emploi.

Le commentaire de ces figures se focalise ici sur les taux de sortie à 6 mois, horizon où les programmes ont eu le temps de produire des effets et pour lesquels des échantillons suffisants sont disponibles.

On examine d'abord les effets de CVE et OPP dans les zones où les deux dispositifs sont présents conjointement. Comme il s'agit des mêmes zones, les taux de sortie vers l'emploi hors programmes d'accompagnement renforcés sont les mêmes ; à 6 mois, ce taux est de 23,2%. Cependant, même si la population de départ est la même, les populations qui rentrent en accompagnement CVE d'une part, et en accompagnement OPP d'autre part, ne semblent pas avoir les mêmes caractéristiques initiales. En effet, *s'ils n'avaient pas été accompagnés*, les bénéficiaires d'OPP auraient eu des taux de sortie plus élevés que les bénéficiaires de CVE (à 6 mois, 21.5% comparés à 17%). En d'autres termes, les OPP attirent ou retiennent des demandeurs d'emploi dont les chances initiales de sortir du chômage sont plus élevées. Sur ces demandeurs d'emploi, la plus-value du programme OPP (une augmentation des chances de sortie vers l'emploi de 1.2 points de pourcentage) est plus faible que la plus-value du programme CVE sur ses propres bénéficiaires (une augmentation des chances de sortie vers l'emploi de 9.2 points de pourcentage). Attention cependant : ces effets estimés, comme le montre le tableau 9 qui les présente avec leurs intervalles de confiance, sont entachés d'une incertitude statistique importante. Ils demandent donc à être confirmés lorsque des échantillons plus importants seront disponibles.

On examine ensuite les effets dans le flux indemnisable lorsque les OPP ou CVE sont présents seuls. Il ne s'agit pas des mêmes zones géographiques, et on constate des différences de départ importantes : le taux moyen initial de retour à l'emploi à 6 mois est beaucoup plus élevé là où CVE est implanté que là où les OPP sont implantés (29.4% contre 20.1%). Ceux qui entrent en accompagnement renforcé auraient eu des chances de sortie vers l'emploi plus faibles que les autres (en l'absence d'accompagnement). Cette fois, cet écart

Effet d'OPP - zones communes CVE-OPP

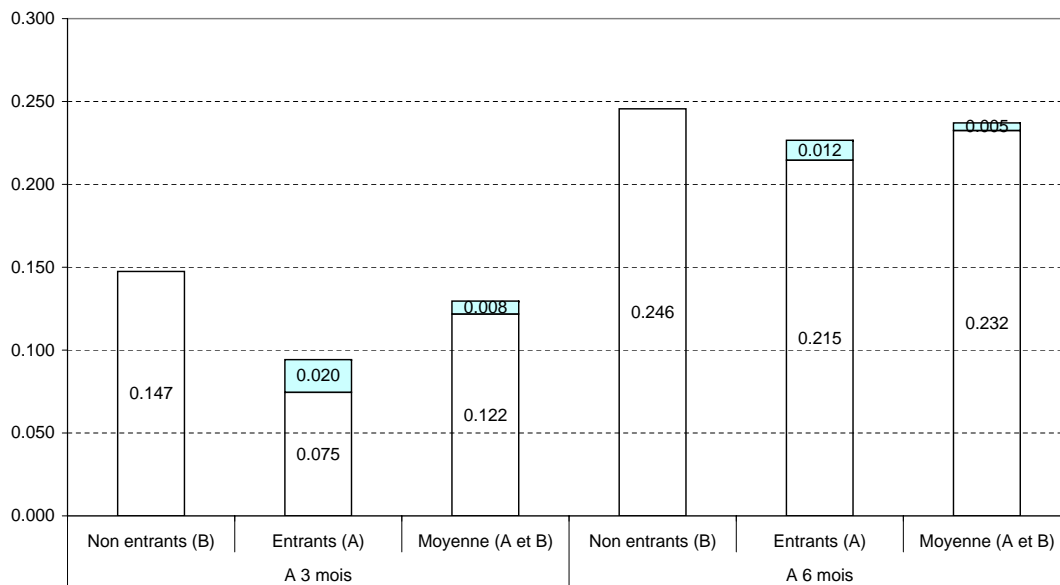


FIG. 1 – Taux de sortie vers l'emploi observés et contrefactuels dans le flux indemnisable OPP - zone commune OPP et CVE

La partie droite du graphique concerne l'effet à 6 mois - la partie gauche l'effet à 3 mois.

Le premier bâton indique le taux de retour à l'emploi des DE orientés vers le programme et qui n'y sont pas rentrés.

Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l'emploi observé.

Le troisième bâton indique le taux de retour à l'emploi de la population orientée vers le programme en l'absence de participation (partie inférieure) et totale (y compris partie supérieure)

Effet de CVE - flux indemnisable, zones communes CVE-OPP

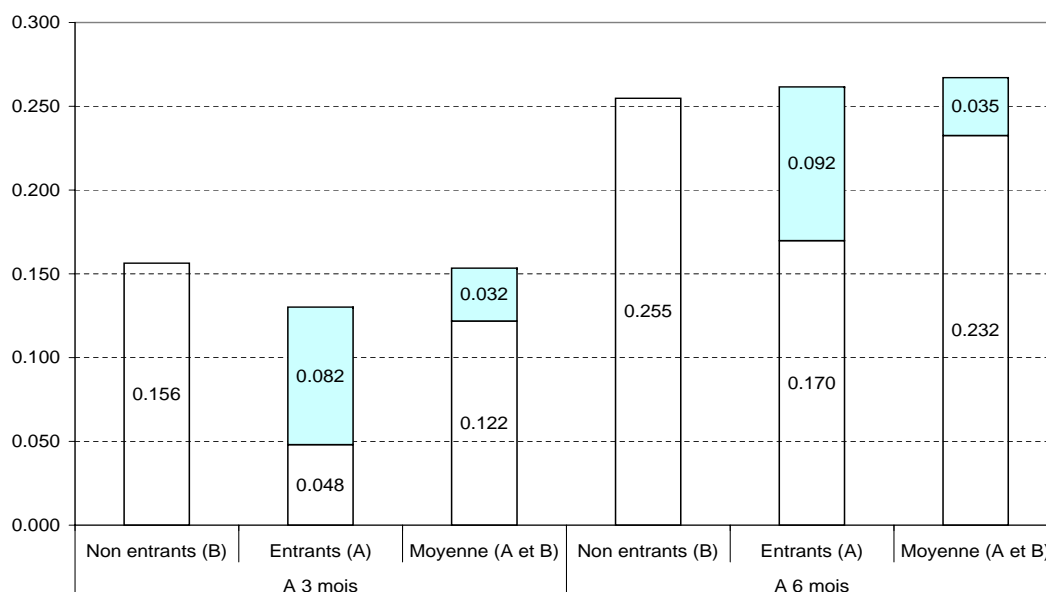


FIG. 2 – Taux de sortie vers l’emploi observés et contrefactuels dans le flux indemnisable CVE - zone commune OPP et CVE

La partie droite du graphique concerne l’effet à 6 mois - la partie gauche l’effet à 3 mois

Le premier bâton indique le taux de retour à l’emploi des DE orientés vers le programme et qui n’y sont pas rentrés... Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l’emploi observé.

Le troisième bâton indique le taux de retour à l’emploi de la population orientée vers le programme en l’absence de participation (partie inférieure) et totale (y compris partie supérieure)

Effet d'OPP - zones OPP seulement

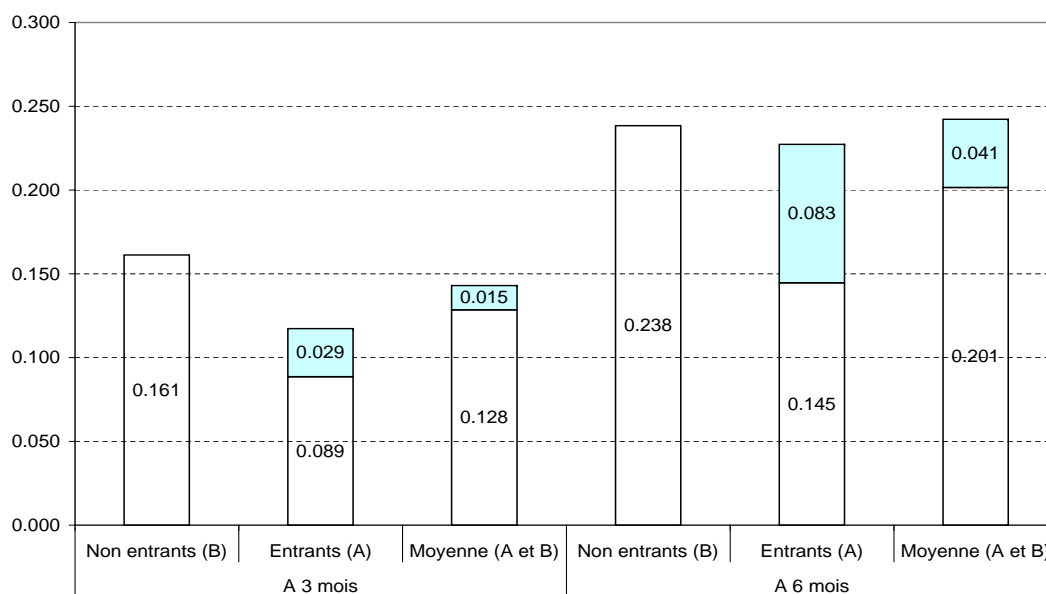


FIG. 3 – Taux de sortie vers l'emploi observés et contrefactuels dans le flux indemnisable OPP - zone OPP seul

La partie droite du graphique concerne l'effet à 6 mois - la partie gauche l'effet à 3 mois

Le premier bâton indique le taux de retour à l'emploi des DE orientés vers le programme et qui n'y sont pas rentrés.

Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l'emploi observé.

Le troisième bâton indique le taux de retour à l'emploi de la population orientée vers le programme en l'absence de participation (partie inférieure) et totale (y compris partie supérieure)

Effet de CVE - flux indemnisable, zones CVE seulement

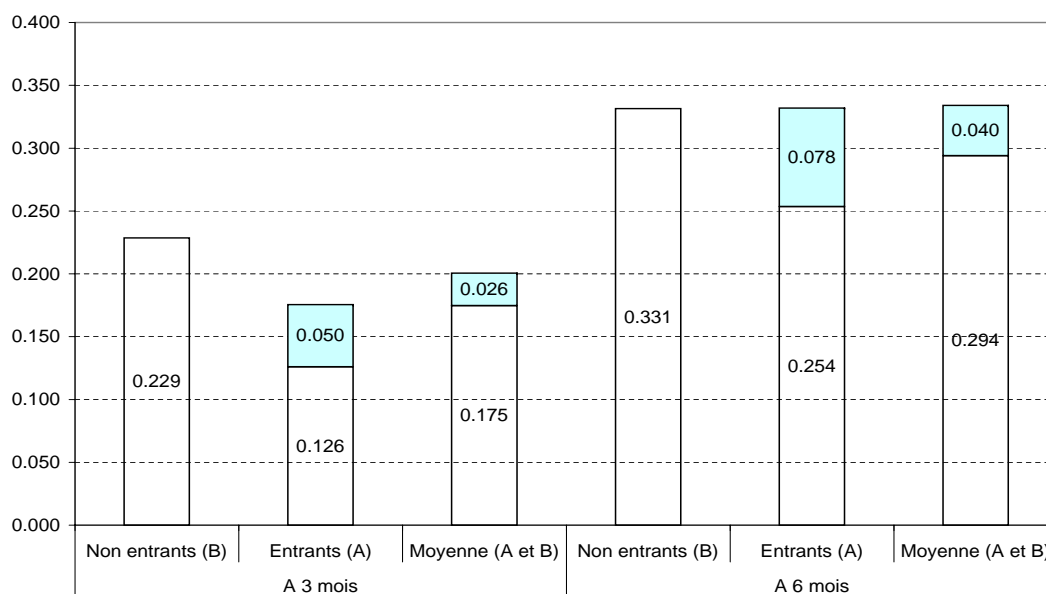


FIG. 4 – Taux de sortie vers l’emploi observés et contrefactuels dans le flux indemnisable CVE - zone CVE seul

La partie droite du graphique concerne l’effet à 6 mois - la partie gauche l’effet à 3 mois

Le premier bâton indique le taux de retour à l’emploi des DE orientés vers le programme et qui n’y sont pas rentrés.

Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l’emploi observé.

Le troisième bâton indique le taux de retour à l’emploi de la population orientée vers le programme en l’absence de participation (partie inférieure) et totale (y compris partie supérieure)

Effet de CVE dans le flux non indemnisable

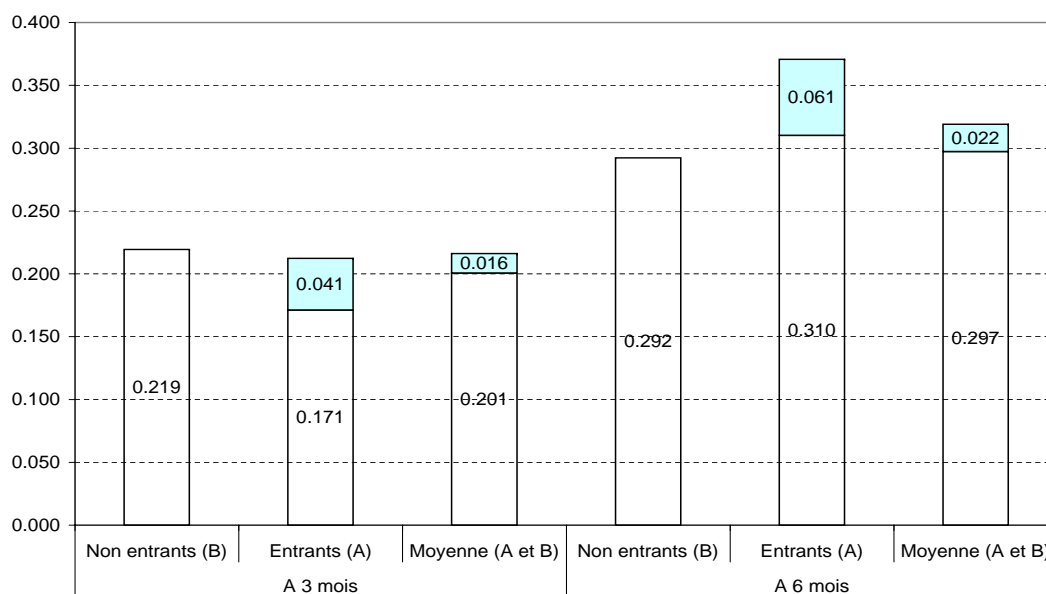


FIG. 5 – Taux de sortie vers l'emploi observés et contrefactuels dans le flux indemnisable CVE - Flux Non Indemnisable

La partie droite du graphique concerne l'effet à 6 mois - la partie gauche l'effet à 3 mois

Le premier bâton indique le taux de retour à l'emploi des DE orientés vers le programme et qui n'y sont pas rentrés.

Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l'emploi observé.

Le troisième bâton indique le taux de retour à l'emploi de la population orientée vers le programme en l'absence de participation (partie inférieure) et totale (y compris partie supérieure)

Effet de CVE dans le stock

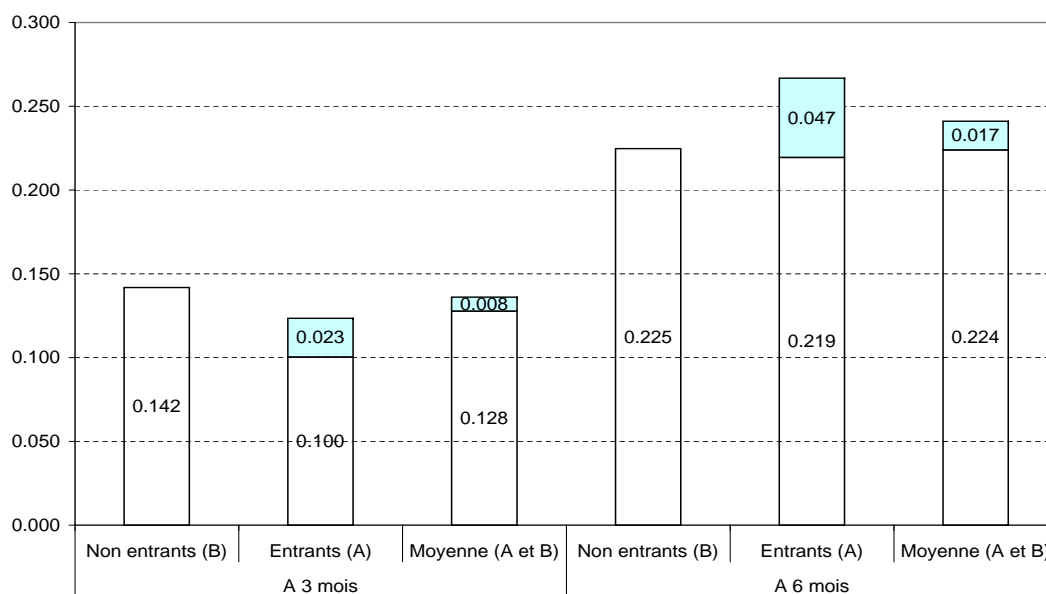


FIG. 6 – Taux de sortie vers l'emploi observés et contrefactuels dans le flux indemnisable CVE - Stock

La partie droite du graphique concerne l'effet à 6 mois - la partie gauche l'effet à 3 mois

Le premier bâton indique le taux de retour à l'emploi des DE orientés vers le programme et qui n'y sont pas rentrés.

Le deuxième concerne les DE orientés vers le programme et participant au programme. La partie inférieure est le contrefactuel, la partie supérieure la plus-value et la somme des deux le taux de retour à l'emploi observé.

Le troisième bâton indique le taux de retour à l'emploi de la population orientée vers le programme en l'absence de participation (partie inférieure) et totale (y compris partie supérieure)

entre entrants (*compliers*) et non entrants est du même ordre pour CVE et OPP. On observe également, à l’horizon de 6 mois, une plus-value comparable des deux accompagnements. Cette plus-value permet aux *compliers* de *rattraper* leur retard initial. A nouveau, il faut prendre garde à l’incertitude statistique qui subsiste (voir les intervalles de confiance au bas du tableau 9).

On peut noter au passage la bonne illustration du biais qu’il y aurait eu si on avait cherché à déduire l’effet de l’accompagnement de la simple comparaison entre bénéficiaires (entrants) et non bénéficiaires (non entrants). En comparant la colonne de gauche et celle du milieu (effet de l’accompagnement inclus), on aurait en effet cru à tort que l’effet des accompagnements CVE et OPP est à peu près nul. L’effet semble au contraire bien positif ; mais il bénéficie à des demandeurs d’emploi qui partaient de situations défavorables, et qui rattrapent ainsi leur retard.

La situation est assez différente pour les demandeurs d’emploi du flux non indemnisable et du stock. Ceux qui entrent effectivement dans un accompagnement CVE ont des chances de sortie vers l’emploi similaires à ceux qui n’entrent pas dans CVE (31% dans le flux non indemnisable, 22% dans le stock). La plus-value de l’accompagnement CVE leur permet au final d’obtenir des taux de sortie vers l’emploi supérieurs (37% dans le flux non indemnisable, 27% dans le stock). Les mécanismes de sélection qui gouvernent l’entrée dans le dispositif semblent donc être différents de ceux à l’oeuvre dans le flux indemnisable.

En résumé, cette analyse des contrefactuels, même si elle doit être envisagée avec prudence (les intervalles de confiance sont larges), suggère que les publics effectivement touchés par les accompagnements renforcés diffèrent d’un programme à l’autre, et d’une population à l’autre. Par ailleurs, elle permet de mieux évaluer l’ampleur des effets, en les rapportant à la situation qu’auraient connue les bénéficiaires s’ils n’avaient pas été accompagnés. Cela suggère que les effets obtenus sont relativement importants : à 6 mois, dans la plupart des cas (et toujours avec une marge d’incertitude importante), l’accompagnement renforcé a pour effet de multiplier les taux de sortie vers l’emploi par un facteur de 1.2 à 1.3.

5 Annexe : Estimateurs considérés

On présente ici les paramètres estimés et les estimateurs utilisés. Ces éléments font référence aux travaux théoriques développés récemment en économétrie pour effectuer des évaluations d'impact de programmes.

5.1 Le cadre d'analyse standard

On s'intéresse à une variable d'output y , comme par exemple la sortie vers l'emploi à une date donnée. Pour définir la plus-value sur la variable y associée à la participation au programme, on introduit deux variables de résultat (ou output) potentiel : $y(0)$ et $y(1)$. Il s'agit des outputs individuels sous chacune des alternatives correspondant à la participation ou non au programme. $y(0)$ est l'output si on ne participe pas au programme, $y(1)$ est l'output lorsque l'on y participe. Ces deux outputs existent pour chaque individu indépendamment de la participation ou non au programme. On définit l'effet causal de la participation comme la différence entre ces deux outputs :

$$c = y(1) - y(0)$$

Cet effet causal ainsi défini a deux caractéristiques centrales :

1. Il est individuel. Il n'y a pas a priori d'homogénéité de l'effet du programme dans la population. Il y a une distribution de l'effet du programme dans la population. Pour certains l'effet est peut être important et pour d'autres plus faible.
2. L'effet causal est inobservable. On n'observe jamais que la situation sous l'une ou l'autre des alternatives de participation au programme. Pour les bénéficiaires, on observe $y(1)$ mais pas $y(0)$ et pour les non-bénéficiaires on observe $y(0)$ mais pas $y(1)$. La difficulté principale de l'évaluation consiste à reconstituer cette situation inobservée.

Les méthodes d'évaluation standards pour évaluer l'effet d'un programme reposent sur des hypothèses fortes nécessaires à la reconstitution de cette situation inobservée. La fiabilité des résultats se mesure de ce fait à l'aune de la vraisemblance de ces hypothèses.

A contrario, le cadre expérimental qui a été développé permet de reconstituer cette situation alternative sans faire d'hypothèses fortes. Il permet en

particulier d'identifier l'effet moyen du programme sur une sous-population particulière : ceux qui se conforment à l'affectation initiale. Il s'agit de ceux qui entrent dans le dispositif lorsqu'on le leur propose et qui n'y seraient pas entrés si on ne le leur avait pas proposé.

Le cadre expérimental constitué est basé sur le fait que la population éligible initiale est ventilée au hasard en deux sous-populations repérées par une variable $z \in \{0, 1\}$. La population $z = 1$ se voit proposer d'entrer dans le programme. La population $z = 0$ ne reçoit pas cette incitation. Il s'agit par exemple de la ventilation des demandeurs d'emploi du parcours 3 du flux non indemnisable, et on les ventile en deux sous-population à l'aide de l'Outil de constitution des cohortes (OCC). La décision d'entrer dans le programme est notée $T \in \{0, 1\}$.

L'intérêt de cette ventilation initiale au hasard est de produire deux sous-populations statistiquement identiques, c'est-à-dire deux sous-populations dans lesquelles n'importe quelle variable est distribuée de façon identique : il y aura la même proportion d'individus aux cheveux châtain, la même proportion d'individus ayant un diplôme donné, la même proportion d'individus prêts à tout pour reprendre un emploi, etc. La seule différence est liée au fait que l'exposition aux programmes a été différente. Lorsqu'on mesurera des résultats différents pour ces deux populations, la différence ne pourra être attribuée qu'à cette différence d'exposition aux programmes. Toutes les autres sources de différences sont neutralisées par le tirage au sort.

On introduit deux comportements potentiels d'entrée dans le dispositif associés aux valeurs de z . On a ainsi $T(1)$ et $T(0)$. $T(1)$ représente le comportement d'entrée dans le dispositif lorsque l'on est dans la cohorte affectée au programme. Comme la participation est libre, une partie des individus va accepter d'entrer, une autre n'entrera pas. C'est cette décision qui est retracée par la variable $T(1)$. Lorsque le tirage OCC donne $z = 0$, l'individu est affecté au groupe de contrôle et normalement n'entre pas dans le dispositif. C'est cette situation qui est retracée dans la variable $T(0)$. On a normalement $T(0) = 0$ pour tout le monde.

L'orientation aléatoire implique la relation d'indépendance

$$y(0), y(1), T(0), T(1) \perp z$$

L'estimateur de référence considéré est l'estimateur dit de Wald qui compare la situation moyenne des individus affectés au groupe $z = 1$ avec la

situation moyenne des individus affectés au groupe $z = 0$ et normalisé pour tenir compte des différences d'entrée dans le dispositif entre les deux populations :

$$\hat{b}_W = \frac{\bar{y}^{z=1} - \bar{y}^{z=0}}{\bar{T}^{z=1} - \bar{T}^{z=0}}$$

Cet estimateur identifie le paramètre

$$\lim \hat{b}_W = \frac{E(y|z=1) - E(y|z=0)}{E(T|z=1) - E(T|z=0)}$$

Ce paramètre est aussi le paramètre identifié par le système de conditions d'orthogonalité :

$$E\left(\begin{pmatrix} 1 \\ z \end{pmatrix} (y - a - bT)\right) = 0 \quad (1)$$

Proposition 1 *Le paramètre b identifié par le système (1) s'exprime à partir de la distribution des outputs potentiels comme*

$$b = \frac{E(y(1) - y(0))(T(1) - T(0))}{E(T(1) - T(0))} \quad (2)$$

Lorsqu'il y a monotonie, c'est-à-dire lorsque pour tout individu on a $T(1) \geq T(0)$ ce paramètre identifie l'effet moyen du programme sur une sous-population particulière appelée compliers :

$$b = E(y(1) - y(0)) | (T(1) - T(0) = 1) \quad (3)$$

Dans le cas actuel pour lequel il n'est pas possible pour un individu affecté au parcours classique d'entrer dans le programme, on a nécessairement monotonie puisque $T(0) = 0$. Les compliers représentent simplement les individus décidant d'entrer dans le programme lorsqu'on le leur propose : $T(1) = 1$.

La décision observée d'entrer de l'individu est simplement $T = T(0)(1 - z) + T(1)z$. De même l'output observé y est simplement $y = y(1)T + y(0)(1 - T)$. On peut écrire ces deux quantités en fonction seulement des outputs et participations potentielles :

$$y = y(0) + T(0)(y(1) - y(0)) + (y(1) - y(0))(T(1) - T(0))z \quad (4)$$

$$= g_0 + g_1 z \quad (5)$$

$$T = T(0) + (T(1) - T(0))z \quad (6)$$

$$= h_0 + h_1 z \quad (7)$$

En considérant les expressions des équations (5) et (7), on a

$$(y - a - bT) = (g_0 - a - bh_0 + (g_1 - bh_1)) = f_0(\theta) + zf_1(\theta) \quad (8)$$

où $\theta = (a, b)$, $f_0(\theta) = g_0 - a - bh_0$ et $f_1(\theta) = g_1 - bh_1$

Les conditions d'orthogonalité (1) se réécrivent donc

$$E \left(\begin{pmatrix} 1 \\ z \end{pmatrix} (f_0(\theta) + zf_1(\theta)) \right) = E \begin{pmatrix} f_0 + zf_1 \\ zf_0 + zf_1 \end{pmatrix} = 0$$

Compte tenu de la relation d'indépendance, on a donc

$$\begin{pmatrix} E(f_0) + E(z)E(f_1) \\ E(f_0)E(z) + E(f_1)E(z) \end{pmatrix} = 0$$

Comme $E(z) \neq 1$ on voit que les conditions d'orthogonalité identifient le paramètre θ tel que

$$E(f_0(\theta)) = E(g_0 - a - bh_0) = 0 \quad (9)$$

$$E(f_1(\theta)) = E(g_1 - bh_1) = 0 \quad (10)$$

On en déduit en particulier que le paramètre b identifie

$$b = \frac{E(g_1)}{E(h_1)} = \frac{E(y(1) - y(0))(T(1) - T(0))}{E(T(1) - T(0))} \quad (11)$$

On trouve donc sans surprise que lorsqu'il y a monotonie, les conditions d'orthogonalité précédentes identifient le paramètre LATE : Local Average Treatment Effect.

5.2 Indépendance conditionnelle à des observables X_c

En pratique les probabilités d'envoi dans les programmes ont fluctué d'une ALE à l'autre et d'une période à l'autre. Cette fluctuation était nécessaire

pour satisfaire les objectifs quantitatifs d'envoi aux OPP. Il en résulte qu'on n'a pas la condition d'indépendance précédente sur toute la population dans son ensemble, mais seulement conditionnellement aux zones et périodes de constance des probabilités d'envoi dans les programmes.

On considère maintenant que l'on a $y(0), y(1), T(0), T(1) \perp z | X_c$ et non plus $y(0), y(1), T(0), T(1) \perp z$. On a le résultat suivant :

Proposition 2 *Lorsque l'on a indépendance conditionnelle à des observables $y(0), y(1), T(0), T(1) \perp z | X_c$. Les conditions d'orthogonalité (1), convenablement pondérées :*

$$E \left(w(X_c, z) \begin{pmatrix} 1 \\ z \end{pmatrix} (y - a - bT) \right) = 0 \quad (12)$$

avec

$$w(X_c, z) = \left(\frac{C_1}{P(z=1|X_c)} \right)^z \left(\frac{C_0}{1 - P(z=1|X_c)} \right)^{1-z} \quad (13)$$

identifient le même paramètre que précédemment.

$$b = \frac{E(g_1)}{E(h_1)} = \frac{E(y(1) - y(0))(T(1) - T(0))}{E(T(1) - T(0))} \quad (14)$$

Compte tenu de l'expression du résidu de l'équation (8), les conditions précédentes (12) s'écrivent

$$\begin{pmatrix} E(w(X_c, z) f_0) + E(w(X_c, z) z f_1) \\ E(w(X_c, z) f_0 z) + E(w(X_c, z) f_1 z) \end{pmatrix} = 0$$

On a compte tenu de la relation d'indépendance

$$\begin{aligned} E(w(X_c, z) f_0) &= E(E(w(X_c, z) f_0 | X_c)) \\ &= E(E(w(X_c, z) | X_c) E(f_0 | X_c)) \end{aligned}$$

De même pour $f = f_0$ ou $f = f_1$ on a :

$$\begin{aligned} E(w(X_c, z) z f) &= E(E(w(X_c, z) z f | X_c)) \\ &= E(E(w(X_c, z) z | X_c) E(f | X_c)) \end{aligned}$$

On voit donc que pour que le paramètre identifié par les conditions d'orthogonalité soit le même, il suffit que l'on ait $E(w(X_c, z)|X_c) = A$ et $E(w(X_c, z)z|X_c) = B$ avec $A \neq B$.

En effet dans ce cas, les équations précédentes se résument à

$$\begin{pmatrix} A & B \\ B & B \end{pmatrix} \begin{pmatrix} E(f_0) \\ E(f_1) \end{pmatrix} = 0$$

$$\begin{aligned} E(w(X_c, z)|X_c) &= A = w(X_c, 1)P(z=1|X_c) + w(X_c, 0)P(z=0|X_c) \\ E(w(X_c, z)z|X_c) &= B = w(X_c, 1)P(z=1|X_c) \end{aligned}$$

Il en résulte que tout système de poids de la forme

$$w(X_c, z_0) = \frac{C_{z_0}}{P(z=z_0|X_c)} \quad (15)$$

permet d'identifier les paramètres annulant $E(f_0)$ et $E(f_1)$. En particulier pour $A = 1$ et $B = 2$, on a le système de poids proposé.

5.2.1 Forme de l'estimateur et comparaison avec l'estimateur à variables de contrôle

L'estimateur obtenu par solution du système (12) généralise l'estimateur de Wald. On trouve facilement son expression :

$$\hat{b}_W = \left(\frac{\overline{y/P(z=1|X_c)}^{z=1}}{\overline{1/P(z=1|X_c)}^{z=1}} - \frac{\overline{y/P(z=0|X_c)}^{z=0}}{\overline{1/P(z=0|X_c)}^{z=0}} \right) / \left(\frac{\overline{T/P(z=1|X_c)}^{z=1}}{\overline{1/P(z=1|X_c)}^{z=1}} - \frac{\overline{T/P(z=0|X_c)}^{z=0}}{\overline{1/P(z=0|X_c)}^{z=0}} \right) \quad (16)$$

Lorsque les variables X_c forment une partition de la population, les poids peuvent être estimés simplement par

$$\hat{P}(z=1|X_c) = \frac{N_{X_c,1}}{N_{X_c}}$$

L'utilisation de ces poids empiriques conduit à une expression simple de l'estimateur

$$\hat{b}_W = \frac{\sum_{X_c} \pi_{X_c} (\bar{y}^{X_c, z=1} - \bar{y}^{X_c, z=0})}{\sum_{X_c} \pi_{X_c} (\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0})} \quad (17)$$

où $\pi_{X_c} = N_{X_c}/N$ mesure le poids de la cellule X_c .
Cet estimateur peut être comparé utilement à l'estimateur à variable instrumentale \hat{b}_c du modèle

$$y = X_c a_c + b_c T + u \quad (18)$$

avec pour variables instrumentales X_c et z , défini comme annulant les conditions d'orthogonalité :

$$E \left(\begin{pmatrix} X'_c \\ z \end{pmatrix} (y - X_c a_c - b_c T) \right) = 0 \quad (19)$$

On trouve aisément l'expression de cet estimateur :

$$\hat{b}_c = \frac{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} (\bar{y}^{X_c, z=1} - \bar{y}^{X_c, z=0})}{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} (\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0})} \quad (20)$$

$\hat{p}_{1,c} = N_{X_c,1}/N_{X_c}$ représente la probabilité empirique d'affectation au programme et $\hat{p}_{0,c} = N_{X_c,0}/N_{X_c}$ représente la probabilité empirique d'affectation au groupe de contrôle.

Par rapport à l'estimateur précédent, cet estimateur donne donc d'autant plus de poids à une cellule que sa repartition entre affectation au traitement et aux contrôle est équilibrée.

Un autre estimateur possible est l'estimateur à variable instrumentale \hat{b}_c du modèle

$$y = X_c a_c + b_c T + u \quad (21)$$

avec pour variables instrumentales X_c et $z \times X_c$, défini comme annulant les conditions d'orthogonalité :

$$E \left(\begin{pmatrix} X'_c \\ z X'_c \end{pmatrix} (y - X_c a_c - b_c T) \right) = 0 \quad (22)$$

L'expression de cet estimateur est la suivante :

$$\hat{b}_c = \frac{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} (\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0}) (\bar{y}^{X_c, z=1} - \bar{y}^{X_c, z=0})}{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} (\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0})^2} \quad (23)$$

Cet estimateur est suridentifié. On peut alors faire le test de Sargan. Un tel test correspond alors à l'hypothèse d'homogénéité des estimateurs de Wald \hat{b}_{X_c} sur chacune des population X_c :

$$\hat{b}_{X_c} = \frac{\bar{y}^{X_c, z=1} - \bar{y}^{X_c, z=0}}{\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0}} \quad (24)$$

L'hypothèse nulle associée au test de Sargan est alors $H_0 : \exists b \mid \hat{b}_{X_c} = b$

L'estimateur (23) peut se réécrire facilement à partir des estimateurs \hat{b}_{X_c} . On a en fait :

$$\hat{b}_c = \frac{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} \left(\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0} \right)^2 \hat{b}_{X_c}}{\sum_{X_c} \pi_{X_c} \hat{p}_{1,c} \hat{p}_{0,c} \left(\bar{T}^{X_c, z=1} - \bar{T}^{X_c, z=0} \right)^2} \quad (25)$$

On voit que par rapport aux autres estimateurs, on pondère d'avantage les cellules dans lesquelles on a une différence de take-up importante. (Notons que dans le cas où on considère l'effet hétérogène, on a des paramètres différents car les populations sélectionnées sont différentes).

Compte tenu de l'expression de la variance de cet estimateur, on peut montrer en outre qu'il s'écrit

$$\hat{b}_c = \sum_{X_c} \hat{b}_{X_c} \frac{1/V_{X_c}}{\sum_{X_c} 1/V_{X_c}} \quad (26)$$

où V_{X_c} est sa variance. Il s'agit donc de l'estimateur des Moindres Carrés Asymptotiques du modèle imposant l'homogénéité du traitement

$$b_{X_c} = b \quad (27)$$

où b_{X_c} est la valeur du paramètre dans la cellule X_c .

Le tableau 10 présente les résultats obtenu pour l'évaluation à 6 mois sur les six premières cohortes. Il montre que les estimations ne sont pas très sensibles à la méthode utilisée. On remarque comme on s'y attend que l'estimateur le plus efficace est celui qui est obtenu en combinant les indicatrice d'affectation avec les zones de probabilité constante. Toutefois, ce surcroît d'information ne conduit qu'à des modifications mineures des paramètres. On note en, particulier la grande proximité avec l'estimateur à variables instrumentales simple. On pourrait formellement tester la compatibilité en effectuant un test de Sargan. Ceci n'est pas utile ici en raison de la très grande

proximité des résultats. Les résultats ne sont pas sensibles à la méthode utilisée, et on en déduit en outre qu'il y a homogénéité des paramètres sur l'ensemble des zones à probabilités constantes.

Evaluation CVE et OPP dans le flux indemnisable^a

	Effet de l'accompagnement OPP			Effet de l'accompagnement CVE						
	Est.	Std.	Effectif	Est.	Std.	Effectif				
	Borne			Borne						
	Inf.	Sup.	Effectif	Inf.	Sup.	Effectif				
Estimateur 1 ^b	1,2	(3,2)	-5,2	7,6	23213	9,2	(4,2)	0,9	17,5	23213
Estimateur 2 ^c	1,2	(3,0)	-4,8	7,2	23213	9,7	(3,9)	1,9	17,5	23213
Estimateur 3 ^d	0,0	(3,3)	-6,5	6,5	23213	6,4	(4,4)	-2,4	15,1	23213

Evaluation CVE dans le flux non indemnisable et le stock

	Flux non indemnisable			Stock						
	Est.	Std.	Effectif	Est.	Std.	Effectif				
	Borne			Borne						
	Inf.	Sup.	Effectif	Inf.	Sup.	Effectif				
Estimateur 1	6,1	(2,3)	1,5	10,6	27978	4,7	(1,8)	1,0	8,4	58472
Estimateur 2	5,8	(2,2)	1,3	10,3	27978	4,5	(1,8)	0,8	8,1	58472
Estimateur 3	6,3	(2,3)	1,7	10,9	27978	4,3	(2,0)	0,3	8,2	58472

^aALE où CVE et OPP sont présents conjointement

^bEstimateur avec indicatrices de proba constantes; instruments : ZOPP et ZCVE

^cEstimateur avec indicatrices de proba constantes; instruments : ZOPP et ZCVE, et interactions avec indicatrices de proba constantes

^dEstimateur avec pondérations

TAB. 10 – Estimateurs alternatifs. Effet à 6 mois sur les six premières cohortes

5.3 Introduction de variables de conditionnement X_0 et indépendance conditionnelle à X_c

On considère maintenant la question de l'introduction de caractéristiques observables des individus sous la forme de variables X_0 . Ceci est intéressant car cela permettra d'examiner l'hétérogénéité de l'effet. On considère la situation dans laquelle on a randomisation de z conditionnellement à des variables explicatives X_c . On a donc ici $X_0, y(0), y(1), T(0), T(1) \perp z | X_c$. On considère les conditions d'orthogonalité

$$E \left(w(X_c, z) \begin{pmatrix} X'_0 \\ zX'_0 \end{pmatrix} (y - X_0a - TX_0b) \right) = 0 \quad (28)$$

avec $w(X_c, z_0) = 1/P(z = z_0 | X_c)$.

Proposition 3 *Sous l'hypothèse d'indépendance $X_0, y(0), y(1), T(0), T(1) \perp z | X_c$ et l'hypothèse de monotonie $T(1) \geq T(0)$, les conditions d'orthogonalité 28 identifient le vecteur des coefficients de la projection orthogonal de l'effet causal du programme sur les caractéristiques X_0 pour la population des compliers.*

$$b = E(X'_0 X_0 | T(1) - T(0) = 1)^{-1} E(X'_0 (y(1) - y(0)) | T(1) - T(0) = 1) \quad (29)$$

Là aussi, on peut réécrire l'expression du résidu $(y - X_0a - TX_0b)$ en utilisant les équations 5 et 7, on a

$$(y - X_0a - X_0bT) = (g_0 - X_0a - X_0bh_0 + z(g_1 - X_0bh_1)) = f_0(\theta) + zf_1(\theta) \quad (30)$$

où $\theta = (a, b)$, $f_0(\theta) = g_0 - X_0a - X_0bh_0$ et $f_1(\theta) = g_1 - X_0bh_1$

Les conditions d'orthogonalité 28 se réécrivent donc

$$E \left(w(X_c, z) \begin{pmatrix} X'_0 \\ zX'_0 \end{pmatrix} (f_0(\theta) + zf_1(\theta)) \right) = E \left(w(X_c, z) \begin{pmatrix} X'_0 f_0 + zX'_0 f_1 \\ zX'_0 f_0 + zX'_0 f_1 \end{pmatrix} \right) = 0$$

Compte tenu de la relation d'indépendance, les conditions 5.4.1 se réécrivent comme

$$\begin{pmatrix} 2E(X'_0 f_0) + E(X'_0 f_1) \\ E(X'_0 f_0) + E(X'_0 f_1) \end{pmatrix} = 0$$

Ceci provient du fait qu'en raison des poids choisis on a $E(w(X_c, z)) = 2$ et $E(zw(X_c, z)) = 1$. On voit donc que les conditions d'orthogonalité identifient le paramètre θ tel que

$$E(X'_0 f_0(\theta)) = E(X'_0 g_0 - X_0 a - X_0 b h_0) = 0 \quad (31)$$

$$E(X'_0 f_1(\theta)) = E(X'_0 g_1 - X'_0 X_0 b h_1) = 0 \quad (32)$$

On en déduit en particulier que le paramètre b identifie

$$b = E(X'_0 X_0 h_1)^{-1} E(X'_0 g_1) \quad (33)$$

Comme $h_1 = T(1) - T(0)$ et $g_1 = (T(1) - T(0))(y(1) - y(0))$ Sous l'hypothèse de monotonie, on voit que l'on a

$$b = E(X'_0 X_0 | T(1) - T(0) = 1)^{-1} E(X'_0 (y(1) - y(0)) | T(1) - T(0) = 1) \quad (34)$$

et on voit que le paramètre b s'interprète naturellement comme le vecteur des coefficients de la projection orthogonale de l'effet du programme sur les caractéristiques X_0 . Remarquons que si à partir de ces estimateurs, on voulait remonter à un paramètre global, il faudrait réintégrer par rapport à la distribution des variables de conditionnement sachant que l'on est un complier.

Le tableau 11 présente une tentative de détection d'une source d'hétérogénéité de l'effet entre individus, lorsque l'on s'intéresse à la sortie à 6 mois. On examine l'hétérogénéité de l'effet en considérant la projection de l'effet sur certaines caractéristiques observables. On introduit en particulier le niveau d'éducation (diplôme supérieur au Bac), l'âge (variable centrée réduite) et on considère enfin le sexe. Le tableau présente dans son panel supérieur l'effet suivant la caractéristique considérée et dans sa partie inférieure l'effet de la caractéristique considérée sur la sortie. On voit qu'on ne détecte pas facilement d'hétérogénéité de l'effet bien que les caractéristiques retenues jouent fortement sur la sortie du chômage vers l'emploi. Dans l'ensemble les résultats semblent plus élevés pour les hommes, les jeunes et les non qualifiés, que ce soit pour le programme CVE ou pour le programme OPP. Ces effets ne sont toutefois pas significatifs sauf pour le programme CVE dans le stock où on détecte un effet significatif pour le sexe et l'âge. On note en particulier que l'effet pour CVE dans le stock serait nul pour les femmes, alors qu'il est significativement positif pour les hommes.

	FI ^a		FNI	Stock
	OPP	CVE	CVE	CVE
Effet de l'accompagnement renforcé				
Homme d'âge moyen, > bac	0.0447	0.0505	0.0626	0.0730**
	(0.0574)	(0.0804)	(0.0443)	(0.0336)
Effet additionnel < bac	0.0041	0.1338	0.0052	0.0261
	(0.0622)	(0.0857)	(0.0459)	(0.0355)
Effet additionnel âge normalisé	-0.0358	-0.0351	-0.0137	-0.0373**
	(0.0306)	(0.0429)	(0.0235)	(0.0181)
Effet additionnel femmes	-0.0505	-0.0518	-0.0162	-0.0732**
	(0.0601)	(0.0853)	(0.0457)	(0.0347)
Effet des caractéristiques sur le taux de sortie				
< bac	-0.0618***		-0.0881***	-0.0509***
	(0.0208)		(0.0117)	(0.0088)
Age (normalisé)	-0.0046***		-0.0054***	-0.0063***
	(0.0009)		(0.0005)	(0.0004)
Femme	-0.0328		-0.0511***	-0.0364***
	(0.0200)		(0.0114)	(0.0084)
Nobs	23213		27978	58472

^azones communes

TAB. 11 – Estimateurs alternatifs. Effet à 6 mois sur les six premières cohortes

5.4 Sondage, non-réponse et variance des estimateurs

5.4.1 Sondage et non-réponse

Dans la situation où on a $X_0, y(0), y(1), T(0), T(1) \perp z | X_c$, les conditions d'orthogonalité que l'on considèrerait avec information complète sur les variables d'output seraient

$$E \left(w(X_c, z_0) \begin{pmatrix} X'_0 \\ zX'_0 \end{pmatrix} (y - X_0a - TX_0b) \right) = 0$$

avec $w(X_c, z_0) = 1/P(z = z_0 | X_c)$.

Ces conditions d'orthogonalité s'écrivent donc généralement

$$E(h(y, T, z, X_0, X_c, \theta_0)) = 0 \quad (35)$$

les observations sur la variable y ne sont disponibles que pour les individus sortis et ayant renseigné leur Déclaration de Situation Mensuelle. Pour ceux sortis sans renseigner cette déclaration mensuelle on effectue aussi un sondage et on recherche l'information manquante. D'une façon générale, on effectue un sondage conditionnellement à x_n , la variable indiquant la sortie sans réponse, z et T : $s(y, z, T)$, avec $E(s(x_n, z, T)) = \lambda(x_n, z, T)$. Pour $x_n = 0$ on a l'information, si bien que l'on a $s = 1$.

Les conditions d'orthogonalité que l'on considère sont donc

$$E \left(\frac{s(x_n, z, T)}{\lambda(x_n, z, T)} h(y, T, z, X_0, X_c, \theta_0) \right) = 0$$

et compte tenu du caractère aléatoire du sondage conditionnellement à x_n, z, T , c'est bien le même paramètre que l'on identifie.

Il y a aussi le problème des non réponses. Il y a donc une variable de réponse R indiquant si l'individu répond. On considère que cette variable de réponse est indépendante du résidu $h(y, T, z, X_0, X_c, \theta)$ conditionnellement à x_r, z, T, X_r . Sous cette hypothèse les conditions d'orthogonalité

$$E \left(\frac{r(x_n, z, T)}{\rho(x_n, z, T, X_r)} h(y, T, z, X_0, X_c, \theta_0) \right) = 0$$

avec $r(x_n, z, T) = s(x_n, z, T) R$ une variable indiquant si la variable d'output est renseignée, ce qui résulte du sondage qui est aléatoire et du comportement de réponse qui est supposé aléatoire et $\rho(x_n, z, T, X_r) = \lambda(x_n, z, T) \times (R = 1 | x_n, z, T, X_r)$ identifie le même paramètre que les conditions 35.

Finalement les conditions d'orthogonalité utilisée s'écrivent :

$$E \left(\omega \begin{pmatrix} X' \\ zX' \end{pmatrix} (y - Xa - TXb) \right) = 0$$

où ω est un poids défini par

$$\omega = \frac{s(x_n, z, T)}{\lambda(x_n, z, T)} \frac{R}{P(R=1|x_n, z, T, X_r)} \frac{1}{P(z=1|x_C)^z} \frac{1}{(1-P(z=1|x_C))^{1-z}} \quad (36)$$

5.4.2 Calcul des estimateurs et de leur variance

L'estimateur correspondant est défini comme celui annulant les conditions d'orthogonalités :

$$\overline{\omega \begin{pmatrix} X' \\ zX' \end{pmatrix} (y - X\hat{a} - TX\hat{b})} = 0$$

Ceci donne comme estimateur

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \left[\begin{array}{cc} \overline{\omega X'_0 X_0} & \overline{\omega X'_0 X_0 T} \\ \overline{\omega z X'_0 X_0} & \overline{\omega z X'_0 X_0 T} \end{array} \right]^{-1} \begin{pmatrix} \overline{\omega X'_0 y} \\ \overline{\omega z X'_0 y} \end{pmatrix}$$

En notant

$$\hat{G} = \left[\begin{array}{cc} \overline{\omega X'_0 X_0} & \overline{\omega X'_0 X_0 T} \\ \overline{\omega z X'_0 X_0} & \overline{\omega z X'_0 X_0 T} \end{array} \right]$$

et

$$G = \left[\begin{array}{cc} E(\omega X'_0 X_0) & E(\omega X'_0 X_0 T) \\ E(\omega z X'_0 X_0) & E(\omega z X'_0 X_0 T) \end{array} \right]$$

On a

$$\sqrt{N} \left(\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \right) = \hat{G}^{-1} \begin{pmatrix} \overline{\omega X'_0 u} \\ \overline{\omega z X'_0 u} \end{pmatrix}$$

avec $u = y - a - bT$

La variance de l'estimateur est de ce fait

$$V \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = G^{-1}VG'^{-1}$$

avec

$$V = \begin{bmatrix} E(\omega^2 X'_0 X_0 u^2) & E(\omega^2 X'_0 X_0 T u^2) \\ E(\omega^2 z X'_0 X_0 u^2) & E(\omega^2 z X'_0 X_0 T u^2) \end{bmatrix}$$

qui peut être estimé par

$$\hat{V} = \begin{bmatrix} \overline{\omega^2 X'_0 X_0 \hat{u}^2} & \overline{\omega^2 X'_0 X_0 T \hat{u}^2} \\ \overline{\omega^2 z X'_0 X_0 \hat{u}^2} & \overline{\omega^2 z X'_0 X_0 T \hat{u}^2} \end{bmatrix}$$

et on a donc pour estimateur convergent de la matrice de variance

$$\hat{V} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \hat{G}^{-1}\hat{V}\hat{G}'^{-1}$$

5.5 Le cas multiprogramme

On reprend le cas initial avec affectation à trois groupes : $1(Z = 0)$ le groupe de contrôle, $Z_1 = 1(Z = 1)$ le groupe affecté au programme 1 et $Z_2 = 1(Z = 2)$ le groupe affecté au programme 2. Une fois affectés, les individus peuvent se comporter de différentes façons et entrer dans les programmes ou non. On a une variable d'entrée dans les programmes T : $T = 0$ correspond à l'absence de programme, $T = 1$ correspond à l'entrée dans le programme 1 et $T = 2$ correspond à l'entrée dans le programme 2. On a aussi des variables d'output correspondant au programme reçu. $y(0)$ lorsque l'on ne suit pas de programme, $y(1) = y(0) + \Delta(1)$ lorsque l'on suit le programme 1 et $y(2) = y(0) + \Delta(2)$ lorsque l'on suit le programme 2. On a de même des variables de programme potentielles correspondant à l'affectation : $T(0)$ est la variable d'entrée dans le programme lorsque l'on est affecté au groupe $Z = 0$, $T(1)$ celle lorsque l'on est affecté au groupe $Z = 1$ et $T(2)$ celle lorsque l'on est affecté au groupe $Z = 2$.

Proposition 4 *Lorsque les schémas d'affectation au programme satisfont les contraintes $(T(2) = 1) = (T(0) = 1) \leq (T(1) = 1)$ et $(T(1) = 2) = (T(0) = 2) \leq (T(2) = 2)$, les paramètres annulant les conditions d'orthogonalité sont les coefficients de la projection linéaire des effets moyens du traitement sur les caractéristiques X_0 pour la population $(T(1) = 1) - (T(0) = 1) = 1$ en ce*

qui concerne c_1 et pour la population $(T(2) = 1) - (T(0) = 2) = 1$ en ce qui concerne c_2 .

Pour que les contraintes soient satisfaites, il faut que la population ne contienne que :

- des never takers $T(0) = 0, T(1) = 0, T(2) = 0$
- des always takers programme 1 $T(0) = 1, T(1) = 1, T(2) = 1$
- des always takers programme 2 $T(0) = 2, T(1) = 2, T(2) = 2$
- des compliers $T(0) = 0, T(1) = 1, T(2) = 2$
- des selective takers programme 1 $T(0) = 0, T(1) = 1, T(2) = 0$
- des selective takers programme 2 $T(0) = 0, T(1) = 0, T(2) = 1$

La population $(T(1) = 1) - (T(0) = 1) = 1$ est composée des compliers et des selective taker 1, la population $(T(2) = 2) - (T(0) = 2) = 1$ est composée des compliers et des selective takers 2.

On a :

$$T = T(0) + (T(1) - T(0))Z_1 + (T(2) - T(0))Z_2$$

soit

$$\begin{aligned} 1(T = 1) &= 1(T(0) = 1) + (1(T(1) = 1) - 1(T(0) = 1))Z_1 + (1(T(2) = 1) - 1(T(0) = 1))Z_2 \\ &= h_0 + h_1Z_1 + h_2Z_2 \end{aligned}$$

et

$$\begin{aligned} 1(T = 2) &= 1(T(0) = 2) + (1(T(1) = 2) - 1(T(0) = 2))Z_1 + (1(T(2) = 2) - 1(T(0) = 2))Z_2 \\ &= i_0 + i_1Z_1 + i_2Z_2 \end{aligned}$$

on a en outre pour la variable y

$$\begin{aligned} y &= y(0) + \Delta_1 1(T = 1) + \Delta_2 1(T = 2) \\ &= y(0) + \Delta_1 (h_0 + h_1Z_1 + h_2Z_2) + \Delta_2 (i_0 + i_1Z_1 + i_2Z_2) \\ &= y(0) + \Delta_1 h_0 + \Delta_2 i_0 + (\Delta_1 h_1 + \Delta_2 i_1) Z_1 + (\Delta_1 h_2 + \Delta_2 i_2) Z_2 \\ &= g_0 + g_1Z_1 + g_2Z_2 \end{aligned}$$

Si on considère maintenant les conditions d'orthogonalité

$$\left(E \begin{pmatrix} X'_0 \\ X'_0 Z_1 \\ X'_0 Z_2 \end{pmatrix} (y - X_0 b - X_0 c_1 1(T=1) - X_0 c_2 1(T=2)) \right) = 0 \quad (37)$$

soit

$$\begin{aligned} y - X_0 b &- X_0 c_1 1(T=1) - X_0 c_2 1(T=2) = g_0 + g_1 Z_1 + g_2 Z_2 - X_0 b \\ &- X_0 c_1 (h_0 + h_1 Z_1 + h_2 Z_2) - X_0 c_2 (i_0 + i_1 Z_1 + i_2 Z_2) \\ &= g_0 - X_0 b - X_0 c_1 h_0 - X_0 c_2 i_0 \\ &+ (g_1 - X_0 c_1 h_1 - X_0 c_2 i_1) Z_1 + (g_2 - X_0 c_1 h_2 Z_2 - X_0 c_2 i_2) Z_2 \\ &= f_0 + f_1 Z_1 + f_2 Z_2 \end{aligned}$$

Compte tenu des hypothèses d'indépendance, on a

$$\begin{pmatrix} E(X'_0 f_0) + E(X'_0 f_1) E(Z_1) + E(X'_0 f_2) E(Z_2) \\ E(X'_0 f_0) E(Z_1) + E(X'_0 f_1) E(Z_1) \\ E(X'_0 f_0) E(Z_2) + E(X'_0 f_2) E(Z_2) \end{pmatrix} = 0$$

On en conclut donc que l'on doit avoir

$$\begin{pmatrix} E(X'_0 f_0) \\ E(X'_0 f_1) \\ E(X'_0 f_2) \end{pmatrix} = 0$$

Ceci conduit donc en particulier aux deux équations :

$$\begin{aligned} E(X'_0 \Delta_1 h_1) + E(X'_0 \Delta_2 i_1) &= E(X'_0 X_0 h_1) c_1 + E(X'_0 X_0 i_1) c_2 \\ E(X'_0 \Delta_1 h_2) + E(X'_0 \Delta_2 i_2) &= E(X'_0 X_0 h_2) c_1 + E(X'_0 X_0 i_2) c_2 \end{aligned}$$

Le cas intéressant est celui dans lequel on a $h_2 = 0$ et $i_1 = 0$. Cela correspond à $(T(2) = 1) = (T(0) = 1)$ et $(T(1) = 2) = (T(0) = 2)$. On vérifie que ceci n'est possible que si on a des taker uniformes (choix de 0 quelle que soit l'affectation, choix de 1 quelle que soit l'affectation et choix de 2 quelle que soit l'affectation), des "compilers" : $T(0) = 0$, $T(1) = 1$ et $T(2) = 2$ et des *selective takers* $T(0) = 0$, $T(1) = 0$ et $T(2) = 2$ ou $T(0) = 0$, $T(1) = 1$ et $T(2) = 0$. Alors les paramètres s'interprètent comme précédemment comme les coefficients de la projection des effets sur les X pour les $h_1 = 1$ et $i_2 = 1$.

Proposition 5 *Soit le cas où il y a des réallocations à partir d'une situation initiale $T^*(j)$ satisfaisant les restrictions d'affectation :*

$$\begin{aligned} T(0) &= P_{00}T^*(0) + P_{01}T^*(1) + P_{02}T^*(2) \\ T(1) &= P_{10}T^*(0) + P_{11}T^*(1) + P_{12}T^*(2) \\ T(2) &= P_{20}T^*(0) + P_{21}T^*(1) + P_{22}T^*(2) \end{aligned}$$

avec P_{kj} des variables prenant la valeur 0 ou 1 et représentant la transition entre l'affectation initiale k et l'affectation finale j et satisfaisant : $P_{k0} + P_{k1} + P_{k2} = 1$. Si ces réaffectations sont purement aléatoires, elles ne perturbent pas l'identification des paramètres. Les coefficients c_1 et c_2 s'interprètent comme les coefficients des projections linéaires des effets des programmes sur les mêmes populations qu'avant, mais définies à partir des variables $T^*(j)$

On a en effet

$$\begin{aligned} h_1 &= (T(1) = 1) - (T(0) = 1) \\ &= (P_{11} - P_{01})((T^*(1) = 1) - (T^*(0) = 1)) + (P_{12} - P_{02})((T^*(2) = 1) - (T^*(0) = 1)) \\ &= (P_{11} - P_{01})((T^*(1) = 1) - (T^*(0) = 1)) \\ h_2 &= (P_{21} - P_{01})((T^*(1) = 1) - (T^*(0) = 1)) + (P_{22} - P_{02})((T^*(2) = 1) - (T^*(0) = 1)) \\ &= (P_{21} - P_{01})((T^*(1) = 1) - (T^*(0) = 1)) \\ i_1 &= (P_{11} - P_{01})((T^*(1) = 2) - (T^*(0) = 2)) + (P_{12} - P_{02})((T^*(2) = 2) - (T^*(0) = 2)) \\ &= (P_{12} - P_{02})((T^*(2) = 1) - (T^*(0) = 2)) \\ i_2 &= (P_{21} - P_{01})((T^*(1) = 2) - (T^*(0) = 2)) + (P_{22} - P_{02})((T^*(2) = 2) - (T^*(0) = 2)) \\ &= (P_{22} - P_{02})((T^*(2) = 2) - (T^*(0) = 2)) \end{aligned}$$

et on en déduit le résultat facilement.

Pour mémoire, il y a bien sûr aussi la situation dans laquelle les effets Δ_1 et Δ_2 sont hétérogènes mais indépendants des variables d'entrée $T^*(j)$

Proposition 6 *Sous l'hypothèse $\Delta_1 = X_0c_1 + \varepsilon_1$, avec $E(\varepsilon_1|X_0) = 0$ et $T^*(j) \perp_{\varepsilon_1} | X_0$ et $\Delta_2 = X_0c_2 + \varepsilon_2$, avec $E(\varepsilon_2|X_0) = 0$ et $T^*(j) \perp_{\varepsilon_2} | X_0$, les paramètres identifiés par les conditions d'orthogonalité 37 sont les effets moyens du traitement dans la population.*

5.6 Calcul des contrefactuels

On présente ici le mode de calcul des contrefactuels des figures de la section 4. On distingue le cas des zones avec un seul traitement, où les formules sont exactes, du cas des zones avec deux traitements, où il s'agit d'approximations.

5.6.1 Zones avec un seul traitement

Dans les zones où un seul dispositif a été expérimenté, les contrefactuels se calculent de la façon standard. Le taux de sortie vers l'emploi des *never-takers* est directement observé :

$$E(y(0)|T(1) = 0) = E(y|Z = 1, T = 0).$$

Le taux de sortie vers l'emploi des *compliers* en présence de traitement peut être reconstitué à partir des quantités observées de la façon suivante :

$$\begin{aligned} E(y(1)|T(1) - T(0) = 1) &= \frac{Pr(T = 1|Z = 1)}{Pr(T = 1|Z = 1) - Pr(T = 1|Z = 0)} \times E(y|Z = 1, T = 1) \\ &- \frac{Pr(T = 1|Z = 0)}{Pr(T = 1|Z = 1) - Pr(T = 1|Z = 0)} \times E(y|Z = 0, T = 1). \end{aligned}$$

Enfin, le taux de sortie vers l'emploi des *compliers* en l'absence de traitement est déduit de l'effet du traitement sur les *compliers*, qui est estimé par régression par variables instrumentales (estimations présentées dans les tableaux 6 à 9) :

$$E(y(0)|T(1) - T(0) = 1) = E(y(1)|T(1) - T(0) = 1) - E(y(1) - y(0)|T(1) - T(0) = 1).$$

Les taux moyens des graphiques sont calculés en excluant les demandeurs d'emploi qui entrent en accompagnement bien qu'orientés vers le groupe de contrôle (les *always-takers*). En effet, ces demandeurs d'emploi ne sont jamais observés sans accompagnement, même indirectement : on ne peut donc évaluer leur taux de sortie contrefactuel. Ces demandeurs d'emploi ne représentent une part significative de la population que dans certaines régions où le programme OPP a été déployé seul.

5.6.2 Zones avec deux traitements

Dans les zones où les deux dispositifs sont présents, des hypothèses additionnelles sont nécessaires. Dans la mesure où les proportions de demandeurs d'emploi qui sont orientés par OCC vers le groupe de contrôle et qui se retrouvent dans les accompagnements OPP et CVE sont très faibles (de l'ordre de 2% pour OPP, moins de 1% pour CVE), on en fait abstraction dans le calcul des contrefactuels. Le taux de sortie vers l'emploi des *compliers* en présence d'un accompagnement et le taux de sortie des *never-takers* sont alors directement observés dans chacun des groupes de traitement :

$$E(y(0)|T(t) = 0) = E(y|Z = t, T = 0),$$

$$E(y(1)|T(t) - T(0) = t) \approx E(y(1)|T(t) = t) = E(y|Z = t, T = t),$$

avec $t = 1$ (traitement CVE) ou $t = 2$ (traitement OPP). On en déduit les taux de sortie de l'emploi des *compliers* en l'absence de traitement, en retranchant l'effet estimé du traitement :

$$E(y(0)|T(t) - T(0) = t) = E(y(1)|T(t) - T(0) = t) - E(y(1) - y(0)|T(t) - T(0) = t).$$