



ANTHROPIC

Analyse des données probantes sur la reconversion des travailleurs

Publié

12 août 2026

Auteurs

David Roodman, indépendant

Maxim Massenkoff, Anthropic

Remerciements

Nous remercions David Fein, Richard Hendra et Evan Rose pour leurs précieux commentaires sur une version préliminaire et Shad Ahmed, Jeremy Avins, Jon Baron, Michael Bernick, Adam Farina, Martha Gimbel, Rebecca Hiscott, Anders Humlum, Benjamin Hyman, Tamara Johnson, Molly Kinder, Sang Lee, Frieda Molina, Kathleen Nazar, Laura Peck, Kerry Persen, Sarah Pollack, Mette Rasmussen, Santi Ruiz, Szymon Sacher, Chase Sackett, Kelsey Schaberg, Monika Tuchowska, Garrett Warfield, Kim Withee, Heather Whitney, Nathan Wilmers, Peter McCrory et Jack Clark pour documents, conversations et commentaires.

Le code et les données de ce rapport sont disponibles sur github.com/droodman/job-training-meta-analysis. Une interface web donnant accès à de nombreux autres résultats de méta-analyse est disponible à l'adresse droodman.github.io/job-training-meta-analysis.

Résumé exécutif

De l'avis des experts et du grand public, la reconversion professionnelle des travailleurs est la mesure la plus plébiscitée pour atténuer les perturbations du marché du travail causées par l'IA. Analyse ¹ Dans ce rapport, nous nous demandons si ces efforts seraient couronnés de succès. Nous présentons un ciblé des données probantes sur la formation professionnelle et bref historique des principaux programmes.

Nous avons examiné la littérature scientifique afin d'identifier les essais randomisés de haute qualité, en nous appuyant sur 146 estimations d'impact issues de 56 essais randomisés menés aux États-Unis de 1973 à nos jours. Ces programmes formaient des demandeurs d'emploi – principalement des adultes et des jeunes à faibles revenus – à des métiers tels qu'aide-soignant, technicien de support informatique et soudeur. D'une durée moyenne de six mois, ils coûtent environ 13 000 dollars par personne.

Globalement, nous constatons :

- Les programmes de formation professionnelle ont un impact limité. Dans le cadre d'essais randomisés menés aux États-Unis, Les programmes ont permis d'augmenter l'emploi de 1,7 point de pourcentage en moyenne, comparativement à un taux d'emploi de référence de 63 % dans le groupe témoin. Ils augmentent les revenus de 800 \$ par année en moyenne. Ces chiffres sont par personne bénéficiant du traitement, ce qui est important car l'observance est souvent faible dans ce type d'études. Les programmes réussissent le test de rentabilité, mais ce résultat repose sur des hypothèses incertaines quant à la durée des bénéfices.
- Les programmes à grande échelle ont également donné des résultats décevants. Aux États-Unis, les grands programmes – d'une envergure suffisante pour atténuer les perturbations généralisées du marché du travail – ont fait l'objet d'évaluations rigoureuses. Le constat est généralement négatif. Une évaluation du principal système fédéral de formation dans les années 1980 a révélé des résultats comparables aux moyennes mentionnées précédemment. En revanche, une évaluation de son successeur et du programme Job Corps destiné aux jeunes n'a constaté pratiquement aucun impact.
- Certains « programmes sectoriels » multiplient les revenus par 10. Quelques programmes ont Ces programmes ont permis d'augmenter les salaires de 5 000 à 10 000 dollars par an grâce à une sélection rigoureuse des candidats, à la participation des employeurs à la définition des formations et au suivi des tendances locales en matière de demande de compétences. Du point de vue du gouvernement, ces programmes s'autofinancent clairement grâce à l'augmentation des recettes fiscales et à la diminution du recours aux prestations sociales.
- Les programmes sectoriels peuvent contribuer à une réponse d'urgence aux bouleversements liés à l'IA, mais ils se heurtent à de nombreux obstacles. Leur impact est soumis à plusieurs conditions :
 - Ils éliminent plus de 80 % des candidats. Leurs processus d'admission sélectionnent des travailleurs que les employeurs ne trouveraient probablement pas par eux-mêmes, mais qui possèdent la ténacité et l'intelligence nécessaires pour réussir.

¹ Karger et coll. (2026), graphique 16.

- Ils transmettent des compétences qui s'acquièrent en quelques mois. La rapidité permet de réduire les coûts et offre

Ces programmes offrent aux gestionnaires une meilleure visibilité sur les compétences recherchées à la sortie de l'école. Ils ne sont pas destinés aux professionnels qualifiés qui pourraient avoir besoin de plusieurs années de formation pour retrouver leur niveau de revenus antérieur et qui savent déjà s'orienter sur le marché du travail.

- Les tentatives de mise en place de programmes sectoriels échouent souvent. Il faut parfois des années pour les développer.

Il est essentiel d'établir des relations avec les employeurs locaux et de développer des compétences pour suivre l'évolution de la demande de compétences au niveau régional. Une expansion trop rapide risquerait de masquer la réalité au détriment de l'essentiel.

- La mise à l'échelle et la randomisation des programmes les plus prometteurs permettraient d'obtenir les résultats nécessaires

En appliquant ces leçons à l'IA, nous sommes bien sûr contraints d'extrapoler dans un monde marqué par une grande incertitude. Nous ignorons si, et quand, les progrès technologiques entraîneront une forte hausse du chômage ; il est difficile de prévoir quels emplois seraient les plus touchés, et la population au chômage pourrait différer des sujets de ces études (principalement des personnes à faibles revenus ou des chômeurs de longue durée).

Mais il semble douteux que nous disposions actuellement de programmes capables de répondre aux besoins actuels.

Aucun des grands programmes de formation gouvernementaux n'a produit de bénéfices importants et durables, et les rares programmes sectoriels performants — une option probable compte tenu de l'enthousiasme qu'ils suscitent chez les plus grands experts² — reposent sur des relations étroites avec les employeurs, une sélection rigoureuse et des coûts initiaux élevés. Une évaluation rapide, menée en situation d'urgence, visant à déployer et tester rapidement les programmes de reconversion professionnelle les plus performants, pourrait s'avérer extrêmement fructueuse à long terme.

² Voir, par exemple, Katz et al. (2022).

Aperçu

La principale préoccupation des Américains concernant l'IA est la perte d'emplois (Anthropic 2026). Si l'on se fie aux données actuelles sur l'utilisation de l'IA, les travailleurs diplômés de l'enseignement supérieur – développeurs de logiciels, assistants juridiques, comptables – sont les plus menacés (Massenkoff et McCrory 2026). Mais cette hypothèse reste très incertaine. Les progrès rapides de la robotique et des véhicules autonomes affecteraient d'autres emplois. Et l'emploi des cols blancs moins qualifiés, comme ceux qui travaillent dans la saisie de données ou le service client, pourrait diminuer plus tôt (Richmond 2026).

En cas de perturbation généralisée du marché du travail, les travailleurs déplacés, quelle que soit leur profession, devraient rapidement trouver un emploi en dehors de leur domaine d'expérience et de leur formation. Un gouvernement ambitieux, agissant avec célérité comme lors de la pandémie de COVID-19, envisagerait probablement d'étendre les programmes de reconversion. De fait, dans une récente enquête menée auprès d'économistes, d'experts en IA et de spécialistes de la prévision, le « soutien à la reconversion » est la mesure politique la plus plébiscitée face aux pertes d'emplois induites par l'IA, devançant ainsi le renforcement des assurances chômage, la garantie d'emploi et le revenu universel (Karger et al., 2026).

Ce rapport prend cette possibilité au sérieux en examinant les résultats obtenus par le passé en matière de formation. Nous espérons ainsi contribuer à ajuster les attentes quant à un déploiement à grande échelle des programmes de recyclage en cas de diffusion rapide et disruptive de l'IA, et identifier les axes de recherche les plus urgents.

Nous constatons l'existence d'un nombre important d'initiatives de formation à étudier. Depuis l'administration Kennedy, le gouvernement fédéral américain consacre chaque année des milliards de dollars à la formation professionnelle.

En 2017, ce montant s'élevait à 14 milliards de dollars et concernait l'aide à la recherche d'emploi, le conseil et la reconversion professionnelle, bénéficiant ainsi à environ 10,7 millions de personnes (GAO 2019, Figure 4).

Comme nous le verrons, le bilan de ces dépenses est mitigé. Un récent rapport de la Maison-Blanche qualifie ces programmes de « largement inefficaces » (CEA 2019). Cependant, quelques programmes privés ont obtenu des résultats remarquables grâce à une collaboration étroite avec les employeurs, résultats validés par des études randomisées de grande qualité.

Bien entendu, les données probantes et l'expérience acquise ne correspondent pas parfaitement à un avenir bouleversé par l'IA.

Si les licenciements touchent principalement les cols blancs, la nouvelle vague de chômeurs de longue durée serait la plus diplômée et la mieux rémunérée de l'histoire. Ces travailleurs disposent en moyenne d'une épargne plus importante et savent s'orienter sur le marché du travail (Manning et Aguirre, 2026). S'ils reprennent leurs études, la reconversion pourrait leur prendre des années. Les données recueillies ici concernent des programmes de formation de quelques semaines ou mois, destinés principalement aux personnes à faibles revenus en marge du marché du travail. La reconversion des travailleurs les plus qualifiés pourrait s'apparenter davantage à un master accéléré en milieu de carrière, et les recherches futures devraient examiner comment les cursus universitaires traditionnels pourraient être adaptés et réorientés vers ce contexte. Néanmoins, les programmes étudiés ici offrent des enseignements précieux, et l'essentiel des efforts de reconversion pourrait se concentrer sur les travailleurs occupant des emplois précaires et mal rémunérés.

Ce projet n'est pas une revue systématique à proprement parler, ni plus précisément une méta-analyse. Il n'a pas été conçu dès le départ pour filtrer et synthétiser mécaniquement les résultats, même si nous réalisons finalement une méta-analyse. Notre approche principale consiste à faire appel à notre jugement pour examiner de manière critique les résultats et les conclusions des études individuelles, ainsi que le tableau qu'elles dressent.

Parallèlement à une analyse approfondie des résultats des meilleures études menées dans ce domaine, nous présentons un historique des initiatives et de l'évaluation de la formation professionnelle aux États-Unis. En examinant en détail la mise en œuvre des programmes, nous tirons des enseignements sur leur conception et les difficultés rencontrées, au-delà des simples chiffres d'efficacité apparents. Plus précisément, dans cette analyse, nous :

- Exiger un niveau de preuve élevé. Nous privilégions les essais randomisés, ainsi que les études qui présentent des affirmations a priori plausibles afin de constituer des expériences naturelles convaincantes. Parmi ces dernières, on retrouve la régression par discontinuité et la randomisation par un juge.

3

- Rechercher activement des études répondant à cette norme : études de discontinuité⁴ Nous trouvons 7 randomisés ou sur des interventions en Europe impliquant une formation ; deux études de randomisation par un juge menées aux États-Unis et au Danemark ; et 56 études randomisées aux États-Unis.

- Synthétiser les expériences américaines — 146 estimations d'impact issues de 56 études — avec une nouvelle Méta-analyse. La plupart des expériences étant documentées par de longs rapports soumis aux agences gouvernementales, le logiciel Claude est utilisé pour extraire les caractéristiques et les impacts. Aucune méta-analyse antérieure n'a inclus un nombre aussi important d'études randomisées menées aux États-Unis.

- Examiner individuellement les études nationales les plus importantes aux États-Unis, ainsi que les Des études européennes et des études de randomisation par les juges, afin de mettre en évidence les principaux défis liés à la mise en œuvre et à l'évaluation des programmes.

- Approfondissez vos recherches sur les « programmes sectoriels », qui sélectionnent rigoureusement les candidats, travaillent en étroite collaboration avec les employeurs sur la conception des programmes d'études et présentent certains des résultats les plus prometteurs, augmentant dans certains cas les revenus de 5 000 à 10 000 \$ par an.

³ Nous ne prétendons pas avoir appliqué parfaitement notre « critère de preuve élevé ». Notamment, nous n'avons inclus aucune étude d'appariement, même si certaines sont tout aussi crédibles que les recherches présentées ici. Vers l'an 2000, un vif débat a opposé les économistes quant à la fiabilité des méthodes d'appariement (Dehejia et Wahba, 1999, 2002 ; Smith et Todd, 2005). Les deux camps ont analysé la formation professionnelle. Tous s'accordent à dire qu'une étude d'appariement est plus crédible lorsqu'elle réussit le test de l'effet placebo, en estimant l'absence d'impact sur les résultats avant traitement ; et que cela est particulièrement vrai lorsque la série chronologique avant traitement est longue. Pourtant, la littérature continue de remettre en question les appariements ayant réussi le test de l'effet placebo (Arceneaux, Gerber et Green, 2010 ; Griffen et Todd, 2017). Un bon exemple récent de ce type d'appariement, que nous n'avons pas inclus, est l'évaluation des programmes de formation en Californie réalisée par Rothstein et al. (2022). Ces études méritent sans doute autant de poids que celle de Hyman (2018), que nous abordons au §5.1, et dont la crédibilité (partielle) repose en partie sur une vérification par placebo.

⁴ La recherche « agressive » comprenait la consultation de rapports qui n'étaient pas auparavant disponibles en ligne, à la Bibliothèque du Congrès et à la bibliothèque du Département du Travail ; la prise de contact avec le personnel des organisations qui produisaient certains rapports, afin d'en obtenir des copies ; la recherche d'auteurs désormais retraités ; l'examen approfondi du Digest of Social Experiments de 2005 ; et le suivi des études mentionnées dans des revues antérieures comme étant en cours.

Voici les principaux résultats.

1.1 Aux États-Unis, les impacts moyens sont positifs mais non transformateurs.

Dans les études randomisées menées aux États-Unis, les effets sur l'emploi et les revenus varient de négatifs à positifs, avec une tendance plutôt positive. Toutes les études incluses dans notre méta-analyse ne proposent pas exclusivement une formation ; certaines combinent plusieurs éléments. Les programmes dont la formation était un volet principal ont permis d'augmenter l'emploi de 2,8 points de pourcentage en moyenne la deuxième année suivant l'attribution aléatoire au traitement, et de 1,7 point entre la troisième et la cinquième année. Les impacts sur les revenus avant impôt s'élèvent à 1 139 \$ et 791 \$ par an (en dollars de 2025). Toutes ces moyennes sont hautement significatives statistiquement. (Il s'agit de moyennes par personne ⁵ Voir le tableau 2 au paragraphe 6.5.2. (Les impacts sont exprimés par personne pour laquelle l'offre est présentée) formée, et non par personne formée, les deux variables différant selon que certains membres du groupe de traitement abandonnent le programme ou que des membres du groupe témoin l'intègrent.) Ces résultats, ainsi que de nombreux autres résultats méta-analytiques, peuvent être explorés [graphiquement sur cette page GitHub](#).

Les données brutes relatives aux impacts sur les revenus sont présentées dans la figure 1, qui illustre ces impacts (en dollars par an) en fonction du nombre d'années écoulées depuis le début de l'expérimentation. À partir de chaque rapport d'étude, les estimations ponctuelles sont extraites à la résolution temporelle la plus fine disponible, généralement mensuelle, trimestrielle ou annuelle. Les interventions dont la formation était un élément principal (interventions « axées sur la formation ») sont représentées par des cercles pleins, les autres par des cercles vides. Toutes les estimations sont représentées en gris, à l'exception de quelques programmes dont les points sont reliés et colorés pour les mettre en évidence : bleu pour les grands programmes fédéraux et magenta pour les programmes sectoriels à fort impact ; ces derniers seront présentés plus loin. On constate que les impacts positifs sur les revenus sont plus nombreux que les impacts négatifs, et que la plupart des études cessent le suivi des participants au bout de six ans.

1.2 Les avantages monétaires dépassent les coûts en moyenne, selon des hypothèses raisonnables mais discutables.

Parmi les 67 % d'interventions de formation dont les descriptions incluent des informations sur les coûts, ces derniers s'élèvent en moyenne à 13 598 \$ par participant. Pour une évaluation complète des bénéfices, il est nécessaire de projeter les impacts à long terme, car la plupart des études n'observent les sujets que pendant quelques années. Nous estimons que la valeur actuelle nette (VAN) des impacts sur les revenus, en suivant un flux de bénéfices jusqu'à la retraite, est de 14 146 \$. Elle atteint 19 525 \$ en incluant les charges sociales patronales et autres avantages sociaux (voir § 6.5.3 et l'annexe G). Compte tenu des grandes incertitudes liées à la projection des impacts, ce résultat doit être interprété comme un léger avantage par rapport au seuil de rentabilité. Il est intéressant de noter que les programmes de formation génèrent, à terme, des recettes publiques équivalentes à environ les trois quarts de leurs coûts, grâce à une augmentation des recettes fiscales et à une diminution du recours à l'aide sociale. Ainsi, lorsque l'État les finance, le coût budgétaire à long terme ne représente qu'environ 24 % du coût initial.

⁵ Chacun de ces résultats provient d'un ensemble d'études sous-jacentes légèrement différent, car toutes ne rendent pas compte des effets pour les deux critères d'évaluation et les deux périodes d'observation. Les moyennes sont des moyennes à effets aléatoires. Le choix des périodes d'observation est adapté de Card, Kluve et Weber (2018), qui divisent le suivi en année 1, année 2 et années suivantes. Les effets de l'année 1 sont omis par souci de concision, car ils contiennent souvent des artefacts liés à la phase de formation, par opposition à la formation elle-même.

⁶ Le retour sur investissement est nettement inférieur si les cotisations à la Sécurité sociale ne sont pas considérées uniquement comme des impôts, mais comme une garantie de prestations futures dont l'État est redevable.

1.3 Les impacts des programmes américains à grande échelle sont faibles.

Le gouvernement américain a commandé des évaluations randomisées de trois programmes nationaux de main-d'œuvre :

- À la fin des années 1980, la loi sur le partenariat pour la formation professionnelle (JTPA) a permis d'augmenter le taux d'emploi des adultes à faibles revenus d'environ 2,3 points, partant d'un niveau de base d'environ 70 %. Le salaire ⁷ Cela a augmenté le salaire annuel de moyen des femmes s'élève à 1 100 \$ en dollars de 2025, contre environ 13 000 \$ pour les femmes et 17 500 \$ pour les hommes. Aucun avantage significatif n'a été constaté pour les jeunes issus de familles à faibles revenus (GAO 1996).
- Le Job Corps est un programme résidentiel intensif pour les jeunes, mis en place dans le cadre de la « guerre contre la pauvreté » des années 1960. Il n'a eu aucun impact sur l'emploi ou les revenus lors des suivis s'étendant sur 20 ans, à l'exception d'une hausse transitoire de l'emploi de 1 à 2 points de pourcentage au cours des années 3 à 5 (Schochet 2021).

Une évaluation de la loi sur l'investissement dans la main-d'œuvre (Workforce Investment Act), qui a succédé à la loi sur la planification familiale (JTPA), a abouti à des résultats quasi nuls pour les programmes destinés aux adultes à faibles revenus et aux travailleurs déplacés (Fortson et al., 2017). Les estimations étaient aussi souvent négatives que positives et n'atteignaient pas de signification statistique. L'une des difficultés résidait dans le faible écart de participation : de nombreuses personnes auxquelles une formation avait été proposée ne l'ont pas suivie, tandis que beaucoup, dans le groupe témoin, ont trouvé une formation malgré tout. De ce fait, la différence de participation entre le groupe traité et le groupe témoin n'était que de 15 points de pourcentage.

Les estimations de l'impact sur les bénéfices issues de ces études sont mises en évidence en bleu dans la figure 1.

1.4 Les effets moyens sont plus élevés par stagiaire que par personne à qui la formation est proposée.

Les données dont nous nous basons sont faussées par le non-respect du protocole, c'est-à-dire l'abandon du programme par les participants traités ou la participation du groupe témoin à la formation. Dans notre ensemble d'essais, les personnes traitées ont 66 points de pourcentage de chances supplémentaires de recevoir la formation ciblée et seulement 28 points de pourcentage de chances supplémentaires de recevoir une formation quelconque, comparativement au groupe témoin.

Nos principales estimations donnent l'impact moyen d'une offre, que les gens s'y soient conformés ou non.

Cela nous renseigne sur l'efficacité d'un programme dans son contexte. On peut toutefois envisager un scénario où l'adhésion serait plus forte : une participation bien plus élevée parmi ceux qui ont reçu une offre, et une participation bien plus faible parmi ceux qui n'en ont pas reçu, faute de programmes. Nous pouvons calculer l'impact par stagiaire à partir des études existantes en ajustant les estimations à la proportion de personnes concernées par l'offre. Nous constatons que cette estimation, probablement une limite supérieure, est environ 2 à 4 fois plus élevée que nos estimations principales, sans toutefois modifier le rapport coût-bénéfice.

1.5 Les données probantes sur les programmes d'aide aux travailleurs déplacés – dont le chômage ou le faible salaire est davantage attribuable à des changements économiques à grande échelle tels que la désindustrialisation – sont plus rares et mitigées.

Une étude danoise, qui exploite le caractère arbitraire de l'attribution des travailleurs sociaux (selon le jour du mois de naissance du client), révèle qu'une formation en salle de classe d'une durée moyenne de 52 jours a aidé des personnes à changer de profession ; cet effet était probablement plus marqué chez les jeunes sortant de l'école.

⁷Figure présente les résultats moyens approximatifs pour les années 3 à 5. Voir la figure 4.

des professions exposées à la délocalisation. L'étude ne détecte pas d'effets similaires pour la formation en cours d'emploi (Humlum, Munch et Rasmussen 2025).

Une étude similaire menée aux États-Unis sur le programme d'aide à la reconversion professionnelle (TAA), mais avec une affirmation moins catégorique concernant l'affectation arbitraire, indique que les travailleurs ayant pu bénéficier de cette aide ont effectivement augmenté leur activité professionnelle et leurs revenus. Ils y sont parfois parvenus en changeant de profession ou de lieu de travail. Étant donné que l'aide comprenait à la fois une formation et une prolongation de l'assurance chômage, les effets de la formation en soi ne peuvent être distingués (Hyman 2018).

Comme indiqué précédemment, l'évaluation du programme d'aide aux travailleurs déplacés de la loi de 1998 sur l'investissement dans la main-d'œuvre (WIA) n'a pratiquement rien donné. L'étude n'a pas inclus les personnes admissibles à l'aide temporaire aux travailleurs (TAA) (Fortson et al., 2017).

- Des études randomisées de programmes pilotes pour les travailleurs déplacés dans le New Jersey (Corson et al. 1989) et au Texas (Bloom 1990) montrent qu'une combinaison d'aide à la recherche d'emploi et de formation n'était pas meilleure que l'aide à la recherche d'emploi seule pour réduire le chômage dans les 12 mois.

Nous avons également constaté des indices suggérant que certains programmes qui intègrent d'autres services à la formation obtiennent de bons résultats. Les formations assorties d'un engagement d'embauche de la part des employeurs augmentent les revenus des étudiants admis d'un facteur dix fois supérieur à la moyenne, bien que ce résultat ne concerne que trois programmes.

1.6 Certains « programmes sectoriels » se distinguent par des avantages importants et durables.

Dans les études portant sur les programmes américains, un groupe se distingue par sa rentabilité immédiate. Ces programmes collaborent étroitement avec les employeurs, sélectionnent les candidats, leur enseignent les compétences professionnelles et les techniques d'insertion professionnelle, et les placent dans des emplois de qualité au sein de secteurs spécifiques à forte demande. Les experts s'accordent sur les termes de « stratégie sectorielle » et de « programme sectoriel » pour désigner ces approches de formation (Conway et Giloth, 2014 ; Katz et al., 2022).

⁸ Certains d'entre eux, comme l'année

Up (Fein et Dastrup 2022) et Per Scholas (Maguire et al. 2010 ; Schaberg et Greenberg 2020) sont marqués en nuances de magenta dans la figure 1.

⁸ Il est difficile d'en donner une définition concise. Comme le font remarquer Conway et Giloth (2014), « la diversité des tactiques et des institutions pouvant être impliquées dans une initiative rend complexe la délimitation précise du champ du développement sectoriel de l'emploi ». Katz et al. en proposent une caractérisation utile.

(2022) : « Les programmes d'emploi sectoriels forment les demandeurs d'emploi à des emplois de qualité dans des secteurs et des groupes professionnels spécifiques présentant une forte demande locale de main-d'œuvre et des perspectives d'évolution de carrière à long terme... Ces programmes s'efforcent de nouer des relations solides avec les employeurs, effectuent une présélection des candidats, associent formation aux compétences générales (ou à l'employabilité) à la formation aux compétences professionnelles, participent au développement et au placement des employés, offrent un accompagnement complet pour aider les participants à mener le programme à terme et proposent souvent un suivi aux participants après la fin du programme et aux employeurs après le placement. » Hendra et al. (2016) précisent davantage l'implication des employeurs : « Pour être qualifiée de programme sectoriel, une initiative doit réunir plusieurs employeurs d'un même secteur afin de collaborer au développement d'une main-d'œuvre qualifiée. »

Dans notre méta-analyse, nous utilisons l'implication des employeurs dans les programmes de formation comme facteur de différenciation clé.

Notre échantillon d'études randomisées comprend les programmes sectoriels marqués en magenta, ainsi que d'autres dont les résultats ont été moins probants, du moins durant les périodes d'évaluation. Globalement, le coût des programmes sectoriels n'est pas supérieur à celui des programmes de formation initiale : 11 602 \$ contre 13 598 \$ par membre du groupe d'intervention, en moyenne. Pourtant, la valeur actuelle nette de leur impact sur les revenus avant impôt s'élève à 60 319 \$ au lieu de 14 146 \$. Le ratio avantages-coûts, de 7, est certes inférieur à celui des programmes les plus performants, mais il reste suffisant pour que, si le gouvernement gérait efficacement de tels programmes, ils s'autofinanceraient grâce à l'augmentation des recettes fiscales et à la diminution du recours aux prestations sociales.

Les programmes sectoriels formaient des personnes à un large éventail de métiers, notamment le support informatique et la réparation d'ordinateurs, le développement de logiciels, l'assistance médicale, les soins infirmiers, la facturation médicale, la construction, l'entretien des bâtiments, la production, la comptabilité et la tenue de livres. Bien que des recherches supplémentaires soient nécessaires pour déterminer les qualités essentielles à leur réussite — et que tous les programmes sectoriels n'aient pas obtenu des résultats aussi impressionnants —, plusieurs facteurs sont probablement importants :

- Les programmes sont sélectifs. Une évaluation de Per Scholas et de trois autres programmes indique que ces programmes acceptent environ 1 candidat sur 5 (Hendra et al. 2016, p. 34).
Les processus de candidature exigeants, comprenant de multiples entretiens et/ou du travail à domicile, sélectionnent les personnes qui ont le besoin, la ténacité, les capacités et le soutien familial nécessaires pour s'investir pleinement dans cette expérience stimulante.
 - Sur des périodes se mesurant en mois, les programmes enseignent à la fois des compétences professionnelles, telles que celles nécessaires pour un emploi de débutant dans les centres de données et les hôpitaux, et des « compétences générales » concernant le comportement à adopter lors des entretiens d'embauche et sur le lieu de travail.
- Ces programmes travaillent en étroite collaboration avec les employeurs pour identifier les besoins en compétences à combler et même mettre en place des filières de recrutement. Le Wisconsin Regional Training Partnership de Milwaukee (Maguire et al., 2010), par exemple, a organisé des formations de 2 à 8 semaines en réponse à des demandes spécifiques d'employeurs.
- Tous ces établissements sont privés. Cela leur permet notamment de mieux sélectionner les candidats et d'adapter leurs programmes aux évolutions du marché du travail.

Les programmes sectoriels performants peuvent être considérés en partie comme des facilitateurs d'opportunités ou des intermédiaires. Ils génèrent des rendements exceptionnels en mettant en relation des employés et des employeurs qui, autrement, ne se seraient jamais rencontrés : des employés qui n'auraient peut-être jamais imaginé pouvoir travailler chez JPMorgan Chase ou Amazon, et des employeurs qui n'ont pas investi les ressources nécessaires, ni développé les compétences en interne, pour recruter efficacement parmi les populations à faibles revenus.

Pour que l'arbitrage soit rentable, il faut à la fois des compétences et des opportunités. Les compétences sont rares : les meilleurs organisateurs de partenariats industriels allient la volonté d'aider les chômeurs, propre au travailleur social, à la réactivité d'un professionnel des RH face aux besoins d'une grande entreprise (Dedrick, 2014). Quant aux opportunités, elles se présentent lorsqu'il existe des employés et des employeurs capables de collaborer, mais qui ne se rencontreraient pas spontanément.

Même avec des approches similaires, les conditions de réussite ne sont pas réunies partout. Une tentative de réplcation d'un programme performant, le CET de San José, a échoué sur les 14 sites pilotes, dont 4 gérés par le CET lui-même en dehors de San José (Miller et al., 2005). Seules deux réplcations de l'approche sectorielle ont permis d'augmenter les revenus des bénéficiaires dans le cadre d'un essai randomisé : l'expansion de Year Up hors de Boston (Fein et Dastrup, 2022) ; et St. Nicks Alliance à New York, l'un des trois nouveaux sites du projet pilote WorkAdvance (Yusim et al., 2025, tableau 3).

Earnings impacts in 56 randomized evaluations of US job training programs

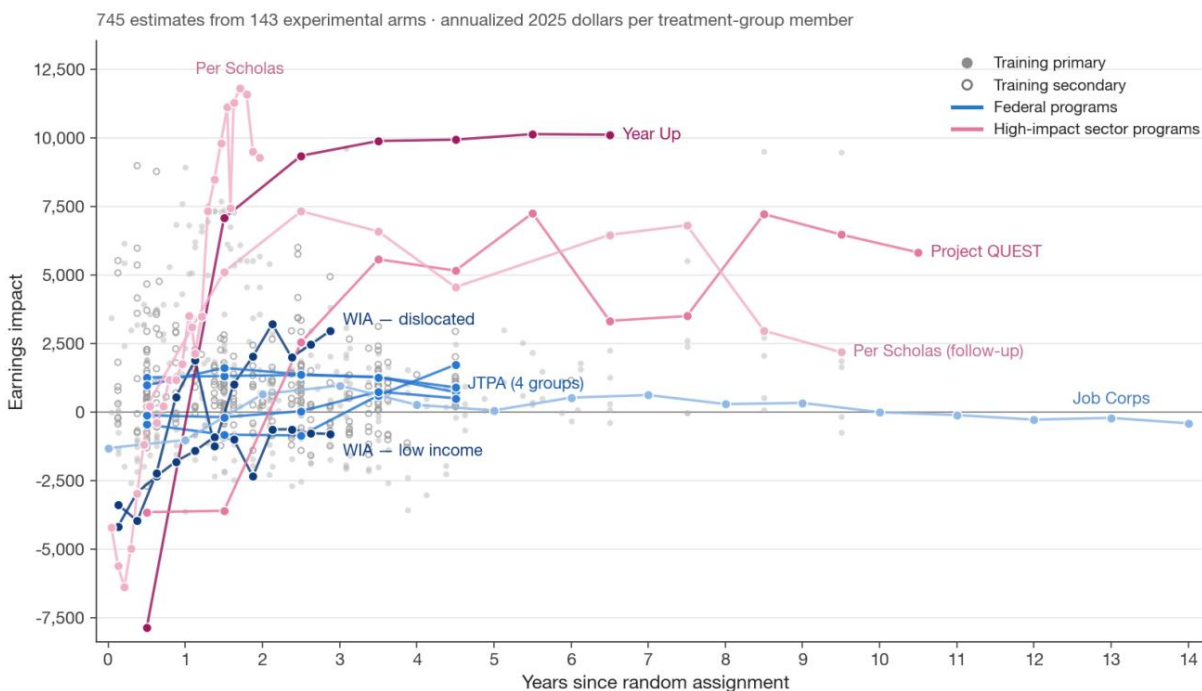


Figure 1. Impacts sur les revenus de 56 évaluations randomisées de programmes de formation professionnelle aux États-Unis. Remarques :

Toutes les estimations d'impact sont extraites et représentées graphiquement avec la résolution temporelle la plus fine des sources primaires, et exprimées en dollars de 2025 par membre du groupe de traitement et par an. Les résultats des programmes fédéraux sont représentés en bleu et certains programmes sectoriels à fort impact en magenta. Au sein de chaque famille de couleurs, les nuances servent uniquement à distinguer les programmes. Tous les autres résultats sont représentés en gris.

Les interventions dont la formation était un élément principal sont représentées par des cercles pleins, les autres par des cercles vides. Chaque série colorée est étiquetée à son extrémité. Per Scholas a fait l'objet de deux études. « JTPA » est le Job Training Partnership Act. « WIA-faible revenu » et « WIA-personnes déplacées » sont deux volets du Workforce Investment Act. Les impacts du JTPA sur les jeunes et ceux du WIA ne sont pas statistiquement significatifs. Par souci de lisibilité, une valeur élevée de l'expérience fondatrice Wildcat des années 1970 a été omise.

1.7 L'approche sectorielle peut probablement aider la société à répondre aux perturbations du travail causées par l'IA. Mais cette approche n'est pas systématique et échouera probablement si on la traite comme telle.

La difficulté de reproduire les programmes sectoriels serait assurément accrue si l'on tentait de les étendre rapidement, que ce soit sous l'égide de bailleurs de fonds publics ou privés. En analysant les échecs de réplcation du CET, les évaluateurs ont noté que « le CET-San Jose s'est développé de manière organique sur 20 ans ».

« Depuis des années, avec un fondateur et un personnel exceptionnellement engagés, [et] des liens très forts avec la communauté locale. » En conséquence, « il est possible qu'un modèle local comme celui de CET ne puisse pas être facilement exporté de manière verticale vers d'autres régions » (Miller et al. 2005).

Bien menée, une initiative visant à étendre les programmes sectoriels pourrait explorer ces possibilités, en collaboration avec des organisations aux compétences reconnues. Le succès de tout programme de formation à grande échelle n'est jamais garanti. Mais ce risque nous incite à intensifier dès maintenant les expérimentations efficaces, plutôt que d'attendre une crise de l'emploi. Si une formation efficace requiert la légitimité des organisations et des liens étroits avec les employeurs, il est préférable de commencer dès maintenant à étendre les programmes performants plutôt que d'attendre une réponse fédérale unique et imposée d'en haut. Une approche rigoureuse serait profitable, ne serait-ce que pour recenser les échecs.

La section suivante de ce rapport présente le contexte. La section 3 expose ensuite certains thèmes critiques transversaux. La section 4 s'attarde sur l'historique de quatre grandes évaluations randomisées américaines de la formation professionnelle, qui constituent la majeure partie de nos données probantes les plus fiables sur la formation à grande échelle. La section 5 examine certaines études non randomisées qui, à première vue, semblent fournir une estimation d'impact fiable. La section 6 passe en revue les méta-analyses antérieures, puis en présente une nouvelle, portant sur des études randomisées menées aux États-Unis. La section 7 couvre l'historique et les données probantes relatives aux programmes sectoriels. La section 8 conclut en proposant des pistes de réflexion sur la manière d'extrapoler ces données probantes à un avenir où le monde du travail sera bouleversé par l'IA.

Dans les annexes C et D, nous passons en revue les preuves limitées et hautement crédibles provenant d'Europe.

2 Contexte

2.1 Types de formation

Les gouvernements et les organisations privées ont mis en œuvre de nombreuses techniques pour aider les gens à trouver un emploi, et à trouver un meilleur emploi. Parmi celles-ci :

Navigation de carrière 1.

Mise à jour et partage des offres d'emploi.

2. Aide aux aspects pratiques de la recherche d'emploi.

Formation directe de la main-d'œuvre 3.

Enseignement des « compétences non techniques », telles que la manière de s'habiller et de se comporter lors d'un entretien d'embauche et sur le lieu de travail.

4. Enseignement en classe des fondamentaux académiques (éducation de base pour adultes).

5. Enseignement des compétences professionnelles.

Apprentissage en milieu de

travail 6. Formation en cours d'emploi.

7. Apprentissages.

8. Subventions à l'emploi dans le secteur privé.

9. Emploi public direct des personnes qui ont besoin de travail.

Nous limiterons généralement notre analyse des programmes de formation professionnelle aux points 3 à 6. Cependant, cette conception comporte quelques complications :

En pratique, il est difficile de distinguer les points 6 et 8. Richard Hendra, du MDRC, commentant une version préliminaire de ce rapport, nous a indiqué que la plupart des formations en milieu de travail aux États-Unis impliquent rarement une formation formelle. Bien entendu, même sans programme d'études, l'emploi subventionné peut former les personnes et faciliter leur reconversion professionnelle.

La plupart des programmes d'insertion professionnelle combinent différentes approches ; par exemple, une aide à la recherche d'emploi associée à une rémunération pour la formation. Dans la méta-analyse présentée au paragraphe 6.5, nous distinguerons globalement les programmes « axés sur la formation », où la formation est l'élément principal selon l'interprétation des rapports par Claude, et les programmes où la formation est un élément secondaire.

De manière générale, nous nous intéressons aux programmes visant à faciliter la reconversion professionnelle, généralement par le biais d'investissements dans le développement des compétences. Plusieurs points non abordés méritent assurément une attention particulière. Le programme d'apprentissage américain est vaste : on comptait 678 014 apprentis en activité en 2025 (date d'arrivée prévue : 2025). Cependant, nos premières recherches n'ont pas permis de trouver de données probantes de qualité sur son efficacité. De plus, les approches de pointe en matière d'adéquation emploi-travail pourraient constituer un outil crucial (et peu coûteux) pour faire face aux perturbations du marché du travail.

3 Thèmes critiques

La plupart des études que nous analysons utilisent la randomisation pour estimer rigoureusement l'impact de la formation professionnelle sur l'emploi et les revenus. Cependant, même la robustesse méthodologique de la randomisation présente des failles. Pour tirer des enseignements de ces recherches, il est essentiel d'examiner attentivement comment les études peuvent ne pas mesurer ce qu'elles visent à mesurer (ce que l'on appelle la validité interne) et dans quelle mesure leurs résultats sont généralisables (validité externe).

Avant d'analyser la situation en détail, cette section recense donc, avec un regard plus global, certains thèmes transversaux et essentiels. Nous abordons d'autres thèmes connexes dans l'annexe A, notamment la question de savoir si les gains des participants se font au détriment des autres travailleurs (A.4) et si la formation est rentable grâce à l'acquisition de nouvelles compétences ou en mettant en valeur des talents préexistants (A.6).

3.1 L'importance de la randomisation

Pour savoir si un programme social est efficace, il est nécessaire de mesurer comment les participants se seraient comportés en son absence. Souvent, les chercheurs comparent les revenus des participants avant et après la formation. Cette méthode est délicate, car certaines personnes peuvent s'inscrire précisément parce que leurs revenus ont diminué ; la plupart d'entre elles verront leurs revenus se rétablir indépendamment de la formation. (Ce biais est appelé « effet Ashenfelter ». Nous en proposons un bref historique dans l'annexe A.2.)

La solution la plus efficace est aussi la plus simple dans son principe : la randomisation. Si un expérimentateur tire à pile ou face pour décider qui aura accès à un programme de formation, alors les groupes traités et témoins auront...

Il s'agit de faire correspondre les trajectoires passées en moyenne. L'expérience a montré que, sans cette caractéristique, même des études dotées de contrôles rigoureux et de stratégies empiriques de pointe peuvent induire les lecteurs en erreur. Pendant des années, les cardiologues ont pris de la vitamine E car des études observationnelles rigoureuses avaient mis en évidence une corrélation persistante entre la vitamine E alimentaire et la santé cardiaque ; des essais randomisés ont finalement infirmé cette corrélation (Lee et al., 2005 ; Abner et al., 2011). Les bénéfices négligeables constatés dans une récente étude à grande échelle sur le revenu garanti ont été surprenants, notamment parce qu'une multitude d'études non randomisées avaient démontré qu'une petite somme d'argent pouvait avoir un impact considérable (Miller et al., 2024). De plus, même l'un des programmes sectoriels les plus performants que nous analysons a publié des rapports annuels contenant des estimations non causales suggérant une augmentation de salaire huit fois supérieure à celle enregistrée dans son essai randomisé.

C'est pourquoi nous privilégions les évaluations randomisées et quelques études observationnelles se réclamant d'une pseudo-randomisation. Certains programmes récents présentent des résultats observationnels encourageants. Toutefois, en l'absence de données randomisées, et compte tenu du nombre important d'échecs initialement prometteurs, nous pourrions être tentés de conclure à l'inefficacité d'un traitement jusqu'à preuve du contraire.

Bien que la randomisation soit un outil puissant pour l'inférence causale, même les études randomisées doivent être lues et interprétées avec un regard critique. La suite de cette section examine des thèmes essentiels qui s'appliquent même aux meilleures évaluations de la formation professionnelle.

3.2 Conformité imparfaite

L'affectation aléatoire à une formation professionnelle ne garantit pas une participation parfaite, notamment lorsque tous les participants du groupe expérimental y participent et aucun de ceux du groupe témoin. En revanche, une offre pourrait avoir fait passer le taux de participation de 25 % (parce qu'un quart du groupe témoin a sollicité le programme) à 50 % (parce que la moitié du groupe expérimental a ignoré l'offre), soit une augmentation de seulement 25 points de pourcentage.

Cela complique les choses. En l'absence d'une conformité totale, les chercheurs se rabattent sur deux notions d'impact causal : l'effet d'intention de traiter (ITT) ou, avec des hypothèses supplémentaires, l'effet moyen local du traitement (LATE ; Imbens et Angrist 1994).

L'ITT correspond à la simple différence de résultats entre ceux qui ont reçu l'offre et ceux qui ne l'ont pas reçue, indépendamment de leur participation. Ainsi, toutes choses égales par ailleurs, si l'impact sur la participation (25 points de pourcentage dans l'exemple ci-dessus) est plus faible, l'ITT l'est également. Sur un marché saturé de programmes de formation efficaces, l'ITT peut être faible simplement parce que le groupe témoin a opté pour d'autres programmes similaires.

L'effet temporaire moyen (LATE) mesure l'impact sur les personnes dont la participation a été déterminée par l'offre. Il est calculé en divisant l'intention de traiter (ITT) par la différence de participation. Ainsi, si notre expérience a augmenté la participation à la formation de 25 % et que son ITT était une augmentation de revenu de 500 \$, le LATE est de $(500 \$ / 25 \%) = 2\,000 \$$. En supposant que l'offre ne puisse vous aider qu'en vous donnant accès à la formation, ce

Un certain groupe de personnes a dû être à l'origine de l'augmentation de 500 dollars, leur avantage a donc dû s'élever à 2 000 dollars.⁹

Quel indicateur est le plus pertinent pour comprendre ces impacts en cas de perturbation du marché du travail ? Les offres de formation initiales (EFI) constituent peut-être la meilleure mesure de l'impact global d'un programme dans son contexte : si un programme ne parvient pas à recruter davantage de personnes pour une formation, il risque d'être inefficace (surtout si ce programme et ses concurrents sont d'efficacité¹⁰ Mais le LATE pourrait indiquer le potentiel d'un programme à comparable). L'aide est plus pertinente dans un contexte où elle est davantage nécessaire et souhaitée (à condition que ses avantages ne diminuent pas significativement à mesure que la participation s'élargit). Cela pourrait faire du LATE un modèle plus adapté à un monde où les suppressions d'emplois sont induites par l'IA : on peut imaginer une offre de formation limitée et une plus grande propension à se former parmi les personnes ayant obtenu une place. Par ailleurs, cette propension pourrait également accroître le recours à la formation et la débrouillardise générale chez les personnes n'ayant pas obtenu de place.

Une dernière complication de l'étude LATE est que les sujets participent parfois à un programme de formation distinct, de sorte qu'il existe différentes façons de la définir : l'impact de la formation ciblée ou l'impact de toute formation (Kline et Walters 2016).

Dans la suite de cette analyse, nous privilégions le modèle ITT, plus simple à définir et qui, compte tenu du niveau d'incertitude élevé, semble tout aussi pertinent pour l'évolution future du marché du travail. Nous présentons également l'augmentation moyenne du taux de participation afin d'illustrer la disponibilité de la formation dans ces différents contextes et de permettre aux lecteurs intéressés d'estimer les bénéfices dans un scénario de conformité totale.

3.3 Attrition

Un écueil économétrique que la randomisation ne permet pas d'éliminer est le risque d'attrition sélective, c'est-à-dire la disparition de participants qui fausse la comparaison entre les groupes de traitement et de contrôle. Par exemple, certaines études ont suivi l'évolution du temps de travail et des revenus des individus au fil des ans en les interrogeant lors d'entretiens de suivi. Or, si les personnes sont moins disposées ou moins enclines à répondre aux enquêteurs lorsqu'elles sont sans emploi, incarcérées ou en rechute, alors les personnes qui réussissent le mieux sur le marché du travail seront surreprésentées dans les données. La réduction non planifiée de l'échantillon pourrait également atténuer les différences apparentes entre les groupes de traitement et de contrôle, ce qui constitue un exemple de biais d'attrition négatif. On peut aussi envisager des scénarios conduisant à un biais d'attrition positif ; cela dépend des personnes qui répondent aux enquêtes. Ainsi, le biais d'attrition demeure une inconnue connue. Nous ignorons ce qu'il est advenu des personnes qui ont quitté l'étude prématurément, et donc l'impact que leur présence aurait eu sur les résultats.

Heureusement, en matière de recherche sur la formation professionnelle, une solution efficace existe : les bases de données gouvernementales. Aux États-Unis, les agences pour l'emploi au niveau des États, ainsi que le Service des impôts (IRS) et l'Administration de la sécurité sociale (SSA), prélèvent des impôts sur les revenus salariaux, directement sur les salaires. Et depuis le milieu des années 1990, le Répertoire national des nouveaux embauchés (National Directory of New Hires) centralise ces données.

⁹ Cette estimation causale ne s'applique qu'à un sous-ensemble appelé les participants ayant suivi l'offre, et non aux personnes qui auraient toujours participé ou jamais participé, indépendamment de l'offre. Leurs bénéfices pourraient différer. De plus, l'hypothèse LATE repose sur le postulat que l'offre n'a influencé les résultats que par son impact sur la formation.

¹⁰ Voir également Heckman et al. (1999) Section 5.2.

Ces données permettent de suivre les personnes redevables de pensions alimentaires. Les chercheurs relient fréquemment (et de manière confidentielle) les numéros de sécurité sociale des personnes concernées à ces bases de données, avec un taux d'abandon minimal.

Bien que les bases de données salariales couvrent la quasi-totalité de la population, elles ne dressent pas un tableau complet de la vie professionnelle d'une personne. Une partie des revenus est perçue au noir. Les travailleurs indépendants ne cotisent pas à l'assurance chômage et n'apparaissent donc pas dans les bases de données de l'emploi. Cependant, il est probable que les personnes interrogées omettent au moins autant d'informations lors d'un entretien. Globalement, le recours aux données gouvernementales atténue considérablement les inquiétudes liées au biais d'attrition (Ashenfelter, 1978), mais de nombreuses études – dont la moitié des évaluations de programmes sectoriels – se limitent à des enquêtes.

3.4 Publication académique vs publication non académique

Il est largement admis que les puissants mécanismes de pression au sein du monde universitaire faussent les résultats publiés dans les revues (Christensen et Miguel, 2018). Dans des cas extrêmes, il y a fraude. Les phénomènes de manipulation des données (p-hacking) (Simonsohn, Nelson et Simmons, 2014) et de biais de publication (Sterling, 1959) sont probablement plus fréquents. Chacun constitue une forme de filtrage, l'une avant la soumission à une revue et l'autre après.

Les données examinées ici sont relativement objectives, et ce pour deux raisons. Premièrement, la plupart des études sont des essais randomisés. Leur simplicité mathématique – la comparaison des moyennes entre les groupes de traitement et de contrôle – réduit (sans toutefois l'éliminer) la marge d'interprétation des calculs, limitant ainsi la possibilité de déceler des résultats statistiquement significatifs. De plus, les expériences randomisées sont plus coûteuses à mettre en œuvre que l'analyse documentaire de données téléchargées, et cet investissement initial réduit le risque que les résultats ne soient pas exploités.

Une dernière raison expliquant le faible taux de sélection tient au fait que la plupart des études analysées ici ont été financées par des organismes gouvernementaux ou des fondations et réalisées par des entreprises de recherche non universitaires ; pour ces entités, la publication académique est secondaire. En effet, pour la grande majorité des études présentées ici, la publication finale consiste en un rapport d'évaluation mis en ligne ou soumis à un organisme gouvernemental.

3.5 Le succès est éphémère

Contrairement, par exemple, à l'assurance chômage, les programmes de formation constituent une catégorie hétérogène et difficile à définir. Ils diffèrent par leur lieu d'intervention, leurs méthodes de recrutement, les personnes formées, les contenus enseignés, les méthodes pédagogiques, la durée de la formation et le corps professoral.

Les organismes qui gèrent les programmes de formation sont eux aussi diversifiés, et ce, de manière tout aussi subtile. Dans une analyse des recherches sur les tentatives de reproduction de programmes de formation professionnelle jugés efficaces lors de leur mise en œuvre initiale, King (2014, p. 211-213) recense 16 facteurs déterminants pour la réussite ou l'échec d'une telle reproduction. Nombre d'entre eux ne peuvent être imposés par simple décret bureaucratique : « un leadership fort et partagé », « un engagement avéré en faveur du développement et de la mise en œuvre de pratiques innovantes », « une planification préalable rigoureuse et une communication continue entre les partenaires ». Les organismes sont, quant à eux, intrinsèquement liés à ces facteurs.

Des contextes qui engendrent certaines incitations – politiques, administratives, financières – peuvent influencer le fonctionnement d'un programme. C'est pourquoi l'extension d'un programme à un plus grand nombre de personnes peut en modifier complètement le fonctionnement.

Dans le domaine de la formation professionnelle, l'exemple du Job Training Partnership Act (JTPA) de 1982 illustre le pouvoir des incitations. Comme expliqué au paragraphe 4.2, le JTPA a délégué la conception des programmes de formation aux acteurs locaux, tout en les tenant responsables du respect des normes de performance. Cette stratégie a faussé la mise en œuvre et engendré la confusion parmi les instances de supervision. Aucune loi du Congrès ne pouvait contourner la loi de Goodhart : lorsqu'une mesure devient un objectif, elle cesse d'être une bonne mesure.

Lors de l'évaluation expérimentale du JTPA, l'équipe de recherche a découvert que parfois, après qu'une personne ait été répartie aléatoirement dans le groupe de traitement, l'agence locale, appelée SDA, tardait à l'enregistrer comme « inscrite » au JTPA jusqu'à ce qu'elle ait trouvé une place de formation en salle de classe ou sur son lieu de travail, et parfois même jusqu'à ce qu'elle ait démontré une bonne assiduité à la formation.

De nombreux organismes de services aux travailleurs (OST) estimaient avoir la latitude de définir le moment où les individus étaient « comptés » dans leurs indicateurs de performance et ont réagi aux incitations du système en retardant leur inscription (Orr et al., 1994). Et si le personnel réagissait aux incitations après l'inscription, il le faisait certainement avant. Avec des financements limités (Orr et al., 1994, p. 55), les OST étaient fortement incités à pratiquer le « creusement », comme l'avait prédit le sénateur Pete Domenici lors des débats au Sénat sur le JTPA. « Autrement dit, ils allaient utiliser les ressources fournies par le gouvernement fédéral pour travailler avec des personnes temporairement sans emploi, mais possédant de bonnes compétences ou une expérience professionnelle significative et susceptibles de retrouver un emploi sans reconversion. » Pourtant, lorsque le JTPA était critiqué dans la presse – ou lorsque la Maison-Blanche posait des questions en coulisses –, les responsables du ministère du Travail citaient les statistiques de performance transmises par les OST pour démontrer l'efficacité du programme.

Empiriquement, les chercheurs ont constaté que les programmes à grande échelle (Allcott 2015 ; Bold 2018) et les programmes gouvernementaux (Vivalt 2020) ont tendance à obtenir de moins bons résultats. Dans une étude plus récente, Bhatt et al. (2021) réexaminent les baisses spectaculaires de la criminalité obtenues par le programme « Becoming a Man » destiné aux jeunes à risque (Heller et al. 2017). Ils concluent : « Au cours des dernières années, avec des taux de participation nettement plus élevés, nous ne pouvons plus rejeter l'hypothèse nulle selon laquelle l'impact est nul. » Dans notre rapport, nous avons identifié trois programmes prometteurs qui ont été délibérément reproduits ou étendus et soumis à des évaluations randomisées : le Center for Employment Training (CET), reproduit sur 14 sites ; Per Scholas, étendu grâce au projet pilote WorkAdvance ; et Year Up, déployé de trois à huit villes. Le CET a échoué sur tous les sites et les nouveaux sites de WorkAdvance ont obtenu des résultats inférieurs à ceux du site initial ; seul Year Up a enregistré des progrès constants.

Ainsi, une question cruciale se pose pour cette étude : si des organisations caritatives privées, en collaboration avec des entreprises et des fondations, mettent au point un système prometteur, les pouvoirs publics peuvent-ils l'étendre à un public beaucoup plus large sans le dénaturer ? Si les facteurs clés de succès sont non seulement éphémères mais aussi fragiles, un nouveau cadre administratif et incitatif risque de les altérer. Il est néanmoins essentiel que les bailleurs de fonds suivent les performances et fondent leurs financements sur celles-ci. Le défi consiste à identifier les grandes caractéristiques du programme qui sont essentielles à sa réussite, puis à suivre et à récompenser les organismes locaux qui mettent en œuvre ces caractéristiques, tout en leur laissant une certaine flexibilité quant aux détails.

4 grandes évaluations randomisées américaines

Depuis 50 ans, le gouvernement américain est à la pointe de la production de données probantes de haute qualité sur l'impact de la formation professionnelle. Cette section passe en revue quatre études majeures qui constituent le socle de cette histoire. La première, bénéficiant d'un financement fédéral massif, visait à démontrer et à tester l'approche du « travail accompagné » pour faciliter l'insertion professionnelle. Les trois autres ont évalué des programmes fédéraux en cours ; elles témoignent de l'efficacité de la formation professionnelle à grande échelle (voir l'annexe A.1).

4.1 Le projet pilote national de travail soutenu

4.1.1 L'étude initiale

Les programmes de formation professionnelle ont une longue histoire en économie. Probablement aucun autre sujet en économie n'a inspiré autant de recherches randomisées. Et aucun autre n'a autant contribué à la « révolution de la crédibilité » en économie (Angrist et Pischke, 2010), qui a rehaussé les normes relatives aux preuves dans l'étude des relations causales. (Voir l'annexe A.2.)

La première grande expérience randomisée en matière de développement de la main-d'œuvre fut le Programme national de démonstration de travail accompagné (NSWD). Conçu comme une étape intermédiaire prudente entre un programme modeste mais prometteur à New York et un éventuel déploiement à l'échelle nationale, le NSWD s'appuyait sur la méthode du « travail accompagné », perfectionnée par le Vera Institute of Justice dans les années 1970 grâce à des années d'expérimentation et au premier essai randomisé mondial de formation professionnelle. Le projet proposait aux participants 12 à 18 mois de travail rémunéré au sein de petites équipes de pairs, sous étroite supervision, suivis d'un accompagnement vers un emploi non subventionné (Vera Institute of Justice, 1976). Les ¹¹ Typique secteurs d'activité concernés incluaient la construction, l'entretien des bâtiments et la garde¹² d'enfants.

Sous l'impulsion de la Fondation Ford, cinq ministères fédéraux se sont associés à elle pour investir la somme considérable de 82,4 millions de dollars dans la démonstration et l'évaluation randomisée du travail soutenu (quatre fois ce montant en valeur actuelle ; MDRC 1980, p. 15, 19 ; Auletta 1982, p. 47). Afin de mener à bien cette démonstration, les financeurs ont créé une nouvelle organisation à but non lucratif, la Manpower Demonstration Research Corporation (MDRC). Celle-ci a ensuite recruté 14 organisations pour mettre en œuvre le projet dans leurs zones d'intervention respectives. La plupart de ces zones étaient de grandes villes, comme Atlanta et Philadelphie ; quelques-unes étaient rurales, comme le centre-nord de la Virginie-Occidentale. Quatre organisations ont été exclues du projet pour une mise en œuvre incomplète du modèle. La démonstration s'est déroulée de mars 1975 à décembre 1978 (MDRC 1980, p. 22, 30).

¹¹ L'origine du travail soutenu est en elle-même une histoire remarquable. Voir l'article du New York Times Magazine, Porter (1974), et l'article Wikipédia sur Herbert Sturz.

¹² Voir aussi Gary Walker, témoignage du 12 mars 1980, devant le Comité du travail et des ressources humaines du Sénat américain (1980, 342).

La formation professionnelle ne constituait pas un élément majeur du programme (dans ¹³ Cependant, le site de New York (le son modèle original), qui orientait certains participants vers des cours de dactylographie et de compétences de base en milieu de travail. Les étudiants de l'une de ces cohortes deviendraient les personnages principaux de la trilogie de Ken Auletta. une série pour le New Yorker, et son livre de 1982, ¹⁴ La sous-classe.

L'expérience et l'analyse des données ont été menées par une entité distincte, Mathematica Policy Research. L'affectation des participants au groupe de traitement ou au groupe témoin, par tirage au sort informatisé, n'a eu lieu qu'après qu'ils aient manifesté leur intérêt pour le programme et réussi la sélection initiale. Sur 10 sites, 6 616 personnes ont participé à l'expérience, réparties équitablement entre le groupe de traitement et le groupe témoin (MDRC 1980, p. 4).

Les personnes affectées au groupe de traitement représentaient un tiers des 10 043 participants à l'étude. Les sujets ont été sélectionnés de manière ciblée parmi quatre populations difficiles à atteindre : les femmes bénéficiant de l'aide aux familles avec enfants à charge (AFDC) depuis au moins 30 des 36 derniers mois et dont le plus jeune enfant était âgé d'au moins 6 ans ; les personnes toxicomanes ayant récemment suivi un traitement ; et les anciens détenus.

récemment libérés de prison ou de centre de détention ; et des adolescents qui ont récemment abandonné l'école. ¹⁵

Les trois derniers groupes étaient composés à 80-94 % d'hommes. Tous les groupes étaient composés à environ 80 % de Noirs et à 10 % d'Hispaniques ; leurs membres avaient passé en moyenne 10 ans à l'école (MDRC 1980, p. 32).

Mathematica a suivi l'emploi, les salaires et d'autres indicateurs en interrogeant régulièrement les participants : initialement (au moment de leur inclusion et de la randomisation) et lors de suivis allant jusqu'à 36 mois. Plusieurs contraintes pratiques ont compliqué l'évolution des échantillons. Tous les programmes n'ont pas démarré à la date prévue, mars 1975, avec une ampleur suffisante pour justifier le déplacement d'une équipe d'enquêteurs. L'expérience a donc débuté plus tard dans certaines régions. Afin d'obtenir une puissance statistique suffisante et d'éviter de se concentrer exclusivement sur la première année des programmes, période durant laquelle le personnel apprend rapidement par tâtonnement, l'expérience a continué d'inclure de nouveaux participants jusqu'en juillet 1977. Les entretiens de suivi ayant largement cessé le 30 mars 1979, la taille des échantillons diminue à mesure que le suivi s'allonge. Les programmes ayant connu un démarrage rapide, notamment à Philadelphie et à Jersey City, représentent une part plus importante du suivi à 36 mois. De plus, on a constaté des abandons (voir § 3.3) : certaines personnes ne se sont pas présentées aux entretiens, ont refusé d'y participer ou étaient injoignables par téléphone ou à domicile. Ce phénomène était plus fréquent dans le groupe témoin, ce qui n'est pas surprenant compte tenu de leur moindre implication dans le programme. 84 % des femmes du groupe AFDC ont été interviewées à 18 mois, contre environ 70 % des participantes des autres groupes (MPR et SSS 2004, p. 21-22).

Il ressort des informations partielles que la plupart des participants ont respecté leur affectation de traitement. 95 % des femmes du groupe de traitement AFDC ont accepté l'offre d'emploi accompagné. Seules 11 % des femmes du groupe témoin AFDC y sont parvenues ; de même, seulement 3 % des anciens toxicomanes et 4 % des jeunes (Heckman et al., 2000, tableau I). Par conséquent, les délais d'action effectifs (DAE) de participation au programme sont proches des délais d'action effectifs (DAI) (voir § 3.2). Aucune information n'est disponible sur le nombre de participants du groupe témoin ayant suivi une formation par ailleurs, ce qui empêche de calculer les DAE pour la formation proprement dite.

¹³« Les directives relatives au travail accompagné autorisaient l'utilisation de 25 % du temps rémunéré pour des services de soutien définis comme « liés au travail ». Ces services comprenaient l'orientation, la formation professionnelle et les activités de préparation et de placement à l'emploi. Toutefois, les programmes locaux n'ont généralement fait qu'un usage limité de cette disposition. » (MDRC 1980, p. 24)

¹⁴Auletta (1981a, 1981b, 1981c).

¹⁵Les définitions précises des groupes se trouvent dans MDRC (1980), Tableau 2-2.

Les figures 2 et 3 présentent les principaux résultats statistiques de l'évaluation : l'évolution de l'emploi et des revenus dans les groupes de traitement et de contrôle. Sur tous les graphiques, les courbes du groupe de contrôle progressent avec le temps écoulé depuis la randomisation. L'attrition pourrait expliquer cette amélioration ¹⁶ Le pays apparente. L'amélioration de la conjoncture économique a probablement aussi joué un rôle : entre le premier trimestre 1975 et le premier trimestre 1979, le taux de chômage des personnes ¹⁷ Et les revenus et noires a diminué, passant de 14,5 % à 12,6 %. L'emploi augmente avec l'âge au début de l'âge adulte. Mais le facteur le plus important est sans doute la régression vers la moyenne : dans un groupe de personnes suffisamment défavorisées pour être éligibles au programme à un moment donné, il ne serait pas surprenant que la plupart d'entre elles voient leur situation s'améliorer, même sans ¹⁸ emploi accompagné.

16 L'attrition ne peut expliquer à elle seule la hausse. Si aucun des 23 % de sujets témoins ayant abandonné l'étude au 9e trimestre (MPR et SSS 2004, tableau 5) n'était employé à ce moment-là, le taux d'emploi réel au 9e trimestre serait de 26,8 %, soit nettement supérieur au niveau initial de 19,4 %.

17 Bureau des statistiques du travail des États-Unis (s.d.), moyennes trimestrielles de la série mensuelle corrigée des variations saisonnières. Une autre version de cette histoire est que les personnes du groupe témoin ont amélioré leur situation en moyenne grâce à d'autres programmes d'emploi, tels que le programme gouvernemental d'incitation au travail (WIN) et le programme de formation professionnelle complet (CETA). Cependant, les membres du groupe témoin n'ont signalé aucune augmentation de leur participation à ces programmes (MDRC 1980, p. 56).

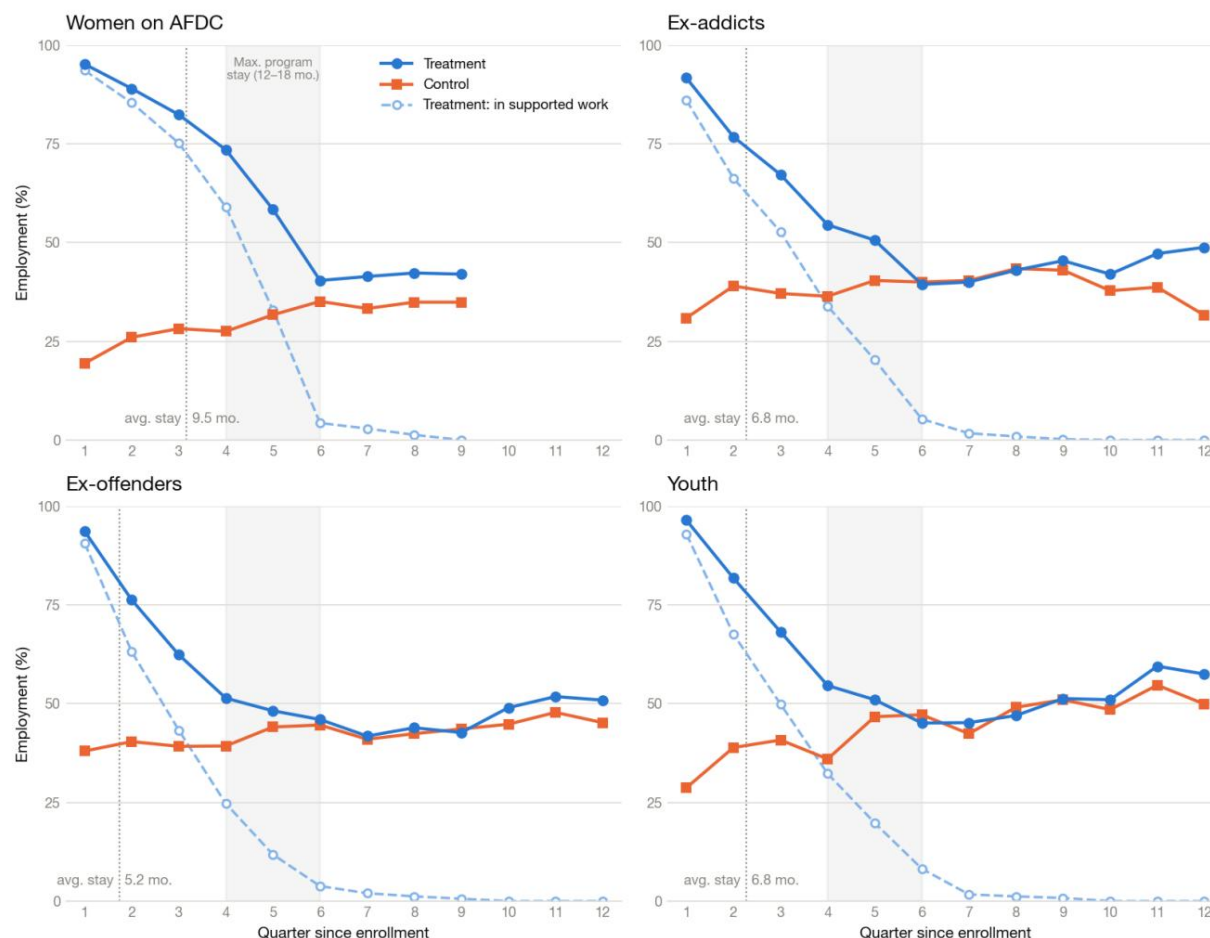


Figure 2. Emploi dans le cadre du programme national de démonstration de travail soutenu (1976-1979)

Remarques : Ce graphique présente la proportion de personnes employées dans les groupes de traitement et de contrôle du Programme national de démonstration de travail accompagné, par trimestre après l'inscription, et ce, séparément pour les quatre groupes cibles. La ligne pointillée indique la proportion de personnes du groupe de traitement occupant un emploi en travail accompagné. Les estimations sont ajustées par régression en fonction de l'âge, du sexe, de l'origine ethnique, du niveau d'études, de l'expérience professionnelle antérieure, de la composition du ménage, du site et de la durée d'activité du site. La zone grisée indique la durée maximale autorisée de participation à un emploi en travail accompagné (12 ou 18 mois, variable selon les sites). La ligne verticale pointillée indique la durée moyenne réelle de participation pour chaque groupe. Les entretiens de suivi de l'AFDC se sont terminés à 27 mois (9^e trimestre) ; les autres groupes ont été suivis jusqu'à 36 mois. Source : taux d'emploi tirés du MDRC (1980), tableaux 4-1, 5-1, 6-1 et 7-1 ; durée maximale de participation tirée de la page 35 ; durée moyenne de participation tirée du tableau 2-7.

Les courbes bleues continues des figures illustrent l'évolution des groupes de traitement. Presque tous les membres des groupes de traitement occupaient un emploi au début du programme, car la quasi-totalité d'entre eux bénéficiaient d'un emploi accompagné (lignes pointillées bleu clair). Cependant, nombreux sont ceux qui ont quitté le programme en quelques mois : démission pour cause d'insatisfaction, licenciement, incarcération, urgences familiales ou obtention d'un emploi stable. Le taux d'emploi se stabilise à un niveau plus élevé.¹⁹ Pourtant, alors que la part des personnes occupant un emploi subventionné chute brutalement en 18 mois, la part de celles qui occupent un emploi quelconque reste inchangée. Pour les femmes bénéficiant de l'AFDC (en haut à gauche de la figure 2), la courbe d'emploi se stabilise juste au-dessus de 40 %. La différence entre le groupe de traitement et le groupe témoin (analyse en intention de traiter) se stabilise à 7,1 points de pourcentage.

¹⁹MDRC (1980), Tableau 2-7 présente des statistiques sur les départs par type pour l'ensemble de la population du programme, et non seulement pour le sous-échantillon expérimental.

(Différence statistiquement significative à $p < 0,1$). Le reste de la figure 2 présente les résultats d'emploi pour les autres échantillons. Si le démarrage tardif de l'expérience pour le groupe AFDC a limité le suivi à neuf trimestres, les autres groupes sont suivis jusqu'à trois ans. Ces derniers présentent²⁰ Tous les trois une évolution particulière : à la sortie du programme, la courbe du groupe expérimental descend jusqu'à celle du groupe témoin – ce qui suggère une absence d'impact immédiat – avant de la dépasser légèrement au cours des derniers trimestres. Le rapport du MDRC juge l'écart observé chez les anciens toxicomanes – 17,2 points de pourcentage au dernier trimestre – statistiquement significatif. Il considère l'impact sur les anciens délinquants comme « marginal » et ne perçoit aucun résultat positif pour les jeunes. Toutefois, la similitude de cette évolution dans les trois groupes suggère qu'il ne s'agit pas d'un simple effet de mode.

Les tendances en matière de revenus, illustrées par la figure 3, sont similaires. Pour les femmes bénéficiant de l'AFDC, un emploi accompagné augmente leur salaire de 75 \$ par mois au cours des derniers trimestres de la période de suivi, soit 900 \$ par an. L'impact sur les revenus est comparable pour les anciens toxicomanes et même les anciens détenus, bien que la différence soit moins significative statistiquement.

Après la publication du rapport principal, le MDRC a effectué un suivi auprès du groupe AFDC 36 mois plus tard. Les données sont disponibles dans des fichiers publics (MDRC 2009, DS7), mais nous n'avons pas encore trouvé de compte rendu et ne pouvons y accéder car nous ne sommes affiliés à aucun membre de l'ICPSR.

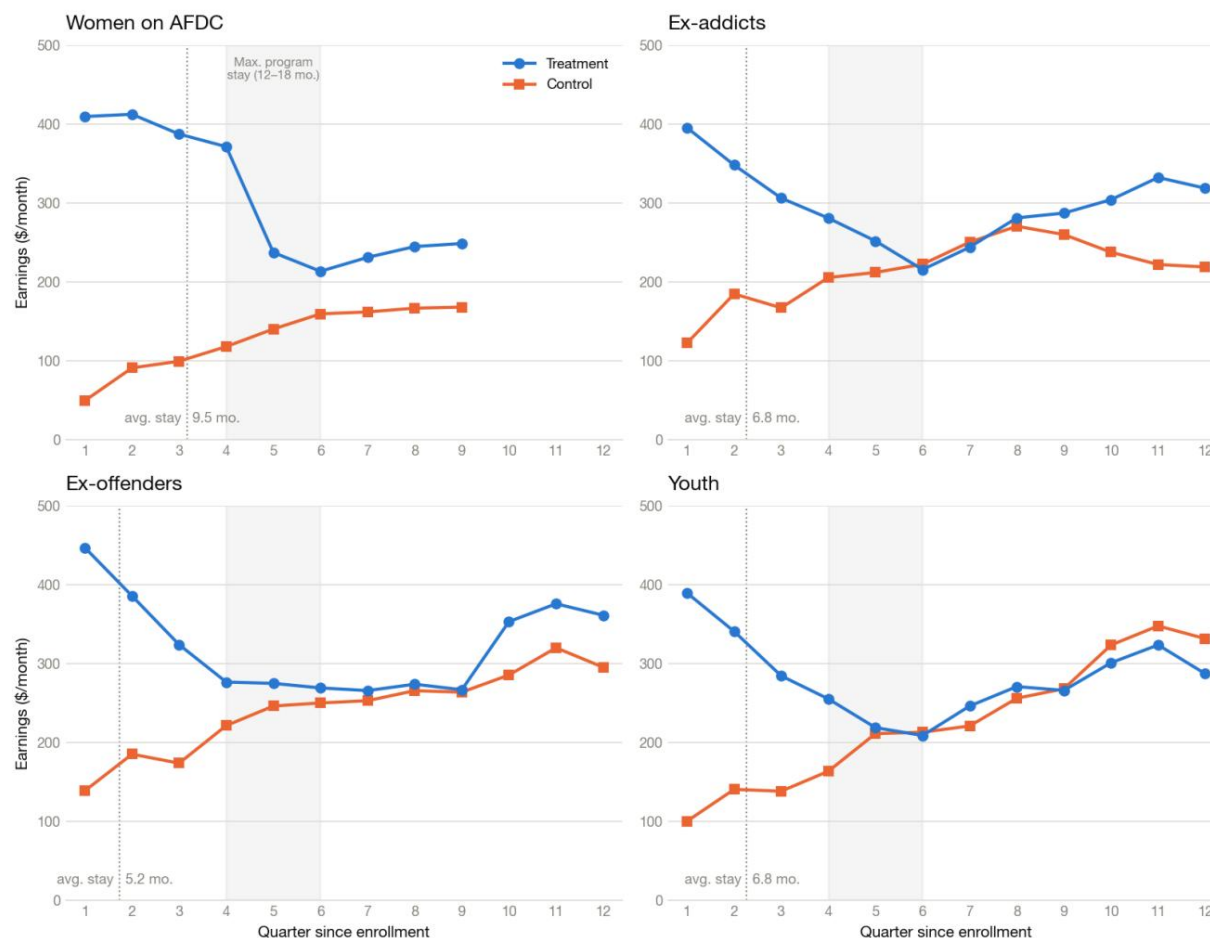


Figure 3. Revenus dans le cadre du programme national de démonstration de travail soutenu (1976-1979)

Remarques : Ce graphique présente les revenus des groupes de traitement et de contrôle du Programme national de démonstration de travail soutenu, par trimestre après l'inscription, séparément pour les quatre groupes cibles. Les estimations sont ajustées par régression ; voir la légende du graphique 2. Source : revenus tirés du MDRC (1980), tableaux 4-3, 5-3, 6-3 et 7-3 ; durée maximale de participation tirée de la page 35 ; durée moyenne de participation tirée du tableau 2-7.

Si le rétrécissement, puis l'élargissement, de l'impact observé chez les anciens toxicomanes et les anciens délinquants constituent un signal et non un simple bruit, qu'est-ce que cela signifie ? Cela pourrait signifier exactement ce que cela semble signifier : le travail accompagné a également aidé les groupes majoritairement masculins, bien qu'avec un délai d'un an après la sortie du programme. Cependant, MDRC (1980, p. 85-88) n'est pas en mesure d'expliquer avec certitude ce qui retarderait l'effet bénéfique. Une deuxième possibilité est un effet de composition : il se peut que les programmes locaux qui ont progressivement dominé l'échantillon au cours des derniers trimestres du suivi, comme ceux de Jersey City et de Philadelphie, aient également été plus efficaces. Toutefois, MDRC (1980, figure 5-3) montre que chez les anciens toxicomanes, le groupe où le signal est le plus fort, les résultats ne changent pas beaucoup si l'on élimine les effets de composition en se limitant aux personnes ayant participé à tous les entretiens. Une troisième possibilité est un biais d'attrition. Le signe d'alerte le plus clair serait alors une chute brutale du taux de participation à l'enquête.

réponse entre le 9e et le 12e trimestre. Mais il n'y a pas eu de chute de ce type au cours de cette période (MPR et SSS 2004, p. 21).

21

Avec une certaine perplexité, nous penchons pour la première explication : les bénéfices différés constatés chez les anciens toxicomanes et les anciens délinquants sont bien réels.

Globalement, l'évaluation suggère qu'à la sortie du programme d'accompagnement, une petite proportion, mais non négligeable, des participants ont trouvé un emploi de meilleure qualité que celui qu'ils auraient pu espérer autrement. Cet impact est particulièrement marqué chez les mères.

4.1.2 Un suivi ultérieur

Après la publication du rapport principal du NSWDC en 1980 – et après la quasi-disparition de la manifestation sous l'administration Reagan en 1981 (Auletta 1999, p. 316) –, un chercheur a étudié ses impacts à long terme en l'associant aux données de la Sécurité sociale (Couch 1992). L'ensemble de données présentait des problèmes. Les informations relatives aux revenus de certaines personnes semblaient manquantes pour 1980-1981, ces données manquantes étant codées de manière trompeuse par des zéros (Couch 1992, p. 383). De plus, la bande de données reçue par Couch ne couvrait que les échantillons AFDC et jeunes. Malgré cela, l'association avec les données gouvernementales semble avoir été relativement exempte d'attrition.

Couch (1992) constate que les effets de l'AFDC sur les salaires des femmes bénéficiaires se sont maintenus jusqu'en 1986, avec toutefois une légère diminution. L'absence d'effet sur les jeunes s'est également maintenue. Pour les femmes, la hausse des revenus s'est élevée en moyenne à 441 dollars par an (en dollars de 1978) entre 1982 et 1986, avec une signification statistique élevée (Couch 1992, tableau 1). Cela représente environ la moitié des 900 \$/an que MDRC a constatés lors d'enquêtes menées vers 1979.

4.1.3 Les impacts en perspective

À bien des égards, le NSWDC a été un succès retentissant : il a permis de recueillir des informations nouvelles et crédibles sur les personnes bénéficiant d'un emploi accompagné. Ces résultats n'auraient pas nécessairement pu être prédits avec certitude ; ils ont donc constitué de véritables enseignements (Ginzberg, Nathan et Solow, 1984, p. 312).

À un autre égard, la démonstration a échoué, car elle n'a quasiment aidé personne. On remarque, sur la figure 2, que les courbes bleues du groupe de traitement chutent rapidement dès le départ : cela correspond aux personnes qui quittent le programme avant de le terminer. Même la différence soutenue de 8 points entre le groupe de traitement et le groupe témoin pour les femmes de l'AFDC est, d'une certaine manière (tendancieuse), exagérée, car elle concerne...

21MDRC (1980, p. 48-49) indique que « des tests statistiques sophistiqués ont été appliqués aux résultats afin de vérifier l'absence de biais lié aux différences de taux de réponse aux entretiens. (Les tests ont montré que les résultats étaient globalement exempts de biais de réponse.) » Ceci semble faire référence à la modélisation de la sélection de Heckman présentée dans l'annexe A du rapport à 18 mois (Maynard, Brown et Schore, 1979). Bien que l'analyse soit pertinente, nous n'y accordons pas une grande importance. Elle ne concerne que le suivi à 18 mois.

Plus important encore, pour garantir la fiabilité du modèle de Heckman, il est nécessaire d'identifier les variables qui influencent la participation à l'enquête et non les résultats étudiés. Or, toutes les variables initiales ayant un impact sur les réponses à l'enquête, comme le fait d'avoir plus de 35 ans et de ne pas avoir été incarcéré au cours des six derniers mois (Maynard, Brown et Schore, 1979, tableau A-4), sont également susceptibles de prédire les résultats. Concernant l'importance des critères d'exclusion, même lorsque le modèle est formellement identifié sans eux, voir Roodman (2011, § 2.2).

le sous-ensemble de femmes qui ont choisi de postuler au programme et donc de participer au tirage au sort. Une mention dans le rapport final (MDRC 1980, p. 43, note 1) suggère qu'environ un tiers des femmes admissibles au programme ont postulé. Ainsi, l'impact sur l'emploi parmi les admissibles a été de :
environ $8/3 \approx 2,5$ points. ²²

L'ouvrage de Ken Auletta, **The Underclass**, donne vie à ces quelques chiffres en partageant ce qu'il a observé dans des classes de travail supervisé à New York :

Au cours des premières semaines de la formation, les effectifs... ont diminué de vingt-deux à seize. Ronald Brooks a été arrêté ; Phillip Rivers, Earl Billings, Larry Pearl et Stanley Lawrence ont été renvoyés du programme pour absentéisme et retards excessifs ; Liza Lance a démissionné car, selon Smith, « elle ne pense pas pouvoir y arriver — elle a l'impression que tous les autres élèves sont plus intelligents qu'elle ». [p. 26]

Plus tard, Auletta nous informe du taux de réussite scolaire. Cinq autres étudiants ont quitté l'université en raison d'une « attitude négative ». « Dans chaque cas, le manque de patience ou de motivation, ou encore une faible estime de soi, ont joué un rôle déterminant » (p. 276).

L'analyse équilibrée est la suivante : a) augmenter l'emploi de 8 points dans un groupe choisi précisément parce qu'il était difficile à aider est une réussite et b) l'intervention n'a pas fonctionné pour tout le monde.

4.1.4 Évaluation comparative des évaluations non randomisées du NSW

Au début des années 1980, le sénateur Dan Quayle a été le principal instigateur de la loi sur le partenariat pour la formation professionnelle (Job Training Partnership Act, JTPA), promulguée par Ronald Reagan et dont l'évaluation sera présentée ci-dessous. À l'instar des lois précédentes, la JTPA chargeait le secrétaire au Travail d'évaluer les programmes autorisés par la loi (loi publique 97-300, article 454). Afin de l'orienter dans cette évaluation, le ministère du Travail a constitué un comité consultatif qui a commandé des rapports à des universitaires et à des évaluateurs professionnels (JTLSRAP 1985 ; Stromsdorfer 1987). L'étude portait notamment sur l'opportunité pour le ministère du Travail de procéder à une évaluation aléatoire de la JTPA.

Les données issues de la démonstration de l'efficacité du travail ont constitué un élément déterminant du débat. Ce débat a, à son tour, influencé notre décision de privilégier nettement les essais randomisés dans cette analyse.

Plusieurs chercheurs ont mené des études analysant les données comme s'il n'y avait pas eu de randomisation. Fraker et Maynard (1984) et LaLonde (1986), notamment, ont constitué des groupes témoins à partir de données extérieures à l'enquête NSW. Ils ont ensuite comparé les estimations non randomisées aux estimations randomisées, jugées plus fiables. Par exemple, selon Fraker et Maynard (1984, tableau IV.2), alors que l'étude randomisée évalue l'impact sur les revenus des jeunes à 7 \$ en 1979, diverses estimations non randomisées varient entre -617 \$ et -1 982 \$. Dans une synthèse conjointe de

²²La note de bas de page du rapport du MDRC indique que s'ils avaient randomisé au sein d'un groupe de personnes éligibles connues plutôt que parmi les candidats, ils auraient eu besoin d'un échantillon de 20 000 plutôt que de 6 600 pour préserver la puissance statistique.

Dans le cadre de projets parallèles, LaLonde et Maynard (1987) concluent que « le scepticisme actuel entourant les résultats des évaluations non expérimentales est justifié ».

Alors que l'article de Fraker et Maynard (1984) a été rédigé sur commande et remis à un bureau du ministère du Travail, celui de LaLonde (1986), paru dans une revue économique de premier plan, a eu une influence bien plus importante. Dans un essai devenu emblématique de la révolution de la crédibilité, Angrist et Pischke (2010) rapportent :

23 Dans

Retracer les origines de la révolution de la crédibilité en économie empirique revient à tenter de retracer la naissance du rock and roll. Les influences initiales sont nombreuses, et chaque fan a son anecdote. Mais, du point de vue de l'économie du travail empirique, on constate qu'une impulsion importante en faveur de meilleures méthodologies et d'essais randomisés plus nombreux émane des études remettant en question la fiabilité des évaluations économétriques des programmes de formation publics subventionnés. L'ouvrage de LaLonde (1986) constitue une référence en la matière...

Dans les années 1980, en tant que membre du comité consultatif du Département du Travail américain (DOL), James Heckman s'est opposé au scepticisme ambiant à l'égard des évaluations non randomisées. Avec de jeunes coauteurs, il a rédigé une réponse à Fraker, Maynard, LaLonde et d'autres : « La récente dénonciation des méthodes non expérimentales d'évaluation des effets de la formation professionnelle est prématurée... Cet article... présente une vision plus optimiste – et réaliste – de la valeur des méthodes non expérimentales pour analyser les effets des programmes de formation sur les salaires » (Heckman, Hotz et Dabos, 1987). Leur argument, développé plus avant dans Heckman et Hotz (1989), n'était pas que le DOL devait se fier aveuglément aux évaluations non expérimentales. Il s'agissait plutôt de souligner que ces évaluations sont souvent les seules disponibles, notamment en raison de leur rapidité et de leur faible coût. De plus, lors de l'examen de recherches non randomisées, il existe des moyens de distinguer les études pertinentes des études non pertinentes. L'outil de sélection principal est désormais appelé « test placebo ». Une méthode d'évaluation fiable, lorsqu'elle est utilisée pour estimer les impacts sur un résultat antérieur à l'intervention, tel que les revenus avant l'entrée dans un emploi accompagné, devrait généralement donner un résultat nul.

Le débat s'est poursuivi pendant des décennies, tandis que les chercheurs élaboraient de nouvelles méthodes complexes pour constituer des groupes témoins en l'absence de randomisation. Parallèlement, le discours scientifique a produit de nombreuses estimations des impacts de l'étude NSW. Un essai récent, d'Imbens et Xu (2025), passe en revue les 40 années écoulées depuis LaLonde (1986). Cette rétrospective conclut à l'amélioration des méthodes non randomisées et rejoint l'avis de Heckman et ses coauteurs quant à l'importance des contrôles par placebo. Ces auteurs ont comparé des méthodes modernes utilisant deux versions de l'ensemble de données NSW. Toutes les méthodes non randomisées échouent aux 24 Pour illustrer ce propos, Imbens et Xu (2025) examinent une multitude de problèmes complexes. contrôles par placebo (Imbens et Xu 2025, Figure 4). Toutes estiment que les travaux soutenus

23 Nous avons publié Fraker et Maynard (1984) sur archive.org.

24 Imbens et Xu (2025) semblent ignorer que Heckman, Hotz et Dabos (1987) ainsi que Heckman et Hotz (1989) ont soulevé ce point. « Heckman et Hotz (1989) ont répondu à la critique de LaLonde concernant les méthodes d'évaluation non expérimentales en préconisant l'utilisation de tests de spécification pour éliminer les estimateurs particulièrement peu performants. Cependant, cette approche n'offrait pas de moyen clair de distinguer les nombreux estimateurs qui, bien qu'ajustant raisonnablement bien aux données, reposent sur des hypothèses d'identification différentes. De ce fait, elle n'a connu qu'un succès limité dans les travaux ultérieurs. » Il me semble que l'un de ces « tests de spécification » était en réalité un test placebo.

Les revenus des hommes ont diminué en 1975 avant leur entrée dans le programme. En revanche, les comparaisons simples et randomisées entre groupe témoin et groupe traité donnent des résultats quasi nuls, comme prévu.

En effet, l'analyse rétrospective d'Imbens et Xu confirme la décision du ministère du Travail à la fin des années 1980 : après avoir entendu le débat des économistes, il a lancé une évaluation randomisée du JTPA.

4.2 Étude sur la loi relative au partenariat pour la formation professionnelle

La loi JTPA a créé, réorganisé ou remanié une série de programmes d'emploi destinés aux anciens combattants, aux Amérindiens, aux personnes handicapées, aux personnes ayant perdu leur emploi suite à des licenciements massifs et aux travailleurs à faibles revenus, correspondant aux catégories visées par la loi NSWD. À la fin des années 1980, le ministère du Travail a commandé une évaluation randomisée pour cette dernière catégorie (titre II-A de la loi), en particulier pour les programmes offerts à l'année, par opposition aux programmes d'été destinés aux jeunes.

Le cadre juridique du JTPA a délégué la responsabilité de ces programmes de formation aux États. Ces derniers ont subdivisé leur territoire en zones de prestation de services (ZPS) et ont créé, pour chacune d'elles, un Conseil de l'industrie privée (CIP) composé majoritairement de représentants de l'industrie, minoritairement d'organisations syndicales et éducatives, et sans aucun représentant du gouvernement. Les CIP ont passé des contrats avec des entités locales pour la mise en œuvre des formations. En contrepartie de la liberté d'élaborer des solutions locales, les CIP étaient soumis à des normes de performance, dont le respect était assorti de la menace de coupes budgétaires.

Cette structure incarnait la philosophie du « nouveau fédéralisme » : il était préférable que le gouvernement fédéral minimise son intervention directe dans l'administration, laisse une plus grande autonomie aux États et responsabilise les prestataires quant aux résultats obtenus (Courty et Marschke, 2011). En pratique, les « résultats » ne désignaient pas les impacts – difficiles à mesurer – mais les retombées, comme le pourcentage de participants quittant un programme de formation pour un emploi à temps plein.

Soumettre un programme national en cours à un test randomisé était, à notre connaissance, sans précédent.

L'administration décentralisée de ce programme national a profondément influencé la signification des résultats.

Indépendants du contrôle des chercheurs, les différents organismes de formation pouvaient recruter des clients différents et leur proposer des services distincts. « Il nous était impossible », écrivent les chercheurs, « à la fois d'atteindre notre objectif d'examiner l'impact des programmes JTPA tels qu'ils étaient mis en œuvre à l'époque... et d'isoler l'effet de la réception d'un service particulier du programme » (Orr et al., 1994). Parallèlement, l'évaluation permettait de répondre plus précisément à une autre question : que se passe-t-il lorsque le gouvernement fédéral dispense des formations à l'échelle nationale et de manière décentralisée ?

L'équipe de recherche espérait mener son expérience auprès d'un sous-ensemble aléatoire des 649 centres de formation d'agents de développement social (ADAS) du pays. Cela aurait permis d'obtenir des résultats plus représentatifs à l'échelle nationale. Or, cela s'est avéré impossible, la plupart des ADAS n'étant pas disposés à exclure aléatoirement une partie de leurs recrues, pour les raisons exposées à l'annexe A.5. « Au final, il a fallu contacter plus de 200 centres de formation pour en trouver 16 acceptant de participer à l'expérience » (Doolittle et Traeger, 1990). Par souci de maîtrise des coûts, l'équipe a généralement dû exclure à la fois les plus petits et les plus grands ADAS : les plus petits, car ils ne permettraient pas de recruter suffisamment de participants, et les plus grands, car de nombreuses organisations distinctes y géraient l'accueil et le recrutement (plus de 50 à Los Angeles ; Orr et al., 1994, p. 41), ce qui compliquait la logistique de l'évaluation. Malgré tout, l'échantillon final

sur 16 SDA, y compris les zones urbaines d'Oakland et de Jersey City, correspondaient presque aux moyennes nationales en matière de chômage, de salaires et de pourcentage de personnes employées dans la fabrication, l'exploitation minière ou l'agriculture (Orr et al. 1994, annexes 3.1, 3.3).

L'attribution aléatoire des sujets a eu lieu entre novembre 1987 et septembre 1989. L'étude a suivi un nombre important de 15 981 participants pendant les 30 mois suivant leur affectation au traitement. Les sujets ont été répartis selon un ratio de 2:1 entre le groupe de traitement et le groupe témoin. Ce déséquilibre a réduit la puissance statistique, mais reflétait la réticence des SDA à refuser des candidats prometteurs. Les résultats ont été analysés pour quatre sous-groupes : hommes et femmes, jeunes (16-21 ans) et adultes (22 ans et plus).

L'expérience a mis en évidence une différence modeste mais perceptible entre le groupe expérimental et le groupe témoin quant à la quantité de traitement reçu. Comme dans l'expérience sur le travail accompagné, la randomisation a eu lieu après le recrutement, la sélection et l'affectation des participants à un service tel qu'une formation en salle ou en situation de travail. Malgré cela, seulement 65 % des personnes assignées aléatoirement au groupe expérimental se sont inscrites officiellement aux services proposés. Les participants du groupe témoin ont trouvé des services similaires ailleurs, souvent subventionnés (Heckman et al., 2000, ²⁵ Pendant ce temps, certaines personnes dans le tableau II). Après 18 mois, ils ont pu intégrer le programme JTPA.

L'adoption partielle et le croisement ont fait que les différences entre le traitement et le contrôle dans la participation au JTPA étaient d'environ 65 points de pourcentage et celles pour la participation à tout programme de formation étaient d'environ 25 points de pourcentage (Orr et al. 1994, annexes 3.15, 4.2, 4.13).

Comprendre l'impact du JTPA sur les résultats en matière d'emploi est étonnamment difficile, car au moins quatre analyses substantielles ont été réalisées :

L'étude officielle révèle que le programme JTPA a augmenté les revenus de 539 \$ pour les femmes adultes et de 550 \$ pour les hommes adultes sur une période de 18 mois. À titre de comparaison, cela représente une augmentation de 4,5 % pour les hommes adultes ; le revenu moyen sur 18 mois dans le groupe témoin était de 12 306 \$ (Bloom et al., 1993, tableau 5.3).

Les chiffres concernant les jeunes femmes et les jeunes hommes sont respectivement de -182 \$ et -854 \$. Le rapport indique que le premier et le dernier de ces quatre chiffres sont significatifs à $p < 0,1$ (Bloom et al. 1993, pièce S.1).

Le rapport de 30 mois (Orr et al., 1994) témoigne des efforts considérables déployés pour comprendre...

L'impact sur 18 mois est négatif pour les jeunes hommes. L'échantillon est divisé selon qu'ils aient déjà été arrêtés ou non. Pour ceux n'ayant jamais été arrêtés, l'impact cumulé sur 30 mois est de -589 \$, une valeur désormais non significative aux seuils conventionnels. Pour ceux ayant déjà été arrêtés, deux valeurs sont rapportées.

La première étude utilise les données des programmes d'assurance chômage, tandis que la seconde repose sur des déclarations faites lors d'entretiens. L'écart entre les deux est frappant : -4 \$ contre -4 209 \$. Malgré de nombreuses recherches, Orr et al. (1994, note 33) n'ont trouvé aucune explication convaincante pour ce dernier écart. Il convient peut-être de l'interpréter comme un exemple du manque de fiabilité des revenus déclarés.

Les estimations d'impact pour les autres groupes démographiques ne varient pas aussi radicalement entre les

²⁵Ces chiffres sont probablement sous-estimés, d'un point de vue technique. Orr et al. (1994, p. 59) estiment que la moitié des personnes déclarées non inscrites l'étaient en réalité, mais n'ont finalement participé que très peu, voire pas du tout, et n'ont donc pas été comptabilisées comme inscrites en raison des primes de performance. Voir § 3.8.

Enquêtes à 18 et 30 mois : les gains totaux sont maintenant de 1 837 \$ et 1 599 \$ pour les hommes et les femmes adultes (tous deux significatifs à $p < 0,1$) et de 104 \$ pour les jeunes femmes (non significatif). (Orr et al. 1994, pièces 4.6, 4.15)

Une analyse de sensibilité réalisée par Heckman et Smith (2000) porte sur l'impact du JTPA sur les revenus des jeunes au cours des 18 premiers mois. L'idée principale est que les résultats sont sensibles à diverses modifications raisonnables de l'analyse, et donc que « l'affirmation selon laquelle les expériences sont supérieures aux méthodes non expérimentales car elles produisent un seul chiffre est fausse ». Cependant, bien que le titre de l'article contienne le mot « sensibilité », le texte ne définit pas précisément ce terme. De plus, nombre de modifications méthodologiques modifient les estimations d'impact de manière significative, ce qui, bien qu'important en valeur absolue, ne représente qu'une fraction des erreurs-types rapportées pour les estimations initiales, les rendant ainsi prévisibles.

- Un suivi à long terme du Bureau général de la comptabilité (GAO 1996) réfute le
Les études originales utilisaient des données déclaratives sur les revenus, contrairement aux études précédentes qui établissaient des liens vers les dossiers de la sécurité sociale. Nous privilégions ce rapport car il introduit une source de données objectives sur les revenus, assure un suivi plus long et évalue également les impacts sur l'emploi.

Les ITT du GAO sur les revenus et l'emploi sont rassemblés dans la figure 4 et la figure B1 de l'annexe.

Ici, l'aspect visuel est un bon indicateur de la significativité statistique. Les différences entre le groupe traité et le groupe témoin sont relativement significatives chez les femmes adultes entre la 1^{re} et la 3^e année suivant le début du traitement, et entre la 3^e et la 4^e année chez les hommes. Pour les deux groupes, la différence s'estompe à partir de la 5^e année. Les impacts significatifs représentent 3 points de pourcentage. L'expérience JTPA a augmenté les revenus salariaux des femmes et des hommes adultes de 500 à 600 dollars par an, avec une légère diminution à partir de la 5^e année. Aucun impact clair n'a été observé chez les jeunes.

Il existe des contrastes et des similitudes entre les résultats du NSWDC et du JTPA. Dans les deux cas, les résultats les plus probants concernent les femmes adultes, et les plus faibles les jeunes. Si le JTPA a moins augmenté le taux d'emploi des femmes (3 points contre 8), il partait d'un niveau plus élevé, ce qui limitait la marge de progression : le taux d'emploi du groupe témoin durant les années où l'impact a été le plus important était d'environ 75 %, contre 35 % pour les femmes bénéficiant de l'AFDC et d'un emploi accompagné. Ce moindre impact pourrait également s'expliquer par un investissement moindre dans la formation. Le JTPA coûtait environ 1 000 \$ par participant (Orr et al., 1994, annexe 6.2), contre 8 000 \$ pour un emploi accompagné (MDRC, 1980, tableau 8-1, ajusté pour tenir compte de l'inflation de 59 % entre 1980 et 1990).

Une dernière cause du faible impact du JTPA mérite d'être soulignée : la faible différence entre les groupes de traitement et de contrôle quant au recours à la formation. Ceci s'explique en grande partie par le fait que les membres du groupe de contrôle ont trouvé une formation par d'autres voies – souvent la même formation, mais moins subventionnée (Heckman et al., 2000, p. 664). Contrairement aux emplois accompagnés coûteux, les services du JTPA étaient probablement assez représentatifs des services de formation en général. Il est donc raisonnable de penser que les alternatives ont eu des impacts similaires. Le recours à ces alternatives a amélioré les résultats de référence dans les groupes de contrôle et a diminué l'impact apparent du traitement. En divisant par l'impact typique de 25 % sur le recours à toute formation, mentionné précédemment, on obtient des LATE pour toute formation quatre fois supérieurs aux ITT – soit environ 12 points de pourcentage d'emploi pour les adultes et 2 000 à 2 400 \$ par an de revenus.

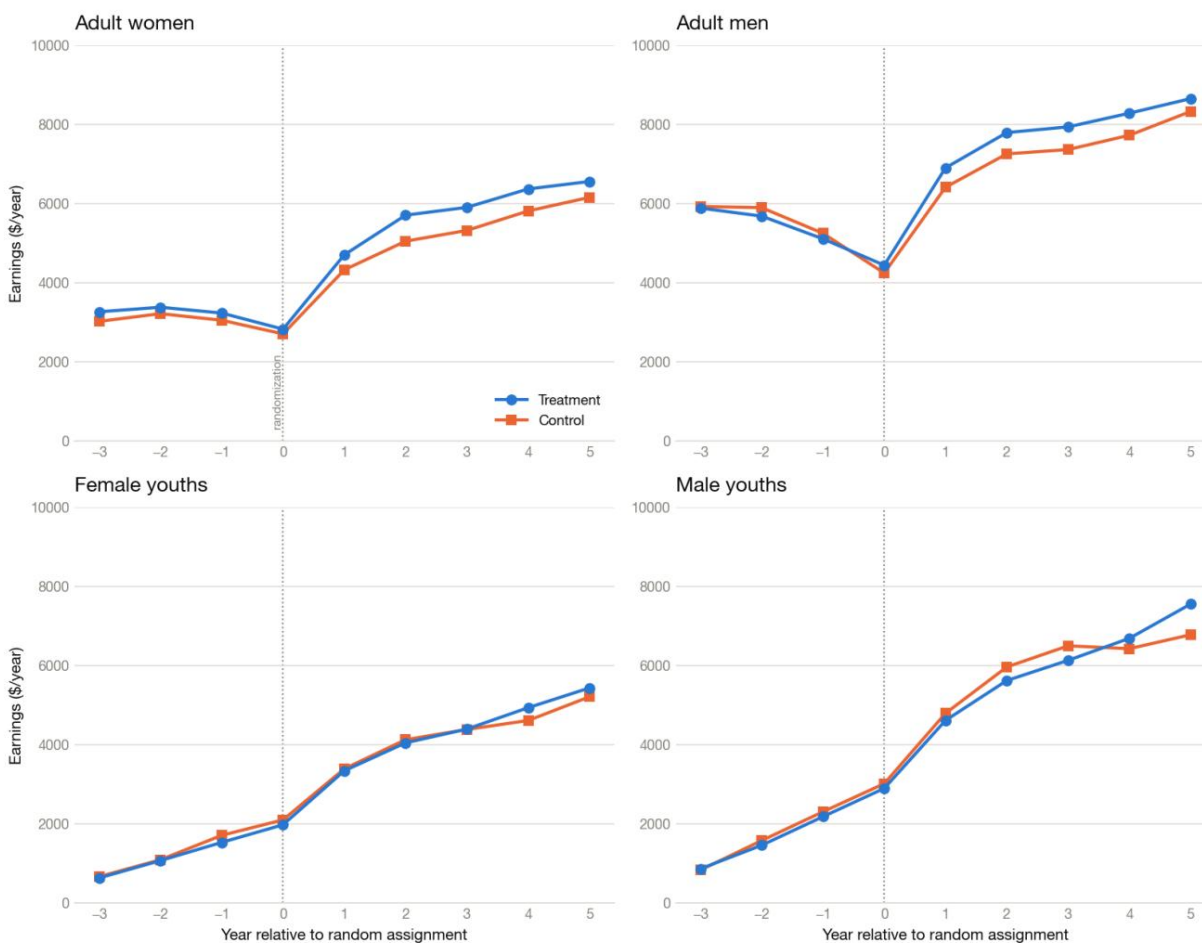


Figure 4. Emploi dans l'étude nationale JTPA par année avant/depuis la randomisation Remarques : Source : GAO (1996), tableaux II.5–II.8.

4.3 Étude nationale sur le Job Corps

La première salve de la guerre contre la pauvreté menée par Lyndon Johnson fut la loi sur l'égalité des chances économiques de 1964. Ce programme a créé plusieurs initiatives sociales, dont le Job Corps, qui, encore aujourd'hui, s'adresse aux adolescents et jeunes adultes ayant abandonné leurs études ou rencontrant des difficultés d'insertion professionnelle. Au début des années 1990, alors que le programme fêtait ses 30 ans, le Département du Travail a décidé d'en évaluer l'efficacité par une étude randomisée. Les responsables du Département étaient sans doute conscients des résultats décevants obtenus par le passé avec des programmes destinés à la jeunesse. Comme nous l'avons vu, le National Supported Work Demonstration et le JTPA ont eu peu d'impact sur les jeunes. De plus, une évaluation financée par le Département du Travail et portant sur 13 programmes de formation locaux pour les décrocheurs scolaires, appelée JOBSTART, n'avait récemment identifié qu'un seul programme ayant un impact significatif (voir §7.1 ci-dessous). Parmi les 12 programmes décevants figuraient les antennes du Job Corps à Atlanta, Phoenix et Los Angeles (Cave et al., 1993).

De toute évidence, la théorie du changement du Job Corps repose sur l'idée qu'immerger les jeunes dans un environnement structuré et favorable les orientera vers une autre voie. Au milieu des années 1990, lors de l'étude nationale sur le Job Corps, 87 % des stagiaires vivaient dans leurs centres de formation. (Nombre de ces bâtiments étaient autrefois des écoles ou des installations militaires ; Schochet (2021), p. 4.) « Au sein des centres », expliquent les chercheurs de Mathematica,

Les participants bénéficient d'une formation professionnelle intensive, d'un enseignement général et d'un large éventail de services, notamment du soutien psychologique, des ateliers de développement des compétences sociales et des programmes d'éducation à la santé. Le programme Job Corps propose des formations professionnelles dans plus de 75 métiers, chaque centre en proposant généralement 10 ou 11. Les programmes de formation professionnelle ont été élaborés en collaboration avec des organisations patronales et syndicales et mettent l'accent sur l'acquisition des compétences spécifiques nécessaires à l'exercice d'un métier (Schochet, Burghardt et McConnell, 2008).

La participation était volontaire mais nécessitait une candidature. Les étudiants pouvaient également choisir la date de leur sortie. Dans l'étude, la durée moyenne de séjour était de 8 mois. Environ un quart des participants ont séjourné moins de 3 mois et un autre quart plus de 12 mois (Schochet 2021, p. 4).

L'étude nationale sur le Job Corps était impressionnante. Elle a fait appel à des méthodes tant qualitatives que quantitatives. L'équipe de recherche a produit une douzaine de rapports totalisant des centaines de pages.

26

Il s'agissait de la première étude randomisée d'un programme de formation professionnelle – et peut-être de tout programme social – dont les résultats étaient, de par leur conception, représentatifs à l'échelle nationale. La quasi-totalité des candidats entre fin 1994 et fin 1996 ont été inclus dans l'étude par randomisation. 5 977 candidats ont été affectés au groupe témoin, 9 409 au groupe expérimental, et les quelque 81 000 candidats restants ont été exclus de l'étude – comme les participants du groupe expérimental, libres de participer, mais sans suivi. La non-participation de la plupart des candidats à l'étude a permis de minimiser les perturbations des programmes en cours et, par conséquent, tout biais d'observation (Annexe A.3).

L'une des difficultés rencontrées par les chercheurs résidait dans la mesure de l'emploi et des revenus. La section 3.3, expliquant le biais d'attrition, aborde le compromis entre le recours aux données autodéclarées et aux bases de données gouvernementales. Un article de synthèse, Schochet, Burghardt et McConnell (2008), ci-après SBM, s'appuie sur les deux sources. Du côté gouvernemental, il utilise deux bases de données : l'une relative à l'assurance chômage et l'autre à la sécurité sociale. Un article ultérieur, Schochet (2021), y ajoute une base de données du fisc américain (IRS) et suit ainsi les sujets pendant 20 ans, soit la plus longue période de cette étude.

Les résultats de toutes les sources de données, chacune disponible sur une période d'année distincte, apparaissent dans la figure 5.²⁷ Le graphique de droite, plus lisible, montre que les jeunes ont déclaré gagner le double de ce que leurs employeurs ont déclaré à la Sécurité sociale. Les trois sources officielles sont plus cohérentes.

²⁶Les rapports sont indexés dans la base de données ERIC (Institut des sciences de l'éducation) sous l'expression « Étude nationale sur le Job Corps ».

Les résultats sont présentés par année civile et non en fonction du nombre d'années écoulées depuis la date de randomisation, qui varie selon les sujets. Pour des raisons de confidentialité, les organismes détenant les données ont limité l'accès des chercheurs à celles-ci. Ceci explique probablement pourquoi les chercheurs n'ont pu extraire que des résultats par année civile.

entre elles. (Les données de l'IRS utilisées ici, qui sont celles qui ont pu être extraites de Schochet (2021), ne comprennent que les revenus salariaux déclarés sur les formulaires W-2, et non les revenus d'activité indépendante.) Hormis une augmentation rapide et autodéclarée des revenus dans le groupe de traitement par rapport au groupe témoin, les courbes des deux groupes sont pratiquement indiscernables.

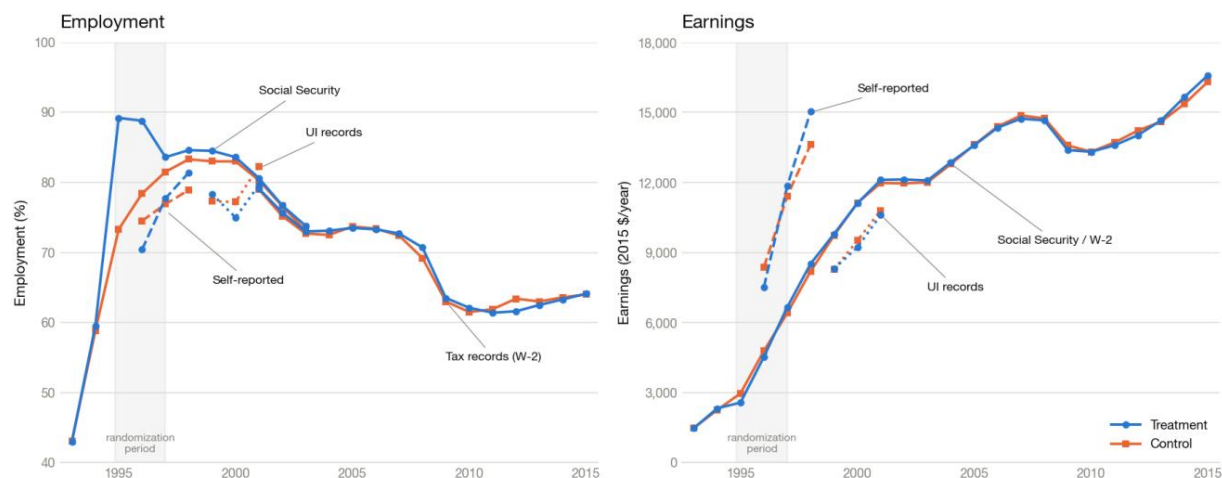


Figure 5. Emploi et salaires dans l'étude nationale Job Corps, tous âges Remarques : Basé sur Schochet, Burghardt et McConnell (2008), tableau 4, et Schochet (2021), tableaux A.1–A.2.

Le constat est sensiblement le même concernant la situation professionnelle (partie gauche de la figure 5). Le seul départ notable survient pendant et juste après la randomisation, lorsque les sujets traités étaient davantage employés, vraisemblablement dans le cadre de stages ou d'autres formes d'emploi proposées par le programme Job Corps.

Dans le graphique de droite, le revenu déclaré entre 1996 et 1998 augmente particulièrement fortement dans le groupe traité. De ce fait, la différence entre le groupe traité et le groupe témoin en 1998 est beaucoup plus importante selon les déclarations des participants que selon les données de la Sécurité sociale (972 \$ contre 218 \$). Cette différence est elle-même hautement significative d'un point de vue statistique (SBM, tableau 4). Le SBM (§7) aborde la question cruciale de la conciliation de ces résultats. Deux mécanismes pourraient être en jeu : des différences dans les informations prises en compte (biais d'attrition) et des différences dans les informations elles-mêmes. Dans un premier temps, le SBM recalcule les moyennes basées sur les données de la Sécurité sociale uniquement pour l'échantillon ayant fourni les déclarations des participants. Cela double presque l'écart entre le groupe traité et le groupe témoin pour la Sécurité sociale, le portant à 393 \$, et suggère fortement la présence d'un biais d'attrition dans les déclarations des participants. Des investigations complémentaires mettent en évidence d'autres facteurs : les personnes interrogées ont peut-être surestimé leur temps de travail, et donc leur rémunération totale compte tenu de leur taux horaire. De plus, comme indiqué précédemment, certains revenus ne sont pas déclarés à la Sécurité sociale. Ces deux effets pourraient agir comme des multiplicateurs, amplifiant les valeurs absolues et, par conséquent, les écarts entre elles.

La publication de l'étude de suivi sur 20 ans, basée sur les données de l'IRS (Schochet, 2021), a suscité la controverse. La cause principale pourrait résider dans la tendance de l'étude à privilégier les résultats positifs et statistiquement significatifs. On observe ce biais à l'œuvre successivement. Premièrement, alors que l'étude de suivi sur neuf ans (SBM, tableau 5) calcule les impacts sur plusieurs sous-groupes – définis par l'âge, le sexe, l'origine ethnique ou le lieu de résidence de l'étudiant sur le lieu de formation –, l'étude de Schochet (2021) se concentre uniquement sur...

Les analyses portent sur différents groupes d'âge. Schochet (2021) explique que, pour des raisons de confidentialité, les responsables du Trésor ont dû effectuer l'analyse sur les données de l'IRS. De plus, « en raison de préoccupations liées à la divulgation des données et au temps consacré par le personnel à l'évaluation des risques associés, les estimations d'impact n'ont pas pu être obtenues pour les autres sous-groupes examinés dans les analyses précédentes ». Cela implique que Schochet (2021) dispose d'une certaine latitude pour prioriser l'analyse par âge. Deuxièmement, parmi les trois groupes d'âge (16-17 ans, 18-19 ans et 20-24 ans), seuls les impacts pour le dernier sont représentés graphiquement. Troisièmement, après avoir vérifié l'existence d'un impact dans cette tranche d'âge pour chaque année de 1993 à 2015, et séparément pour l'emploi et les revenus, le texte se concentre sur la dernière année, où l'impact sur l'emploi est significatif ($p < 0,1$). Quatrièmement, le résumé mentionne ce résultat positif, mais pas les résultats nuls pour l'échantillon complet (comme le montrent les courbes de l'IRS dans la figure 5).

Cette dernière omission a suscité l'ire de Straight Talk on Evidence, un projet d'Arnold Ventures.

Le projet mettait en garde contre le risque que les décideurs politiques, trop occupés, ne lisent que le résumé et soient induits en erreur.²⁸

L'article complet est désormais clair sur les résultats nuls pour l'échantillon complet (Schochet 2021, pp. 12–16). De plus, en réponse à l'inquiétude selon laquelle l'article mettrait en avant de manière sélective les résultats les plus prometteurs, qui pourraient alors n'être qu'un bruit de fond filtré, Schochet (2021) propose implicitement plusieurs réfutations. Toutes s'appuient conceptuellement sur les travaux de Peter Schochet sur les tests d'hypothèses multiples dans l'évaluation des politiques sociales (Schochet, 2009). Cette analyse, ainsi que les autres analyses du SBM,²⁹ Le groupe d'âge avait été planifiée il y a 20 ans. Les changements de classification, de 16-17 ans à 18-19 ans puis à 20-24 ans, ont des conséquences intrinsèques plus importantes qu'on ne le pense.

En effet, à leur entrée au programme Job Corps, les participants les plus âgés avaient davantage d'expérience professionnelle que les plus jeunes (90,4 % avaient déjà travaillé contre 68,4 % au départ), gagnaient plus (5,47 \$ contre 4,71 \$ de l'heure) et étaient plus susceptibles d'avoir des enfants (34,3 % contre 8,6 %) (Schochet 2021, tableau 1). Des éléments indirects corroborent le constat que les participants plus âgés ont davantage tiré profit du programme : ils y sont restés plus longtemps, ont consacré plus de temps aux activités éducatives et de formation, et leur motivation était plus forte, d'après une enquête distincte menée auprès du personnel du programme (SBM, p. 1876-1877). Ajoutons que, compte tenu du constat historique des difficultés rencontrées par les programmes de formation pour aider les jeunes, il serait peu surprenant que Job Corps soit plus bénéfique à ses participants les plus âgés.

Afin d'éviter de filtrer par inadvertance les résultats prometteurs ou significatifs, nous présentons et évaluons les données de manière symétrique les deux critères d'évaluation et les trois groupes d'âge, en présentant les résultats pour chacun. Dans le cadre des données publiées, la méthode la plus appropriée consiste à utiliser les tables de transfert (TOT), telles qu'elles sont reproduites dans Schochet (2021).

28. Parler franchement des preuves (2020).

29L'article Straight Talk comprend également une réponse de Peter Schochet.

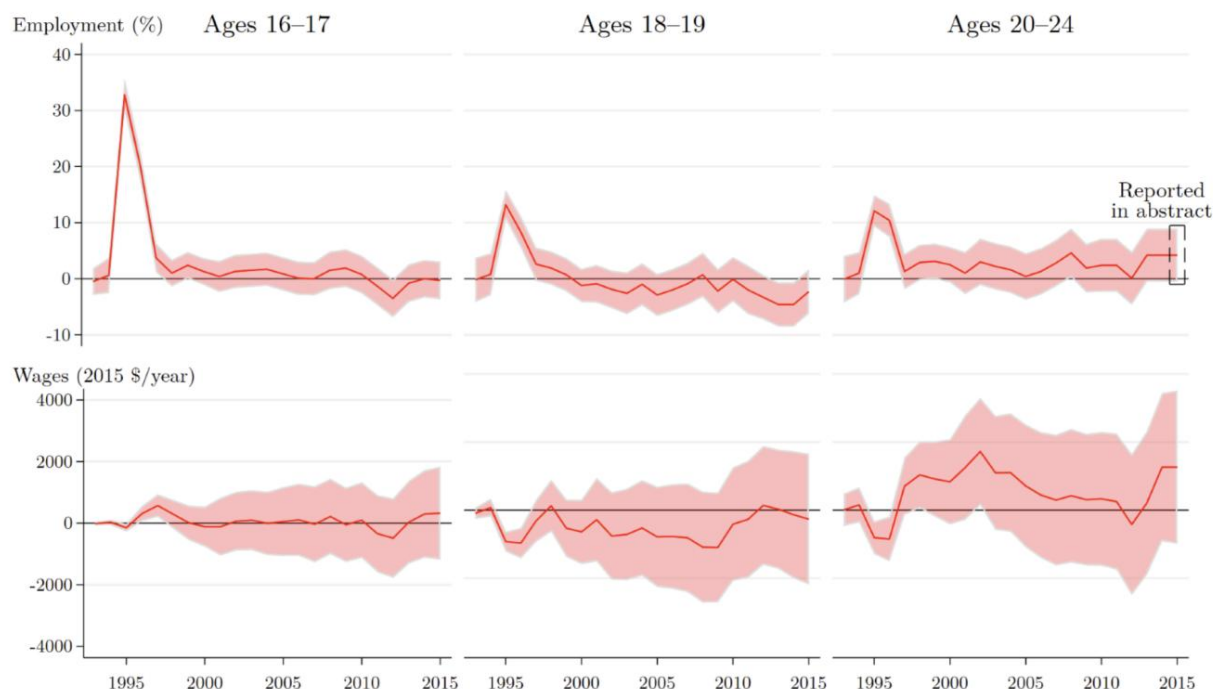


Figure 6. Impacts estimés du traitement sur l'emploi et les salaires dans l'étude nationale Job Corps, par groupe d'âge, d'après Schochet (2021)

Remarques : D'après Schochet (2021), tableaux A.3 et A.4. Sont présentés les impacts estimés du traitement sur les traités (TOT) ainsi que leurs intervalles de confiance à 95 %. Les TOT sont calculés en divisant la différence entre le groupe traitement et le groupe contrôle pour un résultat moyen par la différence entre le groupe traitement et le groupe contrôle pour le taux moyen d'adoption du traitement. Les données proviennent de la Social Security Administration jusqu'en 2000 et de l'IRS ensuite.

La figure 6 présente les TOT (Total Effects) du programme Job Corps par année et par tranche d'âge. Toutes les données sous-jacentes proviennent d'organismes gouvernementaux : la Sécurité sociale américaine jusqu'en 2000, puis l'IRS (Internal Revenue Service). Les intervalles de confiance à 95 % sont également représentés. Il convient de noter que ces intervalles sous-estiment, dans une certaine mesure, la précision des estimations d'impact, car ils concernent chaque année individuellement. De même que la combinaison de deux sondages accroît la précision, calculer l'impact moyen sur la période 2011-2015, par exemple, au lieu d'année en année, permettrait d'obtenir le même résultat.³⁰

Dans le panneau supérieur droit de la figure, à l'extrême droite, on peut observer la statistique mise en évidence dans le résumé de Schochet (2021) : la perte totale d'emploi (TOT) des jeunes de 20 à 24 ans ayant intégré le Job Corps était de 4,2 points de pourcentage en 2015, avec une signification statistique. Cependant, sur l'ensemble de la période de suivi, cette TOT est généralement moins importante et moins significative. En élargissant la perspective, l'impact sur les revenus de cette même tranche d'âge – l'indicateur du travail mis en avant dans le rapport de 2008 (SBM) – est plus difficile à distinguer de zéro. Le résultat le plus significatif de la figure est la baisse de 4,6 points de pourcentage de l'emploi chez les 18-19 ans en 2013 et 2014. Ce résultat est difficile à expliquer et n'est donc pas mis en évidence.

³⁰ L'effet est plus faible que lors de la combinaison de sondages échantillonnés indépendamment, car les impacts au cours d'années successives au sein de la même cohorte ne sont pas statistiquement indépendants.

Nous accordons donc une importance considérable à l'hypothèse que les écarts par rapport à zéro dans les résultats par groupe d'âge soient principalement dus à des fluctuations. La courbe de l'emploi chez les plus jeunes se situe autour de zéro, celle du groupe intermédiaire est légèrement inférieure, et celle des plus âgés est légèrement supérieure. Si l'on attribue le résultat négatif observé chez les 18-19 ans à des fluctuations, la même cause pourrait expliquer l'impact positif constaté chez les plus âgés.

Malheureusement, le signal fort de la figure 5 — pratiquement aucun impact global — semble être le principal verdict de l'évaluation du Job Corps.

4.4 Évaluation de référence selon la loi sur l'investissement dans la main-d'œuvre

En 1998, le Congrès a de nouveau réformé l'ensemble des programmes fédéraux d'aide à l'emploi avec la loi sur l'investissement dans la main-d'œuvre (WIA). Cette loi a atténué l'importance accordée traditionnellement aux personnes à faibles revenus ; a partiellement relégué au second plan la formation professionnelle au profit de l'aide à la valorisation des compétences déjà acquises ; a centralisé les services d'information et de conseil dans des guichets uniques pour l'emploi ; a donné aux demandeurs d'emploi un plus grand contrôle sur le choix des services grâce à un système de bons appelés Comptes individuels de formation ; et a regroupé divers programmes en trois filières : jeunes, adultes et travailleurs déplacés (Blank, Heald et Fagnoni 2011, p. 49 ; Bradley 2013, p. 49).

4) Le volet Jeunesse comprenait le Job Corps. Le volet Adultes a perpétué l'orientation historique vers les personnes à faibles revenus. Le volet Travailleurs déplacés concernait les personnes confrontées à diverses situations difficiles : perte d'emploi suite à une réduction d'effectifs ou à la fermeture d'une usine, baisse de la demande pour leurs compétences en raison de mutations économiques générales, etc. (Fortson et al., 2017, p. xvi).

À l'instar du JTPA qui l'a précédé, le WIA a autorisé le Département du Travail à mener des évaluations. Cela a abouti, à terme, à une autre évaluation nationale et randomisée des activités en cours relatives à l'emploi, notamment des programmes pour les travailleurs adultes et les travailleurs déplacés. Cette étude, baptisée « Évaluation de référence du WIA », porte bien son nom. Le rapport principal consacre une cinquantaine de pages d'annexes à des questions méthodologiques telles que la pondération des observations et l'imputation des données manquantes (Rotz et al., 2017). Il aboutit sans détour à des conclusions décevantes.

Comme pour l'évaluation du programme Job Corps, les 28 sites d'étude ont été choisis aléatoirement. Afin d'améliorer la représentativité nationale, la sélection des sites a été stratifiée selon 6 régions, un nombre prédéfini étant retenu dans chaque région (Fortson et al., 2017, p. 13).

³¹ Contrairement à la plupart des études examinées ici, celle-ci comportait trois groupes de traitement, qui correspondaient aux trois niveaux de service prévus par la WIA (Fortson et al. 2017, Figure 1) :

1. Un groupe s'est vu proposer uniquement des « services de base », qui « consistaient principalement en des informations et des outils en ligne pour aider les clients à planifier leur carrière et à trouver un emploi ».

31 De légères variations par rapport à la perfection ont été constatées. Les bureaux comptant moins de 100 clients par an, qui représentaient ensemble 2 % des clients à l'échelle nationale, ont été exclus pour des raisons de coût. Deux sites initialement sélectionnés ont refusé de participer et ont été remplacés.

2. Le deuxième groupe a eu accès à ces services ainsi qu'à des « services intensifs », qui « comprenaient des évaluations, des ateliers, une aide à la recherche d'emploi, l'élaboration de plans de carrière et de services, des conseils de carrière individuels et une gestion de cas, un placement dans des postes d'expérience professionnelle et une formation préprofessionnelle à court terme ».

3. Le groupe « WIA complet » a eu accès à tout ce qui précède ainsi qu'à des services de formation « conçus pour les préparer à des emplois dans des domaines à forte demande », qu'ils soient fournis par des organismes de formation professionnelle, des collèges communautaires ou des employeurs (formation en cours d'emploi).

La randomisation s'est déroulée de fin 2010 à début 2012. Dans le contexte de la crise financière, le chômage était en baisse mais restait élevé, à 8-9 %. 2 974 personnes ont intégré le volet « Adultes » et 1 983 le volet « Travailleurs déplacés » de l'expérience (Fortson et al., 2017, tableau III.2).

Comme dans l'évaluation du JTPA, des facteurs ont contribué à limiter l'efficacité des traitements. Les restrictions budgétaires ont parfois entraîné l'indisponibilité de certaines formations (Fortson et al., 2017, p. xxi). Les participants pouvaient également choisir de ne pas suivre de formation, car ils étaient libres de dépenser leurs bons d'achat. De plus, pour beaucoup, le WIA n'était pas la seule option de formation professionnelle : des sources privées, caritatives ou commerciales, étaient également disponibles.

Pour toutes ces raisons, durant les trois premiers trimestres de l'étude, les écarts de participation à la formation professionnelle entre les trois groupes étaient modestes : 34 % du groupe principal ont bénéficié d'une formation, contre 41 % dans le groupe intensif et 50 % dans le groupe WIA complet. Dans ce dernier, les dépenses de formation s'élevaient en moyenne à 3 223 \$, dont 1 521 \$ à la charge du gouvernement et le reste à la charge des participants (Fortson et al., 2017, tableau VIII.1). En ajoutant 888 \$ pour d'autres services, le coût pour le gouvernement atteint 2 409 \$ par participant, un montant bien inférieur à la moyenne observée pour les programmes de plus petite envergure (voir tableau 1 et section 6.5.1).

La figure 7 synthétise les principaux résultats. Elle utilise les données sur l'emploi et les salaires issues du Répertoire national officiel des nouvelles embauches (voir §3.3). Elle montre qu'avec la reprise économique, l'emploi et les salaires ont progressé dans les trois groupes étudiés, concernant les travailleurs adultes et les travailleurs licenciés.

Le groupe bénéficiant de l'intégralité du programme WIA – avec accès à une formation professionnelle soutenue par le WIA – a enregistré des taux d'emploi et des revenus inférieurs au cours des deux premiers trimestres, probablement parce que certains participants étaient en formation. Cependant, cet investissement en temps n'a pas clairement porté ses fruits sur les trois années de suivi. Au cours de la dernière année, pour les travailleurs déplacés, on observe une légère tendance à croire que les participants du groupe WIA complet ont gagné davantage que ceux du groupe bénéficiant d'un niveau de services inférieur (services de base et intensifs, graphique en bas à droite). Cet impact positif pourrait encore être pris en compte si les quatre graphiques présentaient une évolution cohérente de cet impact positif, quel que soit le résultat ou la population étudiée. Or, ce n'est pas le cas.

L'évaluation WIA s'avère malheureusement peu efficace pour détecter les effets de la formation professionnelle. La différence entre le groupe traité et le groupe témoin concernant le recours à la formation était modeste, de l'ordre de 16 points de pourcentage. De plus, les dépenses consacrées à la formation étaient relativement faibles. La formation dispensée a peut-être stimulé l'emploi et les revenus, mais lorsque l'ITT est statistiquement indiscernable de zéro, le LATE l'est également.

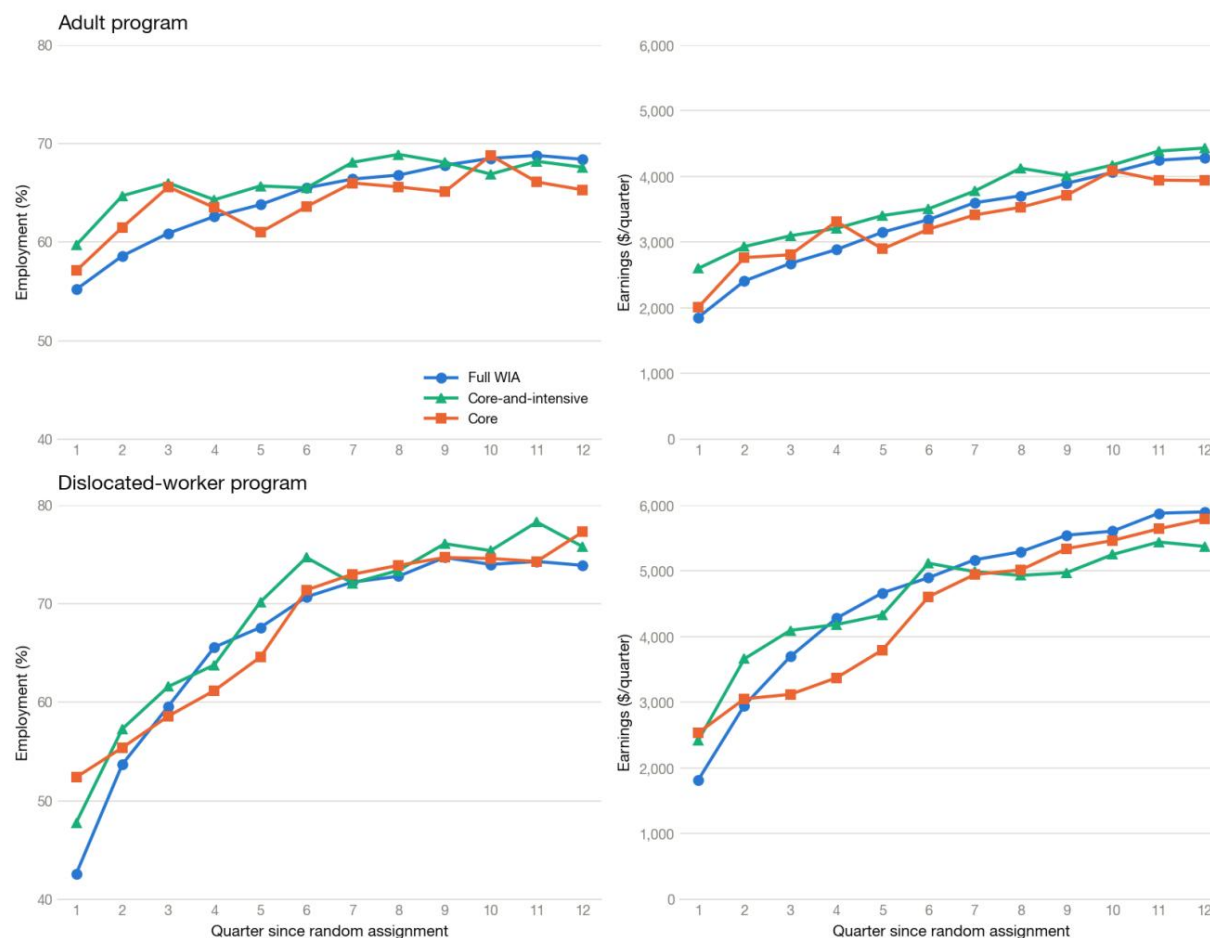


Figure 7. Emploi et revenus par trimestre dans les trois volets de l'évaluation « Gold Standard » de la loi sur l'investissement dans la main-d'œuvre (Workforce Investment Act), pour les programmes destinés aux adultes et aux travailleurs déplacés. Remarques : D'après Rotz et al. (2017), tableaux G.VI.1, G.VI.2, H.VI.1 et H.VI.2. Toutes les données proviennent de sources officielles. Seul le groupe ayant bénéficié de la totalité du programme WIA a reçu une formation professionnelle dans le cadre de la loi sur l'investissement dans la main-d'œuvre. Certains participants du groupe « formation de base et intensive » ont reçu une formation préprofessionnelle de courte durée. Le groupe « formation de base » n'a reçu aucune formation.

4.5 Principales évaluations randomisées américaines : résumé

L'histoire du leadership américain dans la recherche de la vérité empirique sur les programmes d'emploi est marquée par l'ambition et la déception. À notre connaissance, aucun autre gouvernement national n'a consacré autant d'énergie à une évaluation rigoureuse de l'efficacité d'un ensemble de programmes sociaux. Malheureusement, la conclusion, du moins d'après les essais nationaux du JTPA, du Job Corps et du WIA, est que ces programmes n'ont pas eu d'impact significatif sur le marché du travail. Le JTPA, à lui seul, a produit des effets positifs sur certains groupes cibles, les adultes. Dans l'étude du WIA notamment, le faible taux de participation a rendu plus difficile la détection de bénéfices réels pour les personnes ayant suivi une formation. Mais ce faible taux de participation n'est pas un échec que nous devons ignorer.

Comme nous le verrons au paragraphe 7, certains programmes privés de plus petite envergure ont obtenu de bien meilleurs résultats. Cela soulève des questions quant à la possibilité d'étendre ces succès, aux types de financements et de modes de gestion nécessaires.

5 études de « randomisation par les juges »

Certains chercheurs spécialisés dans les interventions sur le marché du travail ont emprunté une technique aux criminologues. Il s'agit de la « randomisation par juge » (Martin, Annan et Forst, 1993 ; Kling, 2006), même en l'absence de juges. Dans notre contexte, l'idée est la suivante : les personnes participant à un programme d'aide à l'emploi sont affectées aléatoirement à un conseiller au sein d'un même bureau pour l'emploi. Si certains conseillers sont plus enclins à orienter leurs clients vers une formation en milieu professionnel, cela constitue une expérience naturelle de formation en milieu professionnel. Ici, la propension d'un conseiller à attribuer un traitement particulier représente l'instrument de mesure. Elle remplace le tirage à pile ou face d'une expérience randomisée.

Cette section passe en revue les deux études de randomisation des juges en formation dont nous avons connaissance. L'une a été menée aux États-Unis, l'autre au Danemark. Comme nous le verrons, il n'est pas évident a priori que l'étude américaine se rapproche de l'idéal d'un traitement randomisé. Dans de nombreuses études impliquant de véritables juges, l'affectation est effectivement aléatoire, même si elle comporte une certaine marge de manœuvre pour influencer cette affectation (par exemple, Kling 2006 ; Green et Winik 2010 ; Loeffler 2013 ; Aizer et Doyle 2015 ; Bhuller et al.). 2020). Le brouillage délibéré est rare dans la formation professionnelle. De toute évidence, les gouvernements veillent davantage à ce que la justice soit impartiale qu'à ce que l'affectation des dossiers par les agents du chômage le soit également.

5.1 Hyman (2018), « Les travailleurs déplacés peuvent-ils être reconvertis ?

Preuves tirées d'une affectation quasi aléatoire à une aide à l'ajustement commercial »

Pendant des décennies, la politique américaine a consisté à accroître l'intégration économique avec les autres pays tout en aidant les personnes ayant perdu leur emploi face à la concurrence étrangère. À partir de 1962, le programme d'aide à l'ajustement commercial (TAA) prenait en charge les coûts de formation professionnelle de ces travailleurs déplacés. Le TAA finançait également l'assurance chômage pendant la formation, étendant la couverture habituelle de 26 semaines à trois ans, et offrait une aide à la recherche d'emploi et à la relocalisation (Collins 2014). (Ce programme est dans l'impasse législative depuis 2022, le Congrès n'ayant pas renouvelé son autorisation. Pour l'instant, le compromis politique qui sous-tendait la promotion du commerce est rompu.)

La principale limite de l'étude, de notre point de vue, est que la prolongation de l'assurance chômage était conditionnée par le fait d'être en formation, ce qui rend impossible de distinguer statistiquement les effets de la formation et du soutien au revenu.

Bien que la majeure partie du programme ait été gérée par les collectivités locales, un bureau du ministère du Travail déterminait quels travailleurs étaient éligibles, c'est-à-dire ceux qui avaient vraisemblablement perdu leur emploi à cause de la concurrence étrangère ou de la délocalisation. En réalité, le ministère du Travail ne prenait pas sa décision au cas par cas, mais au cas par cas, usine par usine. Si un seul travailleur licencié d'une usine bénéficiait de l'aide du programme TAA, tous les travailleurs de la même usine qui en feraient la demande au cours des trois années suivantes en bénéficiaient également (Hyman 2018).

Hyman (2018) étudie si cette aide permet aux personnes de se remettre d'un licenciement. Son analyse repose sur deux hypothèses : 1) parmi les centaines d'enquêteurs du ministère du Travail chargés de déterminer quels licenciements et fermetures étaient dus à la concurrence étrangère et justifiaient donc une aide au titre de cette aide, certains se sont montrés plus indulgents ; et 2) l'attribution des dossiers aux enquêteurs était de fait arbitraire, ou pouvait être perçue comme telle par une analyse ³² La première hypothèse est assez facile à démontrer, car économétrique appropriée. Ce dernier point sera examiné. La seconde hypothèse ne peut être étayée que par des preuves indirectes, car le processus d'attribution était peu transparent et rien ne laisse penser qu'il s'agissait d'un tirage au sort. Hyman cite cette description du ministère du Travail :

Les dossiers TAA sont attribués principalement en fonction de la charge de travail des enquêteurs, ainsi que de leur expérience antérieure au sein d'une entreprise ou d'un secteur d'activité. Les congés du personnel ou d'autres problèmes d'organisation peuvent également être pris en compte.

L'attribution des dossiers en fonction des disponibilités et des congés peut s'avérer aléatoire : deux cas similaires pourraient facilement être confiés à des enquêteurs différents. L'attribution fondée sur l'expérience antérieure n'est pas aléatoire, car elle peut induire des associations durables entre le secteur d'activité d'une personne et les enquêteurs qui lui ont confié des dossiers.

Pour réaliser son analyse, Hyman (2018) combine deux ensembles de données. Il obtient les données relatives à toutes les demandes d'indemnisation pour incapacité de travail (TAA) déposées entre 1974 et 2016. Pour chaque demande, l'ensemble de données indique l'employeur (ancien) qui prétend avoir subi des pressions de la concurrence étrangère, ainsi que l'enquêteur du ministère du Travail ayant statué sur le dossier. Hyman relie ces demandes à la base de données LEHD (Longitudinal Employer-Household Dynamics) du Bureau du recensement, qui regroupe les données des programmes d'assurance chômage des États. Cela lui permet de suivre si un travailleur a occupé – puis perdu ou quitté – un emploi dans une usine ayant fait l'objet d'une demande d'indemnisation ; si les anciens employés de l'usine ont obtenu cette indemnisation ; et combien de temps les bénéficiaires et les non-bénéficiaires ont travaillé et gagné au cours des années suivantes. Pour ce projet, 24 États et le district de Columbia ont autorisé l'accès à leurs contributions à la base de données LEHD. L'échantillon final comprend certains États de la Rust Belt (Illinois, Indiana, Pennsylvanie) ainsi que la Caroline du Sud, où l'industrie textile était autrefois un important employeur. L'instantané LEHD utilisé dans l'étude commence entre 1985 et 2002, selon l'État, et s'arrête en 2011.

Hyman (2018) illustre un schéma quasi universel en économie empirique : la variation du traitement n'étant pas aléatoire, des complications analytiques sont introduites dans l'espoir de l'orienter dans ce sens. Concrètement, lorsqu'une entreprise apprend que, suite à une requête d'un ancien employé, un enquêteur du ministère du Travail a certifié l'éligibilité de l'ensemble de son personnel au programme TAA, elle pourrait se sentir plus libre de licencier davantage d'employés. Or, les employés licenciés ultérieurement pourraient différer systématiquement de ceux licenciés plus tôt : ils sont peut-être plus qualifiés et plus précieux. En revanche, leurs homologues du groupe témoin – des personnes qui ont également échappé aux premiers licenciements, mais dans des usines n'ayant pas bénéficié du programme TAA – pourraient ne pas perdre leur emploi, du moins pas dans les mêmes conditions.

Dans les études de randomisation par les juges (par exemple, Aizer et Doyle, 2015), l'instrument utilisé est le taux d'approbation par validation croisée (méthode du « leave-one-out »), recalculé pour chaque demande. Cette méthode de « jackknife » permet de réduire l'endogénéité de l'instrument liée aux spécificités de chaque demande.

Pour éviter d'introduire un tel décalage, Hyman (2018) limite son échantillon aux personnes qui avaient déjà été licenciées au cours des 12 mois précédant la décision officielle concernant le soutien TAA.

Afin de restreindre davantage les groupes de comparaison et, sans doute, de les rendre plus pertinents, Hyman ajoute de nombreux contrôles : l'année et le trimestre de dépôt de la requête TAA ; le secteur d'activité (un facteur important compte tenu de la déclaration citée précédemment selon laquelle les enquêteurs peuvent avoir tendance à se spécialiser dans certains secteurs) ; le type d'organisation ayant déposé la requête (entreprise, syndicat, association de travailleurs ou bureau d'orientation professionnelle d'un État) ; et si la demande concernait un traitement spécial en lien avec l'Accord de libre-échange nord-américain (ALENA). De plus, Hyman inclut des contrôles relatifs aux caractéristiques démographiques telles que l'origine ethnique et le sexe, ainsi qu'à certains aspects de la formation et de l'expérience professionnelle, comme l'obtention du diplôme d'études secondaires et le salaire avant le licenciement.

L'espoir est qu'à mesure que l'on ajoute des facteurs de contrôle et que l'on restreint les domaines de comparaison, une certaine variation dans la probabilité d'accès au programme TAA subsiste et soit assimilable au hasard. Autrement dit, l'espoir est que, lorsqu'on compare les parcours professionnels de deux femmes blanches, titulaires d'un diplôme d'études secondaires, qui ont quitté leur emploi dans différentes usines automobiles de l'Indiana juste avant le dépôt des demandes de programme TAA pour ces usines durant l'été 1998, les autres différences dans leur vie, non prises en compte, n'aient pas influencé leur carrière, même si, en raison de différences dans la clémence des enquêteurs, l'une avait plus de chances d'obtenir le programme TAA. Dès lors, toute différence observée dans leurs carrières ultérieures est attribuable au programme TAA.

Il n'est pas évident que cette approche fonctionne, c'est-à-dire que l'ajout de nombreux témoins permette d'éliminer la plupart des différences non liées au traitement tout en préservant des différences arbitraires liées au traitement, constituant ainsi une expérience naturelle. Hyman (2018) réalise plusieurs tests statistiques pour valider ou réfuter cette stratégie. L'un de ces tests porte sur la pertinence de l'instrument : même avec tous les témoins, le taux d'approbation moyen d'un investigateur permet-il toujours de prédire ses décisions individuelles ? C'est le cas : chaque augmentation de 10 points de pourcentage du taux d'approbation moyen d'un investigateur accroît la probabilité d'approbation dans les cas individuels de 6 points, et ce de manière statistiquement très significative (Hyman 2018, tableau 4, première ligne).

Deux tests permettent ensuite d'évaluer indirectement l'exclusion, c'est-à-dire de déterminer si, après prise en compte de tous les effets fixes et autres variables de contrôle, la clémence de l'enquêteur n'a aucun lien statistique avec l'emploi et les salaires après traitement, sauf par le biais du traitement TAA. Lorsque cette condition est remplie, la clémence joue effectivement le rôle d'un tirage à pile ou face dans une expérience classique. Le premier test part du constat que si la clémence de l'enquêteur désigné était véritablement aléatoire, elle serait non corrélée avec toutes les autres variables de l'étude. Techniquement, cette propriété n'est pas nécessaire pour que les estimations d'impact soient non biaisées ; toutefois, des indices indirects en ce sens sont rassurants. En l'espèce, Hyman (2018, tableau 3, col. 5-6) constate que la clémence est quasiment indépendante des caractéristiques démographiques, mais qu'elle est plus élevée lorsque les employés ont une plus grande ancienneté ou une rémunération plus élevée, lorsque le taux de chômage du comté est plus élevé, lorsque l'enquêteur est moins expérimenté ou lorsque l'établissement appartient à un secteur plus exposé à la concurrence étrangère. Hyman (2018, p. 20) soutient que même les associations statistiquement significatives ne le sont pas en pratique. Par exemple, une décennie d'expérience supplémentaire n'augmente le taux d'approbation d'un chercheur que de 0,004 point de pourcentage (dans une régression multivariée, tableau 3, col. 6).

Lorsqu'on teste les corrélations impliquant de nombreuses variables, certaines peuvent être déclarées statistiquement significatives par hasard. C'est pourquoi Hyman (2018, tableau 3, col. 6) effectue un test F pour vérifier l'hypothèse que toutes les corrélations sont en réalité nulles. Le résultat est une p-valeur de 0,18, supérieure au seuil conventionnel de 0,05 pour la signification statistique. Cependant, pour ce type de vérification, des seuils bas sont tout sauf conservateurs (Roodman 2009, p. 142). Il n'y a que 18 % de chances d'observer une corrélation aussi forte si le traitement était indépendant de tous les contrôles, ce qui n'est guère rassurant.

Néanmoins, dans un contexte où la perfection n'est ni attendue ni formellement requise, les tests d'aléatoire effectif produisent des résultats plutôt rassurants. Certaines corrélations sont probablement réelles, mais faibles.

Le second test indirect d'exclusion est un contrôle par placebo du type préconisé par Heckman (voir §4.1.4). Si la variable de clémence de l'investigateur génère une expérience naturelle valide, elle ne devrait sembler affecter que les résultats post-traitement, et non ceux pré-traitement.

Ce contrôle par placebo, ainsi que les principaux résultats de l'étude, sont illustrés par les graphiques de la figure 8. Les courbes bleues représentent l'impact estimé de l'accès au programme TAA pour les anciens employés d'une usine donnée, tant sur le nombre de trimestres travaillés par an que sur leur revenu annuel. Au cours des dix années précédant la décision d'accès au programme TAA pour une usine, les courbes bleues restent proches de zéro ; en effet, la plupart des intervalles de confiance à 90 %, représentés par des pointillés, incluent facilement zéro. Le programme ne semble pas avoir d'effet sur les résultats avant qu'il ne puisse en avoir. Ensuite, au cours des deux premières années suivant la décision relative au programme TAA, l'emploi et les revenus diminuent dans le groupe concerné ; cela correspond à une entrée en formation plutôt qu'à une recherche immédiate, et parfois à la découverte, d'un nouvel emploi. Par la suite, les effets deviennent positifs, ce qui suggère que le programme TAA aide les personnes à trouver un emploi plus nombreux et/ou de meilleure qualité. Cependant, les avantages s'estompent à partir de la 10^e année. (Première rangée de la figure 8.) Hyman (2018, p. 24) calcule que les travailleurs gagnent 20 mois de travail supplémentaires et 50 000 \$ de revenus sur 10 ans.

La plausibilité de la tendance générale — une phase relativement stable, une brève et brutale chute, une nette remontée, puis un déclin à long terme — renforce la crédibilité des résultats. Si l'avantage apparent de l'analyse TAA n'est qu'un bruit statistique, cette tendance particulière serait surprenante.

Après avoir établi le résultat principal, Hyman (2018) analyse les données sous différents angles afin de mieux comprendre quels groupes de personnes bénéficient le plus de l'aide et comment. Une question majeure concerne l'impact de la formation dispensée par le programme TAA. Les données ne précisent pas la durée des formations financées par le TAA. Elles fournissent toutefois des moyennes par État et par trimestre. En segmentant l'échantillon selon que la personne travaillait dans un État situé au-dessus ou en dessous du 50^e ^{ème} percentile juste avant son licenciement, il apparaît que les revenus ne diminuent pas pour les personnes originaires des États et des trimestres où elles ont consacré plus de temps à la formation (Hyman 2018, Figure 10). En revanche, dans le groupe complémentaire, les revenus augmentent initialement – ce qui correspond à un temps de formation plus court – puis diminuent. Cette caractérisation des fluctuations n'a pas fait l'objet d'une analyse formelle. ³³ Et ça

33Un autre obstacle à l'évaluation de la robustesse de ce résultat est l'absence d'intervalles de confiance, par souci de lisibilité, dans la figure 12 de Hyman (2018). Pour les erreurs standard, les notes de la figure renvoient le lecteur à une

Cette analyse repose sur une segmentation endogène de l'échantillon : elle montre peut-être surtout que, là où les nouveaux emplois étaient plus difficiles à trouver, les personnes ont davantage opté pour la formation. Globalement, elle suggère légèrement que la phase de formation du dispositif d'aide à l'emploi (TAA) contribue aux avantages apparents de ce dernier. Cependant, comme le maintien des allocations chômage est conditionné par la participation à une formation professionnelle, le chèque ne permet toujours pas de distinguer les effets de la formation de ceux de l'aide au revenu concomitante.

Les résultats concernant la mobilité spatiale et sectorielle sont également importants et connexes. Dorn (2009) divise les États-Unis en « zones de navettage » qui correspondent aux marchés du travail locaux. Le Système de classification des industries de l'Amérique du Nord (SCIAN), à deux chiffres, catégorise les principaux secteurs d'activité tels que la construction et le commerce de détail. Hyman (2018), figure 12 (reproduite ici dans la deuxième ligne de la figure 8), constate que les personnes pouvant bénéficier de l'aide à la transition professionnelle (TAA) avaient 20 à 40 points de probabilité supplémentaires de déménager dans une nouvelle zone de navettage. Les effets étaient similaires pour un changement de secteur d'activité. Si ces résultats se confirment, ils sont significatifs : l'aide à la transition professionnelle aide effectivement les personnes à accéder à de meilleurs environnements de travail après une crise économique.

Dans quelle mesure faut-il croire ces conclusions ? Il y a des raisons de douter et des raisons de...

La fiabilité de cette étude est sujette à caution. Plusieurs raisons incitent à la méfiance : il s'agit d'une étude vaste, complexe, opaque et non expérimentale. Cette complexité augmente le risque d'erreurs et accroît la marge de manœuvre dans la définition de l'échantillon et le traitement des données, ce qui favorise la manipulation des résultats, consciente ou non. Le traitement n'était pas randomisé. L'étude n'a pas été publiée et les données de suivi sont confidentielles, ce qui rend toute analyse approfondie difficile.

En revanche, les contrôles d'exclusion et de pertinence donnent des résultats plutôt rassurants.

Et le profil temporel, y compris l'absence d'impact pendant la période précédant le traitement et la chute brutale mais temporaire, est difficile à expliquer comme de simples artefacts de causalité inverse ou de confusion par une troisième variable.

Nous jugeons les résultats globaux plausibles, mais nous ne savons pas dans quelle mesure ils sont attribuables à la formation plutôt qu'au soutien financier reçu pendant la formation.

L'annexe en ligne semble indisponible. Les astérisques de signification figurent dans le tableau 8 correspondant, mais ce tableau omet la dernière année (l'année 10), dont la faible estimation ponctuelle pour le groupe inférieur à la médiane influence particulièrement l'ampleur de la tendance à la baisse dans ce groupe.

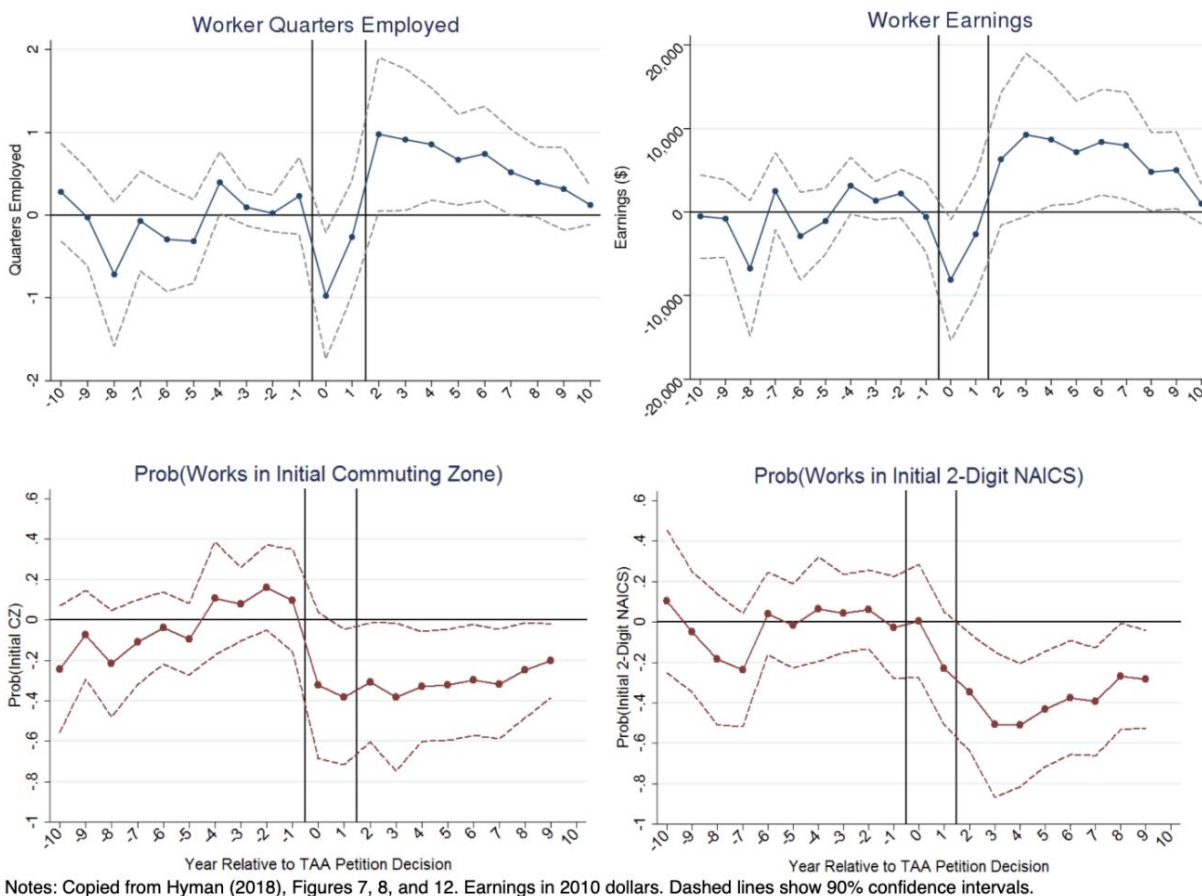


Figure 8. Impacts estimés de la disponibilité de l'aide à l'ajustement commercial, d'après Hyman (2018)

Remarques : Données copiées de Hyman (2018), figures 7, 8 et 12. Revenus en dollars de 2010. Les lignes pointillées indiquent les intervalles de confiance à 90 %.

5.2 Humlum, Munch et Rasmussen (2025), « Qu'est-ce qui fonctionne pour les chômeurs ? Données issues d'affectations quasi aléatoires de travailleurs sociaux »

L'autre étude de randomisation des juges de cette revue diffère de celle de Hyman (2018) à quelques égards : elle se déroule au Danemark.

Les décisions relatives aux options de formation sont prises individuellement, et non pour l'ensemble de l'équipe d'entreprises.

Les décisions sont prises par les conseillers des agences locales pour l'emploi plutôt que par un organisme central d'agence.

Les travailleurs sociaux ne se contentent pas d'autoriser le financement ; ils conçoivent un programme pour chaque client, qui, dans environ la moitié des cas, comprend une formation en salle de classe, une formation en milieu de travail, ou les deux (Humlum, Munch et Rasmussen 2025, tableau 2). Cette variation dans deux traitements

permet à Humlum, Munch et Rasmussen – ci-après « HMR » – d'estimer l'impact de chacun.

- L'article présente un argument plus convaincant concernant l'attribution des dossiers aux travailleurs sociaux contient une composante isolable qui est aussi bonne que du hasard.

Malgré la crédibilité apparente de cette expérience naturelle, les résultats de l'étude, pris dans leur ensemble, sont surprenants, ce qui nous amène à leur accorder un peu moins d'importance.

Conformément à la législation et aux normes danoises (HMR, § 2.2), le HMR distingue deux grandes catégories de formation. La formation en salle de classe englobe l'enseignement général, l'acquisition des compétences de base en recherche d'emploi et les services d'accompagnement destinés à aider les personnes à choisir un nouveau type d'emploi et à trouver un employeur. Elle inclut également la formation professionnelle. La formation en situation de travail, quant à elle, comprend les stages et les contrats à durée déterminée subventionnés.

Cet article exploite une particularité de l'administration du programme danois d'assurance chômage : au sein des agences pour l'emploi, les demandeurs d'emploi étaient souvent affectés à des conseillers en fonction de leur jour de naissance. Par exemple, dans une agence locale, les demandeurs nés la première semaine du mois étaient suivis par un conseiller, ceux nés la deuxième semaine par un autre, et ainsi de suite. HMR propose une illustration à l'aide de données simulées, reproduites ci-dessous (figure 9). Dans la simulation, 58 % des dossiers sont attribués selon la « règle » de la semaine de naissance. Cela crée des pics distincts pour chaque conseiller dans la répartition des dossiers par jour de naissance. Dans les données réelles, que les restrictions de confidentialité nous empêchent de consulter, la situation serait parfois plus simple, parfois moins.

34

L'étude 34HMR emprunte cette stratégie à Cederlöf, Söderström et Vikström (2025), qui présente de nombreuses similitudes et se déroule en Suède. Cette dernière étude, cependant, s'attache à évaluer l'efficacité des travailleurs sociaux pour réduire la durée du chômage de leurs clients et les aider à trouver des emplois mieux rémunérés, indépendamment des effets de la formation. Elle indique que les travailleurs sociaux qui privilégiaient la formation étaient en moyenne moins efficaces (Cederlöf, Söderström et Vikström 2025, tableau 8, ligne « Soutien »). Toutefois, ce résultat n'est pas expérimental.



Figure 9. Simulation : Nombre de clients affectés à quatre travailleurs sociaux dans un centre d'emploi, par jour du mois de naissance, d'après Humlum, Munch et Rasmussen (2025)

Remarques : Copié de Humlum, Munch et Rasmussen (2025), Figure 1.

Puisqu'il est difficile de comprendre en quoi naître le 7^{ème} L'événement du mois aurait un impact sur la carrière. Contrairement à la prédisposition à naître pour ^{ème} autrement que par l'affectation à des travailleurs sociaux ayant des compétences différentes la formation, cette configuration offre de bonnes chances de produire une expérience naturelle solide.

Une difficulté réside dans le fait que l'attribution des demandeurs d'emploi par date de naissance est loin d'être universelle, et les modalités précises de cette pratique ne sont pas documentées : dans quelle mesure les différents centres pour l'emploi l'appliquaient-ils, et comment répartissaient-ils les jours du mois en blocs ? HMR combine donc des techniques automatiques et manuelles pour estimer les règles réellement appliquées. Pour chaque centre pour l'emploi – ou unité au sein du centre où il semble que les différents groupes d'âge soient gérés séparément – HMR suppose provisoirement que le conseiller qui prend en charge le plus de personnes nées, par exemple, le 12^{ème} du mois, est ^{ème} du mois, c'est en fait le responsable de l'attribution. Une fois ces estimations effectuées, des graphiques comme ceux de la figure 9 sont établis. Si, comme sur la figure, un schéma de blocage est évident, le bureau est conservé dans l'échantillon et l'affectation en fonction de la date est déduite du graphique. Après ce processus de construction, l'échantillon HMR comprend 167 222 épisodes de chômage vécus par 127 713 personnes, qui ont consulté l'un des 536 travailleurs sociaux entre 2012 et 2017 (tableau A.1 du HMR).

Ce processus de construction laisse une certaine marge de manœuvre – probablement modeste – à l'exercice non observé d'un pouvoir discrétionnaire. Et dans la mesure où les jugements des chercheurs sont erronés – là encore, probablement modestement –, ils amoindrissent la pertinence des instruments basés sur la date de naissance, c'est-à-dire les tendances du travailleur social prédit à attribuer des formations en salle et/ou en milieu de travail. Enfin, même avec une compréhension parfaite du processus d'attribution des travailleurs sociaux, les effets observés dans ce dispositif pourraient être dus à d'autres facteurs concernant les travailleurs sociaux qui privilégient certains programmes, ce qui enfreindrait l'hypothèse d'exclusion. Cependant, le modèle HMR présente une série de tests assez convaincants suggérant que l'hypothèse d'exclusion est respectée.

Tout comme dans l'étude de Hyman (2018), les auteurs vérifient rigoureusement la pertinence et l'indépendance des instruments. Concernant la pertinence des instruments : après prise en compte des caractéristiques démographiques des clients et des effets fixes temporels et régionaux, les clients dont la prise en charge est prédite, en fonction de leur date de naissance, comme nécessitant une formation en salle de classe (présence 10 points plus élevée), bénéficient effectivement d'un temps de formation en salle de classe plus important (3,8 points supplémentaires). Cette association entre la formation en salle de classe et l'instrument utilisé est hautement significative, statistiquement parlant. Pour la formation en situation de travail, l'association est deux fois moins importante (2,1 points), mais reste significative selon les critères habituels (HMR, tableau C.2).

Concernant l'indépendance des instruments : à l'instar de Hyman (2018), la méthode HMR vérifie si les instruments sont liés à des caractéristiques démographiques et à d'autres variables contrôlées. Rappelons que cette vérification n'est pas techniquement requise pour éviter les biais, mais elle peut renforcer l'idée que l'instrument est quasiment aléatoire. Étant donné que 65 variables de contrôle figurent dans le tableau des résultats de ces vérifications, sous l'hypothèse nulle d'une absence de corrélation entre les instruments et ces variables, on devrait observer environ 6,5 corrélations significatives au seuil de 0,1 pour chaque instrument. On en compte 5 pour l'instrument relatif à la formation en salle et 12 pour celui relatif à la formation en situation de travail.

La principale conclusion de HMR est que, parmi les deux types de formation (en salle et en milieu professionnel), la première semble avoir aidé les personnes à trouver un emploi. Voir la figure 10 ci-dessous, adaptée du tableau D.1 de HMR. Suivre une formation en salle augmente l'emploi d'environ 26 heures par mois la deuxième année suivant l'inscription au chômage (HMR, tableau D.2). En divisant ce chiffre par 160 heures par mois en moyenne, on obtient une augmentation substantielle de 16 % de l'emploi, ce qui est parmi les plus importantes de cette étude. Le constat est le même si l'on considère l'emploi (le fait d'avoir un emploi) ou les revenus comme indicateur, plutôt que le nombre d'heures travaillées. HMR constate que suivre une formation en salle augmente l'emploi d'environ 20 points de pourcentage sept trimestres après le début de la période de chômage (HMR, figure D.2).

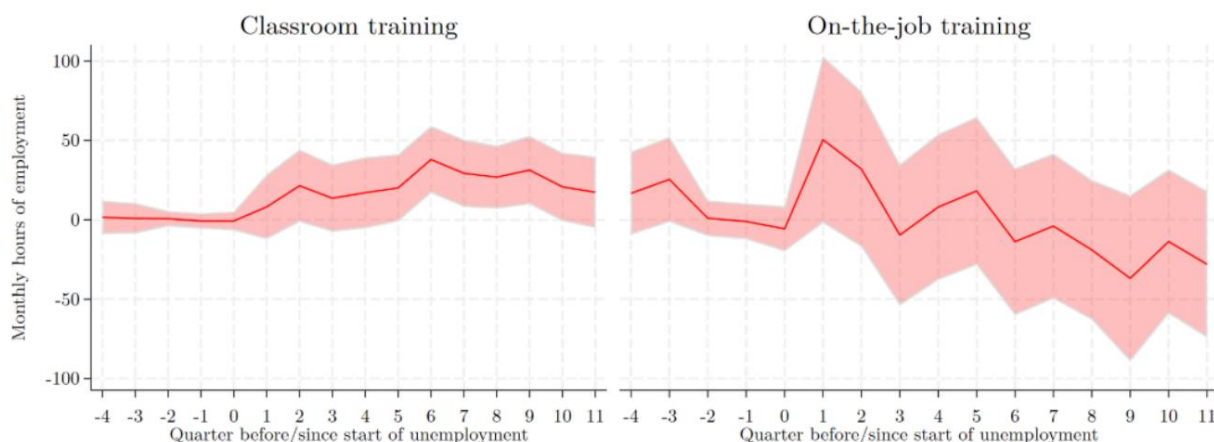


Figure 10. Impact de l'affectation à une formation en salle de classe ou à une formation en cours d'emploi sur les heures de travail par mois, par trimestre avant/depuis le début de l'épisode de chômage, d'après Humlum, Munch et Rasmussen (2025)

Remarques : Basé sur Humlum, Munch et Rasmussen (2025), tableau D.1. L'échantillon est composé de personnes entrées au chômage entre 2012 et 2017. Étant donné que le suivi s'arrête après 2019 (au début de la COVID-19), les périodes de chômage qui ont commencé en 2017 commencent à sortir de l'échantillon après le 7e trimestre.

En revanche, pour la formation en situation de travail, les intervalles de confiance sont environ deux fois plus larges et les estimations de l'impact central sont généralement statistiquement indiscernables de 0. Malgré cette imprécision, la puissance statistique est suffisante pour rejeter l'hypothèse selon laquelle la formation en salle et la formation en situation de travail auraient un impact similaire. Les valeurs p pour cette hypothèse se situent entre 0,03 et 0,22 pour les trimestres 6 à 11 (tableau D.1 du HMR), ce qui est faible. Il est donc peu probable que les deux traitements soient identiques et que la difficulté à détecter l'impact de l'un soit simplement due à une faible puissance statistique.

Pour mieux comprendre l'utilité de la formation en salle, HMR, à l'instar de Hyman (2018), subdivise l'analyse principale selon plusieurs dimensions. Une distinction est faite entre le retour à son ancien emploi et la recherche d'un nouvel emploi. Par ailleurs, les cours en salle sont divisés en « formations à la recherche d'emploi » (d'une durée moyenne de 19 jours) et en « formations aux compétences et services d'accompagnement ». Cette dernière catégorie regroupe les services d'accompagnement, tels que le conseil en emploi, avec la formation professionnelle. HMR explique que « si les cours de recherche d'emploi se distinguent aisément des autres cours, la frontière entre les formations axées sur les compétences et les formations d'accompagnement est moins nette ». Cela surprend, car a priori, la formation axée sur les compétences semble détonner par rapport au conseil et à l'aide à la recherche d'emploi. HMR poursuit en expliquant que la combinaison des formations axées sur les compétences et des formations d'accompagnement « contribue également à renforcer la puissance statistique de notre analyse ». Il s'agit d'un compromis entre la précision de la conception de ce qui est mesuré et la précision de sa mesure.

Le résultat de ces deux distinctions binaires – nouvel ou ancien emploi, cours de recherche d'emploi ou autres matières scolaires – est de quatre ordres. HMR constate qu'environ un an après le début du chômage, les cours de recherche d'emploi et les formations complémentaires augmentent le taux de réinsertion professionnelle.

professions. Aucune des deux ne l'augmente de manière fiable dans les anciennes professions (HMR, ³⁵ Cela suggère Figure 9). que la formation en salle de classe a aidé les travailleurs danois à se remettre de la perte d'emploi en les orientant vers de nouvelles voies professionnelles.

Pour approfondir cette histoire de résilience, HMR analyse si la formation a particulièrement aidé les personnes dont les emplois étaient les plus exposés à la délocalisation. HMR considère comme les moins exposés ceux dont les anciens emplois exigeaient une présence sur site ou un contact direct avec les clients – pensez aux vendeurs et aux mécaniciens automobiles, et non aux programmeurs informatiques et aux téléconseillers. Les estimations de l'impact pour les plus exposés sont plus de deux fois supérieures, bien que la différence ne soit pas statistiquement significative ($p = 0,14$ pour les heures travaillées, et plus faiblement pour les autres indicateurs).

L'analyse HMR décompose également les résultats par étape du processus de formation. Les effets de la formation en salle se manifestant dans les mois suivant l'inscription au système d'assurance chômage, les bénéficiaires passent par quatre étapes : absence d'affectation à une formation ; affectation à une formation ; participation à la formation ; et achèvement de la formation, après un abandon prématuré ou un suivi complet. L'analyse HMR (figure 6(b)) montre que la plus forte augmentation globale de l'emploi concerne les personnes ayant terminé leur formation, et ce, après 14 mois. Ceci corrobore globalement l'hypothèse selon laquelle la formation en salle est efficace, une fois terminée. Cependant, d'autres résultats de cette décomposition s'accordent moins bien avec la théorie. L'abandon prématuré d'une formation et le fait de ne jamais commencer la formation qui lui a été attribuée ont un impact individuel sur le chômage sensiblement équivalent à celui de la réussite de la formation.

36

L'interprétation des résultats du HMR est complexe. Les estimations sont élevées, comparables uniquement à celles du LATE concernant les programmes de formation sectoriels présentés ci-dessous. Comment expliquer que la formation en salle soit nettement plus efficace que la formation en situation de travail pour orienter les individus vers de nouveaux emplois ? Le HMR (§7.1) suggère que cela tient au fait qu'un retour en formation initiale est souvent nécessaire pour une requalification fondamentale. Or, il n'est pas certain que les cours danois destinés aux chômeurs soient si fondamentaux. Ils ne durent en moyenne que 52 jours (HMR, tableau 1). De plus, comme mentionné précédemment, les bénéfices sont sensiblement les mêmes pour ceux qui abandonnent prématurément ou qui ne suivent même pas les cours prévus. Parmi les matières enseignées en salle, la recherche d'emploi a, à bien des égards, un impact plus important, même si elle ne relève pas de la requalification (HMR, tableau H.7, col. 2). Par ailleurs, l'analyse combinant formation professionnelle et conseil en emploi, il est difficile de déterminer si la formation professionnelle a réellement un impact majeur, comme le prédit la théorie de la requalification.

Globalement, l'étude est difficile à évaluer. La méthodologie semble solide. De plus, l'affirmation d'une identification quasi expérimentale des impacts est plus convaincante que dans l'étude de Hyman (2018), grâce à la règle de l'anniversaire. Cependant, à l'instar de celle de Hyman (2018), l'étude demeure intrinsèquement opaque et complexe.

35 Sur 8 trimestres (du 0 au 7 après la randomisation) et deux traitements, l'impact sur le retour à l'emploi dans sa profession d'origine est significatif à $p < 0,05$ dans un cas : les services de recherche d'emploi au trimestre 0. Il est difficile de savoir si ce résultat significatif est un signal ou un bruit.

L'article 7 du règlement 36 HMR évalue les impacts à différentes étapes en estimant la probabilité γ ^{*s*} paramètres pour les γ ₁ ^{**} γ ₂ ^{**} γ ₃ ^{**} γ ₄ ^{**} γ ₅ ^{**} γ ₆ ^{**} γ ₇ ^{**} γ ₈ ^{**} γ ₉ ^{**} γ ₁₀ ^{**} γ ₁₁ ^{**} γ ₁₂ ^{**} γ ₁₃ ^{**} γ ₁₄ ^{**} γ ₁₅ ^{**} γ ₁₆ ^{**} γ ₁₇ ^{**} γ ₁₈ ^{**} γ ₁₉ ^{**} γ ₂₀ ^{**} γ ₂₁ ^{**} γ ₂₂ ^{**} γ ₂₃ ^{**} γ ₂₄ ^{**} γ ₂₅ ^{**} γ ₂₆ ^{**} γ ₂₇ ^{**} γ ₂₈ ^{**} γ ₂₉ ^{**} γ ₃₀ ^{**} γ ₃₁ ^{**} γ ₃₂ ^{**} γ ₃₃ ^{**} γ ₃₄ ^{**} γ ₃₅ ^{**} γ ₃₆ ^{**} γ ₃₇ ^{**} γ ₃₈ ^{**} γ ₃₉ ^{**} γ ₄₀ ^{**} γ ₄₁ ^{**} γ ₄₂ ^{**} γ ₄₃ ^{**} γ ₄₄ ^{**} γ ₄₅ ^{**} γ ₄₆ ^{**} γ ₄₇ ^{**} γ ₄₈ ^{**} γ ₄₉ ^{**} γ ₅₀ ^{**} γ ₅₁ ^{**} γ ₅₂ ^{**} γ ₅₃ ^{**} γ ₅₄ ^{**} γ ₅₅ ^{**} γ ₅₆ ^{**} γ ₅₇ ^{**} γ ₅₈ ^{**} γ ₅₉ ^{**} γ ₆₀ ^{**} γ ₆₁ ^{**} γ ₆₂ ^{**} γ ₆₃ ^{**} γ ₆₄ ^{**} γ ₆₅ ^{**} γ ₆₆ ^{**} γ ₆₇ ^{**} γ ₆₈ ^{**} γ ₆₉ ^{**} γ ₇₀ ^{**} γ ₇₁ ^{**} γ ₇₂ ^{**} γ ₇₃ ^{**} γ ₇₄ ^{**} γ ₇₅ ^{**} γ ₇₆ ^{**} γ ₇₇ ^{**} γ ₇₈ ^{**} γ ₇₉ ^{**} γ ₈₀ ^{**} γ ₈₁ ^{**} γ ₈₂ ^{**} γ ₈₃ ^{**} γ ₈₄ ^{**} γ ₈₅ ^{**} γ ₈₆ ^{**} γ ₈₇ ^{**} γ ₈₈ ^{**} γ ₈₉ ^{**} γ ₉₀ ^{**} γ ₉₁ ^{**} γ ₉₂ ^{**} γ ₉₃ ^{**} γ ₉₄ ^{**} γ ₉₅ ^{**} γ ₉₆ ^{**} γ ₉₇ ^{**} γ ₉₈ ^{**} γ ₉₉ ^{**} γ ₁₀₀ ^{**} γ ₁₀₁ ^{**} γ ₁₀₂ ^{**} γ ₁₀₃ ^{**} γ ₁₀₄ ^{**} γ ₁₀₅ ^{**} γ ₁₀₆ ^{**} γ ₁₀₇ ^{**} γ ₁₀₈ ^{**} γ ₁₀₉ ^{**} γ ₁₁₀ ^{**} γ ₁₁₁ ^{**} γ ₁₁₂ ^{**} γ ₁₁₃ ^{**} γ ₁₁₄ ^{**} γ ₁₁₅ ^{**} γ ₁₁₆ ^{**} γ ₁₁₇ ^{**} γ ₁₁₈ ^{**} γ ₁₁₉ ^{**} γ ₁₂₀ ^{**} γ ₁₂₁ ^{**} γ ₁₂₂ ^{**} γ ₁₂₃ ^{**} γ ₁₂₄ ^{**} γ ₁₂₅ ^{**} γ ₁₂₆ ^{**} γ ₁₂₇ ^{**} γ ₁₂₈ ^{**} γ ₁₂₉ ^{**} γ ₁₃₀ ^{**} γ ₁₃₁ ^{**} γ ₁₃₂ ^{**} γ ₁₃₃ ^{**} γ ₁₃₄ ^{**} γ ₁₃₅ ^{**} γ ₁₃₆ ^{**} γ ₁₃₇ ^{**} γ ₁₃₈ ^{**} γ ₁₃₉ ^{**} γ ₁₄₀ ^{**} γ ₁₄₁ ^{**} γ ₁₄₂ ^{**} γ ₁₄₃ ^{**} γ ₁₄₄ ^{**} γ ₁₄₅ ^{**} γ ₁₄₆ ^{**} γ ₁₄₇ ^{**} γ ₁₄₈ ^{**} γ ₁₄₉ ^{**} γ ₁₅₀ ^{**} γ ₁₅₁ ^{**} γ ₁₅₂ ^{**} γ ₁₅₃ ^{**} γ ₁₅₄ ^{**} γ ₁₅₅ ^{**} γ ₁₅₆ ^{**} γ ₁₅₇ ^{**} γ ₁₅₈ ^{**} γ ₁₅₉ ^{**} γ ₁₆₀ ^{**} γ ₁₆₁ ^{**} γ ₁₆₂ ^{**} γ ₁₆₃ ^{**} γ ₁₆₄ ^{**} γ ₁₆₅ ^{**} γ ₁₆₆ ^{**} γ ₁₆₇ ^{**} γ ₁₆₈ ^{**} γ ₁₆₉ ^{**} γ ₁₇₀ ^{**} γ ₁₇₁ ^{**} γ ₁₇₂ ^{**} γ ₁₇₃ ^{**} γ ₁₇₄ ^{**} γ ₁₇₅ ^{**} γ ₁₇₆ ^{**} γ ₁₇₇ ^{**} γ ₁₇₈ ^{**} γ ₁₇₉ ^{**} γ ₁₈₀ ^{**} γ ₁₈₁ ^{**} γ ₁₈₂ ^{**} γ ₁₈₃ ^{**} γ ₁₈₄ ^{**} γ ₁₈₅ ^{**} γ ₁₈₆ ^{**} γ ₁₈₇ ^{**} γ ₁₈₈ ^{**} γ ₁₈₉ ^{**} γ ₁₉₀ ^{**} γ ₁₉₁ ^{**} γ ₁₉₂ ^{**} γ ₁₉₃ ^{**} γ ₁₉₄ ^{**} γ ₁₉₅ ^{**} γ ₁₉₆ ^{**} γ ₁₉₇ ^{**} γ ₁₉₈ ^{**} γ ₁₉₉ ^{**} γ ₂₀₀ ^{**} γ ₂₀₁ ^{**} γ ₂₀₂ ^{**} γ ₂₀₃ ^{**} γ ₂₀₄ ^{**} γ ₂₀₅ ^{**} γ ₂₀₆ ^{**} γ ₂₀₇ ^{**} γ ₂₀₈ ^{**} γ ₂₀₉ ^{**} γ ₂₁₀ ^{**} γ ₂₁₁ ^{**} γ ₂₁₂ ^{**} γ ₂₁₃ ^{**} γ ₂₁₄ ^{**} γ ₂₁₅ ^{**} γ ₂₁₆ ^{**} γ ₂₁₇ ^{**} γ ₂₁₈ ^{**} γ ₂₁₉ ^{**} γ ₂₂₀ ^{**} γ ₂₂₁ ^{**} γ ₂₂₂ ^{**} γ ₂₂₃ ^{**} γ ₂₂₄ ^{**} γ ₂₂₅ ^{**} γ ₂₂₆ ^{**} γ ₂₂₇ ^{**} γ ₂₂₈ ^{**} γ ₂₂₉ ^{**} γ ₂₃₀ ^{**} γ ₂₃₁ ^{**} γ ₂₃₂ ^{**} γ ₂₃₃ ^{**} γ ₂₃₄ ^{**} γ ₂₃₅ ^{**} γ ₂₃₆ ^{**} γ ₂₃₇ ^{**} γ ₂₃₈ ^{**} γ ₂₃₉ ^{**} γ ₂₄₀ ^{**} γ ₂₄₁ ^{**} γ ₂₄₂ ^{**} γ ₂₄₃ ^{**} γ ₂₄₄ ^{**} γ ₂₄₅ ^{**} γ ₂₄₆ ^{**} γ ₂₄₇ ^{**} γ ₂₄₈ ^{**} γ ₂₄₉ ^{**} γ ₂₅₀ ^{**} γ ₂₅₁ ^{**} γ ₂₅₂ ^{**} γ ₂₅₃ ^{**} γ ₂₅₄ ^{**} γ ₂₅₅ ^{**} γ ₂₅₆ ^{**} γ ₂₅₇ ^{**} γ ₂₅₈ ^{**} γ ₂₅₉ ^{**} γ ₂₆₀ ^{**} γ ₂₆₁ ^{**} γ ₂₆₂ ^{**} γ ₂₆₃ ^{**} γ ₂₆₄ ^{**} γ ₂₆₅ ^{**} γ ₂₆₆ ^{**} γ ₂₆₇ ^{**} γ ₂₆₈ ^{**} γ ₂₆₉ ^{**} γ ₂₇₀ ^{**} γ ₂₇₁ ^{**} γ ₂₇₂ ^{**} γ ₂₇₃ ^{**} γ ₂₇₄ ^{**} γ ₂₇₅ ^{**} γ ₂₇₆ ^{**} γ ₂₇₇ ^{**} γ ₂₇₈ ^{**} γ ₂₇₉ ^{**} γ ₂₈₀ ^{**} γ ₂₈₁ ^{**} γ ₂₈₂ ^{**} γ ₂₈₃ ^{**} γ ₂₈₄ ^{**} γ ₂₈₅ ^{**} γ ₂₈₆ ^{**} γ ₂₈₇ ^{**} γ ₂₈₈ ^{**} γ ₂₈₉ ^{**} γ ₂₉₀ ^{**} γ ₂₉₁ ^{**} γ ₂₉₂ ^{**} γ ₂₉₃ ^{**} γ ₂₉₄ ^{**} γ ₂₉₅ ^{**} γ ₂₉₆ ^{**} γ ₂₉₇ ^{**} γ ₂₉₈ ^{**} γ ₂₉₉ ^{**} γ ₃₀₀ ^{**} γ ₃₀₁ ^{**} γ ₃₀₂ ^{**} γ ₃₀₃ ^{**} γ ₃₀₄ ^{**} γ ₃₀₅ ^{**} γ ₃₀₆ ^{**} γ ₃₀₇ ^{**} γ ₃₀₈ ^{**} γ ₃₀₉ ^{**} γ ₃₁₀ ^{**} γ ₃₁₁ ^{**} γ ₃₁₂ ^{**} γ ₃₁₃ ^{**} γ ₃₁₄ ^{**} γ ₃₁₅ ^{**} γ ₃₁₆ ^{**} γ ₃₁₇ ^{**} γ ₃₁₈ ^{**} γ ₃₁₉ ^{**} γ ₃₂₀ ^{**} γ ₃₂₁ ^{**} γ ₃₂₂ ^{**} γ ₃₂₃ ^{**} γ ₃₂₄ ^{**} γ ₃₂₅ ^{**} γ ₃₂₆ ^{**} γ ₃₂₇ ^{**} γ ₃₂₈ ^{**} γ ₃₂₉ ^{**} γ ₃₃₀ ^{**} γ ₃₃₁ ^{**} γ ₃₃₂ ^{**} γ ₃₃₃ ^{**} γ ₃₃₄ ^{**} γ ₃₃₅ ^{**} γ ₃₃₆ ^{**} γ ₃₃₇ ^{**} γ ₃₃₈ ^{**} γ ₃₃₉ ^{**} γ ₃₄₀ ^{**} γ ₃₄₁ ^{**} γ ₃₄₂ ^{**} γ ₃₄₃ ^{**} γ ₃₄₄ ^{**} γ ₃₄₅ ^{**} γ ₃₄₆ ^{**} γ ₃₄₇ ^{**} γ ₃₄₈ ^{**} γ ₃₄₉ ^{**} γ ₃₅₀ ^{**} γ ₃₅₁ ^{**} γ ₃₅₂ ^{**} γ ₃₅₃ ^{**} γ ₃₅₄ ^{**} γ ₃₅₅ ^{**} γ ₃₅₆ ^{**} γ ₃₅₇ ^{**} γ ₃₅₈ ^{**} γ ₃₅₉ ^{**} γ ₃₆₀ ^{**} γ ₃₆₁ ^{**} γ ₃₆₂ ^{**} γ ₃₆₃ ^{**} γ ₃₆₄ ^{**} γ ₃₆₅ ^{**} γ ₃₆₆ ^{**} γ ₃₆₇ ^{**} γ ₃₆₈ ^{**} γ ₃₆₉ ^{**} γ ₃₇₀ ^{**} γ ₃₇₁ ^{**} γ ₃₇₂ ^{**} γ ₃₇₃ ^{**} γ ₃₇₄ ^{**} γ ₃₇₅ ^{**} γ ₃₇₆ ^{**} γ ₃₇₇ ^{**} γ ₃₇₈ ^{**} γ ₃₇₉ ^{**} γ ₃₈₀ ^{**} γ ₃₈₁ ^{**} γ ₃₈₂ ^{**} γ ₃₈₃ ^{**} γ ₃₈₄ ^{**} γ ₃₈₅ ^{**} γ ₃₈₆ ^{**} γ ₃₈₇ ^{**} γ ₃₈₈ ^{**} γ ₃₈₉ ^{**} γ ₃₉₀ ^{**} γ ₃₉₁ ^{**} γ ₃₉₂ ^{**} γ ₃₉₃ ^{**} γ ₃₉₄ ^{**} γ ₃₉₅ ^{**} γ ₃₉₆ ^{**} γ ₃₉₇ ^{**} γ ₃₉₈ ^{**} γ ₃₉₉ ^{**} γ ₄₀₀ ^{**} γ ₄₀₁ ^{**} γ ₄₀₂ ^{**} γ ₄₀₃ ^{**} γ ₄₀₄ ^{**} γ ₄₀₅ ^{**} γ ₄₀₆ ^{**} γ ₄₀₇ ^{**} γ ₄₀₈ ^{**} γ ₄₀₉ ^{**} γ ₄₁₀ ^{**} γ ₄₁₁ ^{**} γ ₄₁₂ ^{**} γ ₄₁₃ ^{**} γ ₄₁₄ ^{**} γ ₄₁₅ ^{**} γ ₄₁₆ ^{**} γ ₄₁₇ ^{**} γ ₄₁₈ ^{**} γ ₄₁₉ ^{**} γ ₄₂₀ ^{**} γ ₄₂₁ ^{**} γ ₄₂₂ ^{**} γ ₄₂₃ ^{**} γ ₄₂₄ ^{**} γ ₄₂₅ ^{**} γ ₄₂₆ ^{**} γ ₄₂₇ ^{**} γ ₄₂₈ ^{**} γ ₄₂₉ ^{**} γ ₄₃₀ ^{**} γ ₄₃₁ ^{**} γ ₄₃₂ ^{**} γ ₄₃₃ ^{**} γ ₄₃₄ ^{**} γ ₄₃₅ ^{**} γ ₄₃₆ ^{**} γ ₄₃₇ ^{**} γ ₄₃₈ ^{**} γ ₄₃₉ ^{**} γ ₄₄₀ ^{**} γ ₄₄₁ ^{**} γ ₄₄₂ ^{**} γ ₄₄₃ ^{**} γ ₄₄₄ ^{**} γ ₄₄₅ ^{**} γ ₄₄₆ ^{**} γ ₄₄₇ ^{**} γ ₄₄₈ ^{**} γ ₄₄₉ ^{**} γ ₄₅₀ ^{**} γ ₄₅₁ ^{**} γ ₄₅₂ ^{**} γ ₄₅₃ ^{**} γ ₄₅₄ ^{**} γ ₄₅₅ ^{**} γ ₄₅₆ ^{**} γ ₄₅₇ ^{**} γ ₄₅₈ ^{**} γ ₄₅₉ ^{**} γ ₄₆₀ ^{**} γ ₄₆₁ ^{**} γ ₄₆₂ ^{**} γ ₄₆₃ ^{**} γ ₄₆₄ ^{**} γ ₄₆₅ ^{**} γ ₄₆₆ ^{**} γ ₄₆₇ ^{**} γ ₄₆₈ ^{**} γ ₄₆₉ ^{**} γ ₄₇₀ ^{**} γ ₄₇₁ ^{**} γ ₄₇₂ ^{**} γ ₄₇₃ ^{**} γ ₄₇₄ ^{**} γ ₄₇₅ ^{**} γ ₄₇₆ ^{**} γ ₄₇₇ ^{**} γ ₄₇₈ ^{**} γ ₄₇₉ ^{**} γ ₄₈₀ ^{**} γ ₄₈₁ ^{**} γ ₄₈₂ ^{**} γ ₄₈₃ ^{**} γ ₄₈₄ ^{**} γ ₄₈₅ ^{**} γ ₄₈₆ ^{**} γ ₄₈₇ ^{**} γ ₄₈₈ ^{**} γ ₄₈₉ ^{**} γ ₄₉₀ ^{**} γ ₄₉₁ ^{**} γ ₄₉₂ ^{**} γ ₄₉₃ ^{**} γ ₄₉₄ ^{**} γ ₄₉₅ ^{**} γ ₄₉₆ ^{**} γ ₄₉₇ ^{**} γ ₄₉₈ ^{**} γ ₄₉₉ ^{**} γ ₅₀₀ ^{**} γ ₅₀₁ ^{**} γ ₅₀₂ ^{**} γ ₅₀₃ ^{**} γ ₅₀₄ ^{**} γ ₅₀₅ ^{**} γ ₅₀₆ ^{**} γ ₅₀₇ ^{**} γ ₅₀₈ ^{**} γ ₅₀₉ ^{**} γ ₅₁₀ ^{**} γ ₅₁₁ ^{**} γ ₅₁

6 Méta-analyses

Aux États-Unis, les études randomisées portant sur de petits programmes de formation professionnelle sont trop nombreuses pour justifier la rédaction – ou même la lecture – d'analyses critiques pour chacune d'elles. Cette section se tourne donc vers la méta-analyse, discipline qui consiste à synthétiser les estimations issues de nombreuses études. Nous passons d'abord en revue quatre méta-analyses antérieures. Nous présentons ensuite notre propre méta-analyse d'études randomisées menées aux États-Unis.

Les cinq méta-analyses examinées ici couvrent des pans distincts mais partiellement superposés de la littérature : Smedslund et al. (2006) examinent les évaluations randomisées des programmes d'aide à l'emploi, presque toutes provenant des États-Unis ; Haelermans et Borghans (2012) les études non randomisées sur la formation en cours d'emploi dans le monde entier ; Card, Kluve et Weber (2018) une vaste collection internationale d'études randomisées et non randomisées sur les interventions sur le marché du travail ; Peck et al. (2021) les études sur les parcours professionnels aux États-Unis, randomisées et non randomisées ; et notre méta-analyse exhaustive des études randomisées aux États-Unis portant sur la formation.

Les études s'accordent sur la question de premier ordre : les impacts moyens sur les revenus ou l'emploi sont positifs, bien que non transformateurs. En revanche, les résultats divergent quant à la question de second ordre portant sur les caractéristiques d'une intervention associées à un impact plus important. Cela s'explique notamment par le fait que les différentes études examinent des caractéristiques différentes. Une autre explication possible réside dans la fragilité de ces résultats de méta-régression, compte tenu de la petite taille des échantillons et des intercorrélations entre les caractéristiques. Nous présentons quelques notions conceptuelles préliminaires dans l'annexe E.

6.1 Smedslund et al. (2006), « Programmes de travail pour les bénéficiaires de l'aide sociale »

L'étude de Smedslund et al. (2006) est issue de la Collaboration Campbell, qui encourage les revues systématiques des données probantes sur les interventions sociales. À l'instar de son homologue Cochrane, axée sur la santé, la Collaboration Campbell privilégie les essais randomisés. Le domaine d'intérêt de l'étude de Smedslund et al. concerne les programmes d'insertion professionnelle des bénéficiaires de l'aide sociale à l'échelle mondiale. Ce domaine recoupe celui de notre revue, à la manière d'un diagramme de Venn : certains programmes d'insertion professionnelle ne proposent pas de formation ; certains programmes de formation ne sont pas expressément destinés aux personnes bénéficiant de l'aide sociale.

En raison de la tradition américaine en matière d'expériences sociales, Smedslund et al. sont naturellement attirés par les États-Unis. 44 des 46 études se déroulent dans ce pays, et les deux autres au Canada.

Smedslund et al. constatent qu'en moyenne, les programmes d'insertion professionnelle produisent des bénéfices réels qui s'estompent progressivement. Ces programmes ont augmenté l'emploi de 9,7 % (et non de 9,7 points de pourcentage) la première année, de 9,2 % la deuxième année et de 3,7 % la cinquième année. Ces chiffres correspondent à des moyennes à effets aléatoires et sont tous statistiquement très significatifs. Les chiffres correspondants concernant l'impact sur les revenus sont de 4,3 %, 4,4 % et 1,1 %, ce dernier étant difficilement distinguable de 0 (Smedslund et al., 2006, p. 2).

Smedslund et al. constatent une certaine asymétrie dans les graphiques en entonnoir relatifs aux impacts sur l'emploi et les revenus, mais ne quantifient pas le biais de publication qui en résulte. Comme expliqué au paragraphe 3.4, lorsque la plupart des études sous-jacentes sont des essais randomisés financés par des fonds publics, on s'attend à un biais de publication moindre.

L'étude estime l'impact moyen des seuls programmes comportant un volet de formation. Elle ne ventile pas cet impact en fonction du temps écoulé depuis la randomisation. Globalement, les programmes incluant une formation professionnelle ont augmenté l'emploi de 5,6 % et les revenus de 4,5 % (Smedslund et al., 2006, p. 2).

Ces estimations sont elles aussi statistiquement différentes de zéro. Elles sont statistiquement indiscernables des moyennes des programmes n'ayant pas intégré de formation aux compétences. Cependant, cette affirmation repose sur une comparaison intergroupes (telle que définie à l'annexe E) et doit donc être interprétée avec prudence.

6.2 Haelermans et Borghans (2012), « Effets salariaux de la formation en cours d'emploi : une méta-analyse »

Cette méta-analyse rassemble 71 estimations d'impact issues de 38 études sur la formation en milieu professionnel. 37 de ces estimations proviennent d'Europe et 13 des États-Unis. Presque aucune ne répond aux critères de preuve utilisés dans notre analyse. En particulier, aucune n'est randomisée. Une étude quasi expérimentale potentiellement convaincante est incluse : l'étude de Leuven et Oosterbeek (2004) sur le crédit d'impôt néerlandais pour la formation des plus de 40 ans, présentée en annexe D.

Sur les 71 périodes étudiées et sur les différentes périodes de suivi, l'impact moyen des effets aléatoires sur les revenus salariaux est de 3,9 %, avec une grande signification statistique.

Cependant, Haelermans et Borghans mettent en évidence un biais de publication important. Le graphique en entonnoir de leur étude figure dans la figure 2. Chaque point représente une estimation de l'impact de la formation en situation de travail, l'erreur standard (semblable à la marge d'erreur d'un sondage) étant portée en abscisse et l'impact en ordonnée. Le biais de publication est manifeste : les estimations précises se regroupent à gauche, tandis que les estimations imprécises sont dispersées à droite et présentent une asymétrie ascendante. Autrement dit, les études présentant des résultats bruités et des impacts importants ont eu beaucoup plus de chances d'être publiées que celles montrant des impacts négatifs ou faibles.

Haelermans et Borghans (2012, Tableau 5, ligne 1, avant-dernière colonne) estiment que la correction de ce biais de publication réduit l'impact salarial moyen de la formation en cours d'emploi de 3,9 % à 2,6 %.

En revanche, dans cet ensemble d'études universitaires non randomisées, le biais de publication augmente de 50 % l'impact généralement rapporté.

6.3 Card, Kluve et Weber (2018), « Qu'est-ce qui fonctionne ? Une méta-analyse des évaluations récentes des programmes actifs du marché du travail »

Dans le petit nombre de méta-analyses des programmes du marché du travail, Card, Kluve et Weber (2018) (CKW) est le programme phare. Son champ d'application dépasse celui des méta-analyses mentionnées ci-dessus et englobe tous les programmes ciblés visant à faciliter l'accès à l'emploi. Cela inclut l'aide à l'emploi

Formation, aide à la recherche d'emploi, subventions à l'emploi dans le secteur privé et dans le secteur public : cette deuxième édition rassemble 857 estimations d'impact issues de 207 articles. 502 estimations proviennent des pays germaniques et nordiques, et 87 des pays anglo-saxons (États-Unis, Royaume-Uni, Australie, Nouvelle-Zélande et Canada ; CKW, tableau 1). Les programmes de formation sont les mieux représentés, avec 418 estimations. Parmi celles-ci, 54 sont issues d'expérimentations randomisées (CKW, tableau 3).

L'important volume de données confère à CKW une puissance statistique accrue pour aborder des questions cruciales. Quels types d'interventions sur le marché du travail sont les plus efficaces ? Leurs effets s'estompent-ils ou persistent-ils ? Comment la réponse varie-t-elle selon le groupe démographique ciblé ? Les programmes sont-ils plus performants en période de récession ?

CKW constate que si les études menées dans les pays anglo-saxons s'intéressent principalement au niveau de revenu, dans le reste du monde – et donc globalement dans l'échantillon – c'est l'emploi qui prime. Cette différence d'approche pourrait refléter une différence dans la structure des marchés du travail (Krugman, 1994), le salaire minimum, la protection des travailleurs et le chômage ayant historiquement été plus élevés dans l'UE qu'aux États-Unis. La prédominance de l'emploi comme indicateur conduit CKW à se concentrer exclusivement sur ce critère.

Pour analyser les différentes estimations d'impact, CKW commence par les rendre comparables. Contrairement à la quasi-totalité des recherches recensées dans cette revue, qui estiment les différences simples entre les groupes traitement et contrôle à différentes dates de suivi (analyse en intention de traiter), certaines études, notamment non randomisées et menées en Europe, expriment leurs résultats autrement. Une unité alternative typique est la probabilité par unité de temps de trouver un emploi si l'on n'en a pas déjà un, ce qui correspond à un « taux de risque ». Par exemple, un programme de formation professionnelle en Allemagne pourrait augmenter les chances d'un chômeur de trouver un emploi de 10 % à 20 % par mois. Afin de rendre tous les résultats comparables, CKW les simplifie en un indicateur trichotomique : l'estimation d'impact est négative et significative, non significative, ou positive et significative (à $p < 0,05$). Cette quantification présentant un inconvénient majeur – elle permet de regrouper davantage d'études pour la méta-analyse, mais avec des informations moins précises pour chacune d'elles –, CKW réalise également la méta-analyse sans cette quantification, en se limitant aux études 37 et 38. Les deux approches aboutissent à des résultats globalement similaires qui rendent compte résultats issus de l'échantillon le plus restreint, présentant d'impacts simples sur l'emploi. Nous nous concentrerons sur les des impacts réels sur l'emploi, car ils sont plus faciles à interpréter.

Une autre subtilité technique concerne la manière dont les résultats sont moyennés entre les études. Il ne s'agit ni d'un modèle à effets communs ni d'un modèle à effets aléatoires. CKW, qui s'inscrit dans la tradition économique plutôt que médicale, opte pour une alternative mathématiquement plus simple : elle attribue le même poids à toutes les études.

indépendamment de la ³⁹ On peut considérer cela comme une application de la méta-analyse à effets aléatoires à un précision. À l'extrême, les impacts varient tellement d'un contexte d'étude à l'autre que les petites études sont aussi instructives que les grandes. Et puisque les programmes du marché du travail varient clairement beaucoup dans leur

37La variable trichotomique est modélisée comme un probit ordonné.

38Cette quantification permet également de contourner un autre problème fréquent dans les études d'impact non académiques : la significativité statistique n'est pas rapportée, sauf par des indications de dépassement de seuils tels que $p < 0,05$. Smedslund et al. (2006) traitent ce problème en imputant $p = 0,005, 0,03, 0,075$ et $0,1$ pour trois, deux, une ou aucune étoile. Nous procéderons de manière similaire au § 6.7.

39Faisant référence à Solon, Haider et Wooldridge (2015), CKW accompagne cette OLS non pondérée d'erreurs standard robustes à l'hétéroscédasticité, en effet regroupées par étude.

Malgré les impacts, la pondération égale de CKW ne devrait pas s'écarter significativement des équilibres obtenus par les méthodes classiques à effets aléatoires. Toutefois, une pondération aussi importante des petites études risque de rendre les résultats plus variables.

CKW constate que :

- La formation professionnelle augmente l'emploi de 1,4 à 2 points de pourcentage la première année suivant le début de la formation, et de 4,0 à 4,6 points de plus la deuxième année et les suivantes (CKW, tableau 5, colonnes 3 et 4).⁴⁰ Cela se comprend aisément, car la formation peut temporairement retirer des personnes du marché du travail.
- L'impact moyen de la formation professionnelle sur les chômeurs de longue durée est supérieur de 12 points à celui des bénéficiaires « réguliers » de l'assurance chômage (CKW, Tableau 9, col. 2).
- L'impact moyen est de 6 points plus élevé pour les femmes (CKW, Tableau 9, col. 2).
- L'impact moyen est supérieur de 3 points pour les moins de 25 ans par rapport aux plus âgés (CKW, Tableau 9, col. 2).

Au Danemark, en France, en Allemagne et aux États-Unis (pays disposant de suffisamment de données pour effectuer des vérifications), les programmes ciblés (et pas seulement les programmes de formation) ont globalement davantage aidé la population en période de récession. Ils ont permis d'accroître l'emploi d'environ 3 points supplémentaires pour chaque point de hausse du taux de chômage national ou pour chaque point de baisse de la croissance économique.

Cette méta-analyse constate peu de biais de publication. De même, les résultats publiés dans les revues académiques ne présentent que peu de biais de publication. Les publications diffèrent en moyenne de celles qui ne diffèrent pas.

41

- Les estimations expérimentales sont en moyenne inférieures d'environ 1 point aux estimations non expérimentales, mais ce résultat n'a pratiquement aucune signification statistique.

Interprétés comme des énoncés de causalité, la plupart des résultats sont globalement plausibles. Bien entendu, ces résultats sont entachés de plusieurs niveaux d'incertitude. La plupart des études ne sont pas randomisées. (De fait, des dizaines d'études randomisées ne sont pas incluses, comme nous le verrons dans la section suivante.) La méthode méta-analytique accorde une importance particulièrement grande aux études de petite taille. De nombreuses hypothèses sont testées, sans que l'on tienne compte du fait qu'un certain nombre de résultats seront statistiquement significatifs par hasard. Et, comme le souligne CKW, lorsqu'on compare les impacts entre groupes, corrélation n'implique pas causalité. Nous serions surpris que le fait de cibler les femmes plutôt que les hommes ait réellement pour conséquence d'augmenter l'emploi de 6 points supplémentaires.

Nous trouvons les résultats concernant l'évolution temporelle des impacts – faibles la première année, puis en hausse – particulièrement plausibles. Nous sommes moins confiants quant aux résultats des sous-groupes, interprétés de manière causale.

⁴⁰L'impact à court terme est le terme constant dans l'estimation de régression, qui n'est pas rapporté dans CKW et est plutôt extrait des données et du code publics.

⁴¹Les graphiques en entonnoir CKW (figure 3) sont visuellement et techniquement compliqués, ils ne sont donc pas reproduits ici.

6.4 Peck et al. (2021), « Une méta-analyse de 46 évaluations d'impact des parcours professionnels »

Ces dernières décennies, deux grands paradigmes ont émergé au sein de la communauté américaine de la formation professionnelle subventionnée : les stratégies sectorielles et les parcours de carrière. Les stratégies sectorielles insistent sur la collaboration avec les employeurs pour concevoir les programmes de formation ; la section 7 leur accordera une attention particulière.

Les programmes d'accompagnement professionnel visent à soutenir les participants à travers des étapes de difficulté croissante, telles que la scolarité, la formation en milieu de travail et l'expérience professionnelle (Fein 2012 ; Sarna et Strawn 2018, annexe 2.1). Dans les années 2010, certains services du gouvernement américain ont investi dans l'élargissement et l'évaluation des données probantes relatives aux programmes d'accompagnement professionnel. L'une des conclusions de ces travaux est la méta-analyse de Peck et al. (2021).

La méta-analyse s'articule autour de deux questions : Quels sont les effets moyens des programmes ?

Quels aspects de la conception et de la mise en œuvre des programmes sont associés à la réussite ? L'étude rassemble les résultats de 46 études sur les parcours professionnels, dont 27 randomisées et 19 non randomisées.

La réponse à la première question est « positive ». Peck et al. (2021, tableau 2-1) estiment que les programmes d'orientation professionnelle ont augmenté l'emploi dans les professions ciblées de manière impressionnante, de 72 %, passant de 26 à 45 points de pourcentage. De ce point de vue, les programmes ont atteint leurs objectifs.

Plus généralement, les programmes ont augmenté l'emploi total d'environ 6 points, partant d'un niveau initial d'environ 60 %. Cependant, ces statistiques correspondent à des moyennes calculées sur les périodes utilisées par les études sous-jacentes pour évaluer les impacts (Peck et al., 2021, p. 123). Par conséquent, il est impossible de déterminer si les impacts ont été initialement élevés puis ont diminué, comme dans l'évaluation du Job Corps, ou l'inverse, comme l'a constaté CKW. Les impacts sur les revenus sont ventilés par période ; ils s'élèvent à 1 040 \$ par an au cours des trois premières années suivant la mise en œuvre du programme, sur un revenu de base de 16 320 \$. Les impacts à plus long terme semblent être d'une ampleur similaire, mais le rapport ne les exprime jamais en dollars, car ils n'atteignent pas le seuil de signification statistique retenu (Peck et al., 2021, annexe D-5).

42

Pour répondre à la deuxième question de recherche, relative aux corrélats d'impact, l'étude dresse une liste de plus de 75 caractéristiques de programmes, allant de la durée de la formation (supérieure à un an) à l'obtention d'un diplôme universitaire. Le codage de ces caractéristiques – l'attribution de 0 et de 1 à environ 75 caractéristiques et 46 études – constitue une contribution originale. Il alimentera notre analyse des programmes sectoriels présentée au paragraphe 7.

Comme le nombre de caractéristiques est supérieur au nombre d'études, il est impossible d'identifier l'impact de chacune. En effet, on considère généralement qu'il faut 10 études pour chaque corrélat potentiel, ce qui, dans le cas présent, ne permettrait d'étudier qu'environ 4,6 caractéristiques. Afin de restreindre le champ d'étude, Peck et al. regroupent les caractéristiques en blocs thématiques, tels que les données démographiques et les critères d'admissibilité. Les résultats, notamment les impacts sur l'emploi et les revenus, sont ensuite analysés par régression sur chaque bloc individuellement. Seules les caractéristiques « survivantes » sont conservées : celles qui sont statistiquement significatives ($p < 0,1$) et dont le coefficient standardisé est supérieur à 0,1.

L'analyse principale étant bayésienne, l'incertitude est exprimée par des intervalles de crédibilité. Peck et al. (2021, annexe D-6) estiment la probabilité d'un effet à long terme (au-delà de trois ans) à 89 %, puis fixent un seuil à 90 %. La méta-analyse porte sur les tailles d'effet standardisées. Ce résultat étant écarté car non significatif, il n'est jamais converti en valeur monétaire.

ampleur. ⁴³ Enfin, les résultats sont régressés uniquement sur l'ensemble des survivants. Le processus est le suivant : réalisée séparément pour chaque résultat et chaque période.

Peck et al. présentent clairement les résultats relatifs aux corrélats de réussite dans un diagramme, dont un extrait figure ici (figure 11). Quelques caractéristiques apparaissent dans la colonne de droite, présentant la plus forte association avec un impact positif sur l'emploi ou les revenus. Les programmes non dispensés par un collège communautaire ont obtenu de meilleurs résultats, de même que ceux qui fixaient des exigences de revenu minimum, ceux qui proposaient des formations de perfectionnement pour aider les travailleurs à conserver leur emploi dans la même entreprise, ou ceux qui impliquaient les employeurs dans la conception des programmes. À l'instar de la méta-analyse de CKW, il est difficile d'évaluer la pertinence de ces corrélations non expérimentales, même si elles semblent globalement plausibles.

⁴³Les coefficients standardisés sont ceux qui apparaissent après avoir standardisé toutes les variables, c'est-à-dire en les divisant par leurs écarts-types d'échantillon.

CHARACTERISTICS	NOT ASSOCIATED	SMALLER IMPACTS	LARGER IMPACTS
ADMINISTRATIVE ARRANGEMENTS	- School or school district (partner) - Labor union (partner) - Any other agency type	- Community or technical college (lead or partner) - Government agency (partner) - University (partner) - Other (partner)	- Workforce agency (lead) - Community organization (lead) - Trade association (partner)
ELIGIBILITY	Any other eligibility criteria	—	Program has income requirements
INSTRUCTIONAL OFFERINGS	- Offers basic skills training - Offers basic/secondary education, English language acquisition, developmental or remedial education - Offers alternative times/places	Offers flexible sequencing, hybrid instruction, online courses	—
PATHWAYS AND TRAINING	—	—	Offers training above entry-level
CREDENTIAL TYPES	- State or local licensure - Certification developed by an employer or industry association	- Occupational certificate or technical diploma - College degree (associate's, bachelor's)	—
TRAINING DURATION	All training durations	—	—
TRAINING INDUSTRY	- Healthcare - Manufacturing or construction - Information technology - Education	—	—
EMPLOYER ENGAGEMENT ACTIVITIES	Formal employer partnership	Program convenes an employer advisory council	Offers incumbent worker training
EMPLOYER ROLE	- Offers work-based learning (paid or unpaid) - Employer provides resources - Employer delivers instruction - Employer provides labor market information	Employers deliver career awareness services	Provide input on curriculum or program development
ONE-ON-ONE SUPPORT	- Case management - Career or college navigation - Staff assistance mandatory/required	Academic advising	—
SUPPORT SERVICES	- Emergency assistance - Food assistance - Internet - Tutoring - Child/dependent care assistance - Transportation assistance - Connection with benefits and social services - Job search and placement	Tuition, training cost, other financial assistance	—
PARTICIPANT COMPOSITION	- Gender - Education levels - Race/ethnicity	—	—
LOCAL CONTEXT	Unemployment rate	—	—

Figure 11. Corrélat des impacts positifs sur l'emploi et les revenus, reproduits à partir de Peck et al. (2021)

Remarques : Reproduit de Peck et al. (2021), pièce 3-2.

6.5 Une nouvelle méta-analyse d'expériences randomisées de formation professionnelle aux États-Unis

Afin d'obtenir une vision systématique des nombreuses données issues d'essais randomisés de formation professionnelle menés aux États-Unis, nous effectuons une recherche exhaustive d'études et procédons à une méta-analyse. Nous utilisons le logiciel Claude Opus, sous la supervision attentive d'un expert, pour extraire les informations nécessaires à chaque expérience parmi les milliers de pages de rapports.

Pour constituer l'échantillon, nous recherchons des essais randomisés américains d'interventions incluant une formation, dans les revues suivantes : Greenberg et Shroder (2004) ; Smedslund et al. (2006) ; Dutta-Gupta et al. (2016) ; Card, Kluve et Weber (2018) ; Peck et al. (2021) ; et Katz et al. (2022). Nous avons ajouté des suivis récents sur les évaluations PACE et WorkAdvance.

L'avertissement mesuré de Card, Kluve et Weber (2010) – « Des erreurs de mesure et d'interprétation sont probables lors de l'extraction des données des études. » – s'applique également à cette initiative accélérée par l'IA. La collecte et le regroupement des études impliquent de nombreux choix, certains effectués par des humains et d'autres par l'IA. Une intervention comprend-elle une formation suffisamment poussée pour être incluse dans l'échantillon ? La formation est-elle un élément central de l'intervention ? Si un impact est observé entre le 7^e et le 18^e mois suivant la randomisation, s'agit-il d'un impact à court terme (année 1) ou à moyen terme (année 2) ?

Pour améliorer l'extraction, nous demandons à Claude : a) de produire un tableau de justification documentant les sources et la logique des valeurs extraites ; b) de procéder à une vérification critique de toutes les extractions, en consultant le tableau de justification et en se référant aux sources. Comme pour les analyses effectuées sans IA, à mesure que nous apprenons à maîtriser les données et que nous les analysons, nous vérifions et corrigeons l'extraction.

Une autre limite, fréquente dans ce type de littérature, est que de nombreuses études d'impact non académiques ne présentent pas de mesures continues de la significativité statistique, telles que les erreurs-types. Elles indiquent seulement si la significativité dépasse certains seuils standard. Lorsque l'indicateur est binaire (adhésion au traitement, emploi) et qu'aucune variable de contrôle n'est incluse, cela ne pose pas de problème car une formule standard relie les estimations ponctuelles aux erreurs-types. Dans le cas contraire, nous utilisons néanmoins cette formule pour les variables binaires, tandis que pour la variable continue des revenus, nous suivons Greenberg, Michalopoulos et Robins (2006, note 9) en imputant les valeurs $p = 0,005$, $0,03$, $0,075$ et $0,55$ pour $p < 0,01$, $p < 0,05$, $p < 0,1$ et $p > 0,1$, respectivement. Ces imputations incertaines influencent les résultats de la méta-analyse à effets aléatoires.

6.5.1 L'échantillon

Notre échantillon comprend 56 études. Parmi celles-ci, 24 présentent des résultats pour différentes populations, organisations de mise en œuvre ou localités. Cela porte le nombre d'estimations d'impact à 146.

L'échantillon couvre son créneau de manière plus exhaustive que les méta-analyses précédentes. Trente-sept des 56 études figurent dans Smedslund et al. (2006), aucune dans Haelermans et Borghans (2012), deux dans Card, Kluve et Weber (2018) et 26 dans Peck et al. (2021). L'échantillon comprend les quatre principales expériences menées aux États-Unis et analysées au paragraphe 4, ainsi que de nombreuses études portant sur des programmes de moindre envergure.

Étant donné la diversité et la complexité des programmes de formation professionnelle, la formation occupe une place plus importante dans certains que dans d'autres, et il est difficile de tracer une ligne de démarcation entre les interventions axées sur la formation et celles qui ne le sont pas. Nous avons privilégié l'inclusion, tout en demandant à Claude d'évaluer si la formation professionnelle constituait un élément central de chaque programme. Environ la moitié des expériences ont satisfait à ce critère : 33, avec 78 estimations d'impact. Nous avons réalisé une méta-analyse de l'ensemble des études ainsi que du sous-ensemble où la formation était l'élément principal.

	Expériences intégrant la formation			Ceux qui ont une formation primaire		
	Signifier	SD	N	Signifier	SD	N
Paramètre						
Année de randomisation	1998	12	146	1998	13	78
Sud (%)	18	38	146	18	39	78
Nord-Est (%)	19	40	146	19	40	78
Midwest (%)	22	42	146	17	38	78
Ouest (%)	23	42	146	17	38	78
Urbain (%)	90	30	146	97	16	78
Rural (%)	8	26	146	3	16	78
Chômage à la randomisation (%)	6,4	1,3	146	6,5	1,3	78
Population cible						
Adultes bénéficiant de l'aide sociale ou à faible revenu (%)	62	49	146	54	50	78
Parents non gardiens (%)	3	18	146	0	0	78
Abus de substances (%)		12	146	3	16	78
Personnes impliquées dans le système judiciaire (%)	1	24	146		11	78
Travailleurs déplacés (%)	6	16	146	1	19	78
Jeunesse (%)	3	42	146	4	48	78
Invalidité (%)	23 1	12	146	36 3	16	78
Données démographiques						
Mâle (%)	37	34	135	40	31	78
Âge moyen	29	7	113	28	8	59
Noir (%)	47	27	132	47	29	75
Hispaniques (%)	26	25	117	28	26	70
Aucun diplôme d'études secondaires (%)	44	30	117	46	38	65
Traitement						
Formation en salle de classe (%)	77	42	146	90	31	78
Formation en cours d'emploi (%)	39	49	146	29	46	78
Aide à la recherche d'emploi (%)	83	38	146	85	36	78
Obligatoire (%)	28	45	146	12	32	78
Durée du traitement (mois)	5,4	2,7	86	5,8	2,9	61
Coût par personne traitée (2025 \$)	11 402 7	8 629	98	13 598 10	9 359	52
Échelle (%)		25	146		31	78
Engagement des employeurs						
Programme sectoriel (%)	12	32	146	22	42	78
Programmes d'études adaptés aux employeurs (%)	18	38	145	33	47	78
Développement/placement professionnel (%)	31	46	145	40	49	78
Contribution de l'employeur au programme (%)	19	39	145	35	48	78
L'employeur s'engage à embaucher (%)	3	18	145	5	22	78
Administration						
Financement : public (%)	77	42	146	64	48	78
Financement : public et privé (%)	19	40	146	31	46	78
Financement : privé (%)	3	18	146	5	22	78
Admin : public (%)	45	50	146	40	49	78
Administration : public et privé (%)	18	38	146	15	36	78
Admin : privé (%)	37	48	146	45	50	78
Caractéristiques de l'étude						
Taille de l'échantillon	2 948	6 738	146	2 083	4 040	78
Données de résultats issues des registres officiels (%)	70	46	146	59	50	78
Taux de réponse (%)	97	11	113	96	12	53
Publication académique (%)	7	25	146	3	16	78

Tableau 1. Caractéristiques des études menées dans le cadre d'expériences de formation professionnelle aux États-Unis

Remarques : L'échantillon complet comprend 146 estimations issues de 56 études. Les moyennes ne sont pas pondérées. La taille des échantillons varie en raison de Données manquantes. Pour les expériences multirégionales, le chômage « régional » correspond au chômage national. Les coûts indiqués sont par traitement membre du groupe.

La plupart des caractéristiques des études que nous avons recueillies sont répertoriées dans le tableau 1, accompagnées de leurs moyennes non pondérées et de leurs écarts types. Les études se déroulent majoritairement en milieu urbain et sont réparties de façon assez homogène sur l'ensemble des régions de recensement. La plupart visent à aider les personnes à faibles revenus, seulement 3 à 4 % étant destinées aux travailleurs déplacés. Les femmes et les personnes noires sont surreprésentées dans les échantillons. Lorsque la formation est l'élément principal, elle se déroule principalement en salle de classe (90 %), contre seulement 29 % pour la formation en situation de travail. Les participants passent environ six mois dans ces programmes, pour un coût de 13 598 \$ par personne prise en charge (en dollars de 2025, et non par personne formée). Les programmes se présentant comme sectoriels sont minoritaires, de même que, plus généralement, les initiatives visant à impliquer les employeurs. Presque toutes les interventions étudiées ont bénéficié d'un financement public, mais un peu moins de la moitié des interventions axées principalement sur la formation ont délégué leur mise en œuvre locale à des organismes privés. Peu de ces expériences ont fait l'objet de publications dans des revues scientifiques.

La figure 1 présente un aperçu graphique des estimations d'impact issues de l'ensemble de données. Elle illustre l'ensemble des impacts annuels (ou annualisés) sur les revenus en fonction du temps de suivi. Les données diffèrent de celles de la méta-analyse car elles sont mises à jour selon la même périodicité que les sources primaires (mensuelle, trimestrielle, annuelle, etc.). Les trois programmes fédéraux (JTPA, Job Corps et WIA) sont représentés en bleu. Certains programmes sectoriels particulièrement performants sont indiqués en nuances de magenta.

6.5.2 Impacts moyens

	Expériences intégrant la formation			Ceux qui ont une formation primaire		
	Contrôle signifier	Impact	N	Contrôle signifier	Impact N	
Taux d'adoption du programme (%)	10,7 (2,8)	58,3 (5,8)	140	8,6 (3,5)	66,1 (5,7)	78
Taux de participation à la formation (%)	40,3 (2,9)	19,2 (4,4)	95	49,1 (2,8)	28,1 (4,9)	58
Emploi, court terme (%)	56,0 (3,0)	4,5 (1,4)	93	59,5 (4,3)	4,8 (1,7)	41
Emploi, moyen terme (%)	56,9 (2,8)	2,6 (0,4)	87	62,6 (4,0)	2,8 (0,6)	40
Emploi, à long terme (%)	59,8 (4,4)	1,8 (0,5)	64	63,4 (5,1)	1,7 (0,5)	36
Revenus à court terme (en dollars de 2025)	10 575 (1 194)	640 (244)	119	13 834 (2 038)	835 (553)	53
Revenus à moyen terme (en dollars de 2025)	12 577 (1 020)	984 (255)	121	14 645 (1 473)	1 139 (446)	67
Revenus à long terme (en dollars de 2025)	14 361 (1 802)	657 (185)	81	16 148 (2 443)	791 (302)	53

Tableau 2. Moyennes des groupes témoins et estimations d'impact dans les expériences de formation professionnelle menées aux États-Unis. Remarques : Les impacts sont estimés par méta-analyse du maximum de vraisemblance restreint (REML). Les moyennes des groupes témoins sont calculées à l'aide des pondérations REML correspondantes. Les nombres entre parenthèses représentent les écarts-types des valeurs des groupes témoins et les erreurs-types des estimations d'impact. La taille des échantillons varie en raison des données manquantes. Les résultats à court terme sont suivis pendant environ un an, ceux à moyen terme pendant environ deux ans et ceux à long terme pendant trois à cinq ans.

Les moyennes d'impact à effets aléatoires figurent dans le tableau 2. Les résultats sont au nombre de quatre : le taux de participation au programme testé ; le taux de participation à toute formation ; l'emploi ; et les revenus. Les impacts sur les deux derniers sont ventilés selon l'horizon temporel. Conformément à la méthode CKW, les horizons temporels sont le court terme (année 1 après randomisation), le moyen terme (année 2) et le long terme. Nous avons retenu le long terme pour les années 3 à 5, sous réserve de la disponibilité des données, car la plupart des études interrompent le suivi avant la 6e année, et beaucoup bien avant.

⁴⁴ Étant donné que certaines études produisent des estimations pour plusieurs populations, qui peuvent ne pas être Les erreurs standard étant totalement indépendantes, nous les regroupons par étude. Nous utilisons l'estimateur à effets aléatoires dominant, le maximum de vraisemblance restreint (REML). Les résultats visent donc à représenter l'impact d'une étude typique, et non l'impact moyen sur les participants dans l'ensemble des études.

Si le terme « travail » désigne une augmentation moyenne et statistiquement significative de l'emploi et des revenus, alors la formation professionnelle a porté ses fruits aux États-Unis. Dans le tableau 2, 11 des 12 estimations d'impact moyen sont 2 à 4 fois supérieures à leur erreur standard, ce qui les rend hautement significatives. (Il s'agit d'analyses en intention de traiter, exprimées par personne ayant bénéficié du traitement.) L'exception concerne l'impact à court terme des programmes de formation sur les revenus, qui peut refléter des pertes de revenus temporaires pendant la formation.

Cependant, un programme de formation standard n'a pas d'impact significatif. Partant d'un taux d'emploi d'environ 60 %, il augmente les chances de trouver un emploi de 2,6 à 2,8 points à moyen terme (deuxième année) et de 1,7 à 1,8 point à long terme (ces chiffres concernent l'ensemble des programmes et les programmes axés sur la formation). En moyenne, un programme de formation standard accroît le revenu avant impôt de 7,8 % à moyen terme, soit de 1 000 à 1 100 \$ par an, à partir d'un revenu de base de 12 500 \$ pour l'ensemble des participants et de 14 700 \$ pour les programmes axés sur la formation. L'impact sur les revenus à long terme est de 4,6 %, soit de 700 à 800 \$, à partir d'un revenu de base de 14 400 \$ et de 16 100 \$ (tous les chiffres sont exprimés en dollars de 2025).

L'une des causes du faible impact sur les résultats du marché du travail – mais aussi une source d'espoir dans certains cas – réside dans le faible impact sur la participation à la formation. Dans les interventions axées principalement sur la formation, l'offre expérimentale d'accès est acceptée dans 75 % des cas, tandis que seulement 9 % du groupe témoin y accède. En divisant cette différence de 66 % par les ITT mentionnées précédemment, on obtient des estimations « programme tardif » supérieures de 50 %. De même, l'impact sur la participation à tout programme de formation n'est que de 28 %, car de nombreux membres du groupe témoin se tournent vers des formations alternatives. Si ces alternatives sont tout aussi efficaces, alors l'impact de la formation en tant que telle – les valeurs de « formation tardive » – est environ quatre fois supérieur (1/0,28) aux ITT figurant dans les dernières lignes du tableau 2. Si, hypothétiquement, un programme de formation moyen était mis en œuvre dans un contexte où les alternatives seraient plus rares, son taux de participation et ses ITT pourraient être jusqu'à quatre fois plus élevés.

⁴⁴Des estimations à plus long terme sont disponibles pour quelques programmes, notamment Year Up, pour 7 ans (Fein et Dastrup 2022), Project QUEST pour 11 ans (Roder et Elliott 2021) et le Job Corps pour 20 ans (Schochet 2021). Toutefois, il peut exister une tendance à suivre plus longtemps les programmes qui réussissent.

6.5.3 Avantages et coûts

De nombreuses études de notre échantillon comparent les avantages et les coûts de la formation. Toutes prennent en compte le coût initial des programmes et leurs répercussions sur les revenus (avant impôt). Elles diffèrent quant aux autres éléments inclus : impact sur la criminalité, transferts fiscaux, valeur des biens et services produits pendant la formation, etc.

Face à cette diversité, pour réaliser une analyse coûts-avantages des résultats méta-analytiques, nous nous basons principalement sur le modèle de la récente expérimentation WorkAdvance menée sur quatre sites (Schaberg et Greenberg, 2020). Ce modèle prend en ⁴⁵ Voir l'annexe G pour plus de détails. L'analyse établit une distinction entre les compte les perspectives du participant, du gouvernement et de la société, cette dernière étant la somme algébrique des deux premières. Il estime les transferts entre le gouvernement et les participants – plus précisément, les membres du groupe de traitement – liés aux impôts et au recours aux aides sociales telles que les bons alimentaires. Les taxes sur les ventes, perçues aux États-Unis par les États et les collectivités locales, sont prises en compte car elles constituent un autre canal de transfert d'une partie des revenus et des dépenses des stagiaires vers les finances publiques. L'analyse ajoute (ou soustrait) les dépenses liées au travail, telles que les uniformes et l'essence, pour obtenir le résultat final. travailler et le coût d'opportunité du temps non rémunéré. ⁴⁶

Nous calculons cinq statistiques descriptives sur les informations présentées dans ce cadre :

- Un ratio avantages-coûts (AC) étroit. Le numérateur est la valeur actuelle nette de la durée de vie.
Les incidences se font sur la rémunération totale (et non uniquement sur le revenu avant impôt). Le dénominateur correspond aux coûts de formation initiaux.
- Un ratio BC plus large, qui déduit du dénominateur la réduction des dépenses sur autres formations.

La valeur marginale des fonds publics (VMFP) correspond au rapport entre le bénéfice net pour le participant et le coût net pour les gouvernements. Dans le cas présent, on suppose que le gouvernement finance la formation.

- Le ratio entre le coût net et le coût brut pour le gouvernement, qui indique le niveau de formation requis
Elle s'autofinance grâce à des recettes fiscales plus élevées et à des dépenses moindres en prestations sociales (en supposant, encore une fois, que le gouvernement prenne en charge les frais).
- Le taux de rendement interne (TRI) du point de vue sociétal, qui est le taux d'actualisation à quel point les coûts et les avantages s'équilibrent.

⁴⁵Le regretté David H. Greenberg a également coécrit un manuel sur l'analyse coûts-avantages.

⁴⁶ Deux études de notre échantillon estiment et monétisent l'impact de la criminalité : MDRC (1980) sur le Programme national de démonstration de travail soutenu (anciens toxicomanes et anciens détenus) et McConnell et Glazerman (2001) sur le programme Job Corps. Nous avons choisi de ne pas le faire car le lien est complexe et encore mal compris. Par exemple, les candidats retenus pour les programmes sectoriels pourraient avoir une propension à la criminalité bien inférieure à celle des décrocheurs scolaires, des anciens détenus et des anciens toxicomanes. Il en va de même pour les travailleurs déplacés par l'IA.

Bien que le ratio avantages-coûts, tel que défini ici, ne tienne pas compte des flux fiscaux et d'autres facteurs, il a l'avantage d'être facile à comprendre, et donc sans doute plus pertinent pour les décideurs politiques. Il est également important pour les financeurs privés, qui ne peuvent pas compenser les économies d'impôt.

Tous les montants en dollars sont exprimés par membre du groupe de traitement, et non par stagiaire. Autrement dit, ils sont calculés selon une approche en intention de traiter (ITT) et non selon une approche en analyse tardive (LATE). Cette distinction n'a cependant aucune incidence sur les cinq statistiques descriptives.

Pour les programmes de formation initiale, nous estimons approximativement que :

- En moyenne, les programmes font un peu mieux que le seuil de rentabilité, avec un ratio avantages-coûts faible. 1,44, une large de 1,80 et un MVPF de 2,52.

En moyenne, un programme financé par l'État couvre finalement environ 76 % de ses coûts. Ce chiffre inclut les économies budgétaires réalisées par les administrations fédérales, étatiques et locales.

- Le TRI sociétal est de 5,97 %.

6.5.4 Corrélats d'impact Pour extrapoler à un

avenir perturbé par l'IA, il est important de savoir comment les impacts diffèrent selon les sous-groupes.

Par exemple, si la formation s'avère plus efficace chez les adultes expérimentés, cela pourrait indiquer (à titre indicatif) son efficacité chez les travailleurs en reconversion professionnelle. Cela dit, il convient de rappeler la mise en garde concernant les comparaisons entre groupes : aussi rigoureuses que soient les études individuelles, les méta-analyses ne permettent pas de déterminer avec certitude quels facteurs rendent la formation plus efficace, plutôt que de simplement y être corrélées.

Pour réduire la liste des corrélats potentiels dans nos méta-régressions, nous avons repris la méthode par blocs de la méta-analyse de Peck et al. (2021) (§ 6.4). De plus, afin de pallier les valeurs manquantes pour certaines variables, nous avons effectué les méta-régressions avec et sans imputation multiple, une méthode permettant de compléter les valeurs manquantes tout en ajustant les erreurs-types en fonction de l'incertitude liée à ces imputations.

47

47 Nous générons 10 ensembles de données imputées, séparément pour chaque résultat et chaque période.

	Emploi				Gains			
	moyen terme		à long terme		moyen terme		à long terme	
Cadre (omis = Sud, urbain)								
Année de randomisation	-0,08**	-0,06	-0,07*	-0,05				
	(0,04)	(0,04)	(0,04)	(0,05)				
Nord-Est	4,36***	4,34***						
	(1,28)	(1,30)						
Midwest	3,61***	3,37***	-0,04	-0,29				
	(0,91)	(0,92)	(0,85)	(0,90)				
Ouest	2,20*	2,13*						
	(1,15)	(1,26)						
Taux de chômage après suivi (%)	1,19***	1,18***			267*	225*		
	(0,36)	(0,38)			(155)	(135)		
Population cible (omise = adultes bénéficiant de l'aide sociale ou à faible revenu)								
abus de substances			7,12***	7,63***			4468***	4468***
			(1,43)	(1,08)			(114)	(114)
impliqué dans le système judiciaire							2701***	2701***
							(114)	(114)
travailleurs déplacés	-4,05***	-3,15***	-2,57**	-2,65**				
	(0,95)	(1,07)	-2,22**	(1,11)				
Jeunesse	-3,00***							
	(0,96)	(0,97)						
Invalidité	6,05**	6,28**	(2,18)	8,18***	8,50***			
	(3,11)			(1,32)	(0,77)			
Traitement (omis = obligatoire)								
Formation en salle de classe		0,08	0,33					
		(0,85)	(1,20)					
Formation en cours d'emploi						690**		
						(333)		
Volontaire				-0,83				
				(0,85)				
Coût par personne traitée (en milliers de dollars)		0,00						
		(0,00)						
Engagement des employeurs								
L'employeur s'engage à embaucher					8830***	8569***	9161***	9161***
					(1813)	(1735)	(135)	(135)
Caractéristiques de l'étude (données omises = source des résultats : documents officiels)								
Source des données sur les résultats : enquête							-315*	-315*
							(176)	(176)
imputation multiple								
R ²	0,36	0,44	0,21	0,24	0,69	0,72	0,87	0,87
Nombre d'expériences	87	87	64	64	120	121	81	81

Tableau 3. Méta-régressions sur les « survivants » des méta-régressions initiales bloc par bloc

Remarques : Estimations REML. Les impacts sur l'emploi sont mesurés en pourcentage et les impacts sur les revenus en dollars de 2025.

Comme expliqué dans le texte, dans chaque régression, les régresseurs inclus sont ceux qui ont survécu aux régressions initiales par blocs. Moyen terme

Les effets se manifestent sur environ 2 ans et à long terme sur plus de 3 ans. Les erreurs standard sont indiquées entre parenthèses et regroupées par étude. Les résultats sont présentés dans les colonnes MI.

Les données () sont regroupées sur 10 imputations ; les autres colonnes utilisent les cas complets. Chaque paire de régressions d'emploi produit

Des résultats différents malgré une taille d'échantillon identique s'expliquent par le fait que les données manquantes pour les variables non survivantes restreignent les échantillons dans certains cas. régressions initiales déterminant les survivants. p < 0,05. p < 0,10, p < 0,01

Les méta-régressions finales — celles portant sur les « survivants » de la sélection par blocs — sont présentées dans le tableau 3. Pour des raisons de puissance statistique, l'échantillon comprend toutes les études, et pas seulement celles de formation primaire. Les résultats concernant les impacts à court terme (année 1) sont omis car ils sont moins significatifs et peuvent être masqués par les effets temporaires de la participation aux programmes. Lors de l'analyse des résultats, il est préférable de se concentrer sur les tendances communes à toutes les colonnes.

On constate que les effets à moyen et long terme sur l'emploi ont diminué d'environ 0,07 point par an depuis les années 1970. Les programmes testés dans le Sud (région de référence du premier graphique) ont été moins efficaces pour l'emploi ; dans les autres régions (Nord-Est, Midwest, Ouest), les effets sur l'emploi étaient supérieurs de 2 à 4 points de pourcentage. La formation en salle n'a pas démontré d'avantage significatif, et il semble plus difficile d'aider les travailleurs ayant perdu leur emploi.

D'autres résultats statistiquement significatifs sont associés à un nombre restreint de programmes. Le seul programme ayant des résultats à long terme ciblant explicitement les personnes ayant des problèmes de toxicomanie ou des antécédents judiciaires — le National Supported Work Demonstration — explique à lui seul les impacts impressionnants observés chez les personnes ayant eu affaire à la justice. Les deux programmes destinés aux personnes ayant une déficience intellectuelle ont augmenté l'emploi d'environ 8 points à long terme. ^{>48,49} L'impact apparemment important de l'engagement d'un employeur à embaucher des diplômés — 8 000 à 9 000 \$ la deuxième année et les suivantes — est dû à trois programmes : le programme Wildcat, qui a inspiré le NSW, Year Up et le Wisconsin Regional Training Partnership. La section 7.1 abordera les deuxième et troisième programmes.</sup>

Il est intéressant de noter que le taux de chômage local est également associé à l'efficacité des programmes de formation. Afin d'étudier le lien avec les conditions du marché du travail, nous avons inclus une covariable indiquant le taux de chômage régional moyen pendant la période de suivi de l'étude. (La « région » correspond à l'une des quatre régions de recensement ou, pour les évaluations multisites, à l'ensemble du pays.) Lorsque le taux de chômage est supérieur d'un point, un programme de formation accroît l'emploi de 1,2 point supplémentaire à moyen terme (deux ans après la randomisation) et les revenus de 250 \$ par an. Toutefois, cette association est beaucoup plus faible à long terme (années 3 à 5). Dans la régression initiale portant sur le bloc de variables caractérisant le contexte de l'étude (non présentée ici), le coefficient relatif à l'emploi à long terme n'est que de 0,37 point et est statistiquement indiscernable de zéro. Il n'apparaît pas dans les régressions à long terme présentées dans le tableau, et les cellules correspondantes sont donc vides.

Plusieurs théories pourraient expliquer la corrélation à moyen terme entre le taux de chômage local et l'impact de la formation. En période de chômage élevé, le participant marginal peut être une personne ayant récemment perdu son emploi, possédant une expérience professionnelle productive et plus à même de tirer profit de la formation et des opportunités qu'elle offre. Il se pourrait également que le chômage lié à l'inadéquation entre l'offre et la demande soit plus élevé lorsque le chômage est plus important, ce qui accroît les gains liés à un changement de secteur d'activité (Şahin et al., 2014). Enfin, une dernière explication est que lors des pics de chômage,

48L'essai de services de transition structurés de formation et d'emploi (Kerachsky et al. 1985) et la démonstration de formation à l'emploi transitoire (Decker et Thornton 1995).

49 Le faible impact observé sur les programmes destinés aux travailleurs déplacés s'explique uniquement par les résultats nuls de l'étude WIA (§4.4). Deux autres programmes de l'échantillon — le programme de démonstration de réinsertion professionnelle de l'assurance chômage du New Jersey (Anderson, Corson et Decker, 1990) et le programme de démonstration d'adaptation des travailleurs du Texas (Bloom, 1990) — ciblaient également ce groupe. Cependant, ces études n'ont fait état que d'impacts à court terme.

La concurrence pour les formations s'intensifie. Cela réduit les alternatives offertes au groupe témoin, par rapport auquel les résultats du groupe expérimental sont comparés. De fait, chaque programme de formation devient plus nécessaire et moins redondant. Comme indiqué dans la section 3.2, cela pourrait amplifier l'effet de l'offre de formation (analyse en intention de traiter) même sans modification de l'impact de la formation sur les participants.

Des données non présentées dans le tableau corroborent cette dernière hypothèse. Dans les régressions initiales par blocs, une hausse d'un point du chômage est associée à une augmentation de 3,4 points de la participation au programme ; l'erreur standard étant de 2,8, ce résultat, bien que suggestif, ne satisfait pas au critère de « survie » de la significativité statistique appliqué ici. Cependant, lorsque l'on considère la participation à tout programme de formation plutôt qu'au programme d'intérêt, l'association à moyen terme avec le taux de chômage devient encore plus forte et statistiquement significative, à 4,7 points (erreur standard de 1,8). En effet, en période de pénurie de main-d'œuvre, proposer un programme de formation supplémentaire augmente le nombre de personnes suivant une formation.

L'interprétation des associations qualitativement différentes entre l'impact de la formation et le chômage à moyen et long terme constitue une énigme majeure. La faible association observée à long terme pourrait simplement s'expliquer par la réduction de la taille de l'échantillon et la perte de puissance statistique. Il se peut aussi que l'impact direct de la formation sur l'emploi se manifeste principalement dans les mois qui suivent la fin du programme, soit à moyen terme. Cet effet peut persister pendant des années, les personnes ayant un emploi le conservant, mais il pourrait être peu dépendant du taux de chômage actuel. Dans l'annexe F, nous présentons les résultats de méta-régressions similaires à celles du tableau 3, avec une différence : l'impact de la formation sur la période précédente est pris en compte. Nous prévoyons que son coefficient sera plus faible dans les méta-régressions portant sur l'impact à moyen terme, car le terme de persistance y concerne le court terme (une année après la randomisation) et inclut des artefacts transitoires liés à la formation. Parfois, la formation retire des personnes du marché du travail. Parfois, elle leur garantit un emploi, temporairement.

Ce terme de « persistance », figurant dans la première ligne du tableau de l'annexe F, est statistiquement significatif sur l'ensemble de la période. Il est plus faible à moyen terme pour l'emploi (comme prévu), mais pas pour les revenus (de façon surprenante). Dans les régressions à long terme sans imputation multiple, le terme à moyen terme Les coefficients d'impact de gain de 0,48 et 0,52 avec le ⁵⁰ Cela suggère que les associations à moyen terme chômage mentionné ci-dessus — 1,2 point d'impact sur l'emploi et 250 \$/an d'impact sur les revenus — persistent dans les années 3 à 5 à la moitié de ces niveaux.

Globalement, les résultats relatifs au taux de chômage indiquent que la formation aide les personnes à retrouver plus rapidement un emploi après une perte. Cependant, l'absence de corrélation à long terme suggère que les conditions du marché du travail local n'ont pas d'incidence sur la capacité de la formation à placer durablement les individus sur la voie d'une meilleure insertion professionnelle.

Les graphiques en forêt montrant les impacts des études individuelles ainsi que ces conclusions générales sont consultables via une interface web à l' [adresse `doodman.github.io/job-training-meta-analysis`](https://doodman.github.io/job-training-meta-analysis). Également publié

⁵⁰ Les régressions par imputation multiple sont ici légèrement moins fiables car elles supposent que les données manquantes dans les estimations d'impact à long terme sont indépendantes de la nouvelle variable explicative, l'impact à moyen terme. En réalité, on observe une certaine tendance à prolonger le suivi des expériences concluantes.

Des tableaux présentent les méta-régressions bloc par bloc à partir desquelles les survivants sont extraits.

6.5.5 Expériences randomisées de formation professionnelle aux États-Unis : résumé La méta-analyse des essais

randomisés menés aux États-Unis appuie les conclusions suivantes :

- Les programmes de formation professionnelle augmentent en moyenne l'emploi et les revenus.
- Les impacts de l'ITT sont faibles, de l'ordre de 2 points d'emploi, et de 1 000 à 1 100 \$.
Les revenus avant impôt la deuxième année et de 700 à 800 \$ par an les années suivantes devraient suffire à compenser les coûts, selon nos estimations.
- Les programmes soutenus par un engagement d'embauche des employeurs augmentent les revenus de
Le nombre d'étudiants admis est nettement supérieur. Ce résultat, bien que basé sur un nombre restreint d'études, est intéressant au vu de l'engouement suscité par les « programmes sectoriels », que nous aborderons prochainement.

7. Cibler les programmes sectoriels

7.1 Histoire

Tandis que le gouvernement américain s'enlisait dans les itérations et les évaluations de ses programmes de main-d'œuvre dotés de milliards de dollars, les entrepreneurs sociaux, les fondations et les administrations locales et étatiques poursuivaient leurs expérimentations à plus petite échelle. Comme mentionné précédemment, un paradigme a émergé : la « stratégie sectorielle », des programmes travaillant en étroite collaboration avec les employeurs, sélectionnant les demandeurs d'emploi, leur formant aux compétences professionnelles et à l'insertion professionnelle, et les plaçant dans des emplois de qualité au sein de secteurs spécifiques à forte demande. Dans une sorte d'évolution convergente, des éléments de cette approche sont apparus dans les années 1980, ici et là à travers le pays (Mangat 2007 ; Conway 2014 ; King et Prince 2015). Finalement, les observateurs ont relevé les similitudes et ont donné un nom à cette approche.

51

Bien que les experts puissent diverger sur certains points, tous s'accordent à dire que les programmes sectoriels jouent un rôle d'intermédiaire actif entre employeurs et employés potentiels. Cette approche contraste avec la stratégie implicite de « formation et espoir » adoptée par la plupart des écoles et des collèges communautaires. Les intermédiaires qui collaborent étroitement avec les principaux employeurs locaux peuvent identifier les compétences recherchées, les intégrer aux programmes d'études, organiser des stages et gagner la confiance des employeurs quant aux compétences de leurs diplômés. King et Prince (2015) expliquent que l'approche sectorielle rompt avec un modèle courant :

...dans lequel de nombreux organismes de formation, avec des degrés variables allant de « presque pas du tout » à « efficacement », ont identifié les compétences recherchées, créé des programmes de formation... puis se sont mis en concurrence... pour que leurs stagiaires soient embauchés. Duplication des efforts, incohérence dans

⁵¹ Voir la note de bas de page 6 pour plus de détails sur la définition.

Les normes de formation et la présence occasionnelle de prestataires de formation peu scrupuleux ont toutes contribué à la méfiance des employeurs... Les établissements d'enseignement et de formation sont peu incités à collaborer avec les employeurs, car leur financement repose sur l'inscription et parfois la réussite aux cours plutôt que sur le placement professionnel.

La plupart des tâtonnements, des erreurs et des enseignements qui ont permis d'élaborer la stratégie sectorielle se sont déroulés sans recours à des essais randomisés. Cependant, l'exemple donné par le Programme national de démonstration de travail soutenu a joué un rôle, car les fondations privées ont cherché à tester rigoureusement les affirmations, et les résultats positifs ont alimenté l'optimisme et les financements. C'est précisément pour ces résultats encourageants que cette section est consacrée aux programmes sectoriels.

Lorsqu'on réfléchit à la manière de tirer parti du succès de certains programmes sectoriels, il est important de reconnaître que leur fonctionnement est intimement lié à leur origine. Intégrer le modèle dans un autre contexte institutionnel, tel qu'un programme fédéral, risque d'en altérer l'essence, selon la manière dont cela est fait. C'est pourquoi nous commençons par retracer l'histoire tout en soulignant le rôle de la recherche randomisée.

7.1.1 Le Centre de formation à l'emploi de San José

Dans les années 1980, après l'arrêt du NSWDC par le gouvernement fédéral, la Fondation Rockefeller a financé quatre études randomisées, baptisées « Projet de démonstration auprès des femmes monoparentales issues de minorités ».

Un site s'est particulièrement distingué : le Centre de formation à l'emploi (CET) de San José. Au cours de la dernière année du suivi de deux ans et demi, le CET a permis d'augmenter le taux d'emploi de 57 % à 66 % et le salaire mensuel de 405 \$ à 506 \$ (en dollars de 1986 ; Burghardt et al., 1992, tableau IV.1).

Parallèlement, grâce à un financement du JTPA, le MDRC a lancé un projet pilote sur 13 sites, baptisé JOBSTART. Ce projet visait à tester des méthodes d'aide aux jeunes décrocheurs scolaires, le groupe le moins aidé par le NSWDC. Une fois de plus, un programme se distingue dans les tableaux de résultats : le CET. Au cours des deux dernières années de suivi, la différence entre le programme CET et le groupe témoin s'élevait en moyenne à 250 \$ par mois. (Cave et al., 1993, tableau 5.12)

Qu'est-ce qui distinguait le CET ? Ce n'était pas tant le fait que les étudiants passaient six mois en classe à acquérir des compétences professionnelles. Ce qui le caractérisait davantage, c'était que le CET partageait des points communs avec ce qui allait devenir les programmes sectoriels, « notamment l'accent mis sur les emplois de qualité et les liens avec les employeurs » (Giloith 2010, p. 20).

Impressionné par ces résultats, le Département du Travail a lancé en 1992 un projet visant à reproduire le modèle du CET dans 14 sites à travers le pays et à l'évaluer par randomisation. L'expérience fut décevante. Seuls quatre sites furent jugés avoir reproduit le modèle avec une grande fidélité – et il s'agissait d'antennes du CET situées hors de San José. Aucun site n'a produit de résultats positifs, significatifs et durables (Miller et al., 2005, tableau 3.2). Les évaluateurs ont provisoirement attribué ce manque d'impact à plusieurs facteurs : les jeunes bénéficiant des répliques étaient légèrement plus favorisés ; la conjoncture économique était plus favorable ; et les membres du groupe témoin disposaient d'autres possibilités de formation (Miller et al., 2005, p. xi).

Ils ont également suggéré que le secret de leur réussite résidait dans les liens tissés au fil des années avec les acteurs locaux :

CET-San Jose est unique à bien des égards : sa croissance organique sur plus de 20 ans, l'engagement exceptionnel de son fondateur et de son équipe, ses liens étroits avec la communauté locale et sa tradition de plaidoyer politique en faveur de la communauté hispanique locale en font un organisme hors du commun. Un modèle comme celui de CET, forgé localement, ne peut sans doute pas être facilement transposé d'une approche verticale à d'autres régions (Miller et al., 2005).

7.1.2 Projet QUEST : Indice tiré d'une recherche non randomisée

Constatant des tendances, les bailleurs de fonds privés ont opté pour le terme « stratégie sectorielle » (Conway 2014). Afin d'en savoir plus sur les programmes relevant de cette catégorie, les fondations Mott, Ford et Casey ont financé deux séries d'évaluations non randomisées, qui se sont contentées de vérifier si les participants gagnaient davantage après avoir suivi le programme qu'avant (Zandniapour et Conway 2002 ; Roder 2008 ; Conway 2014, p. 50).

Il faut s'attendre à ce que la régression vers la moyenne biaise ce type d'études, un phénomène connu en économie du travail sous le nom de « creux d'Ashenfelter » (Ashenfelter, 1978 ; voir annexe A.2). Les personnes dont la situation financière est instable sont plus susceptibles d'intégrer un programme de formation juste après leur dernière période de baisse et, dans de nombreux cas, de se redresser au moins partiellement, même sans formation. Comme pour le CET une décennie plus tôt, un programme figurait dans les deux études et se distinguait particulièrement. Le projet QUEST à San Antonio, au Texas, a enregistré la plus forte hausse de revenus parmi les six programmes analysés par Zandniapour et Conway (2002, figure 2), passant de 5 367 \$ par an avant l'entrée en fonction à 24 907 \$ par an après. Il se classait également premier parmi les neuf programmes de Roder (2008, figure 1). Ce programme ne proposait pas de formations. Après avoir consulté les employeurs locaux du secteur de la santé concernant leurs besoins en personnel, l'entreprise a collaboré avec des collègues communautaires pour structurer et gérer des programmes de certification destinés à ses recrues. Ses diplômés ont obtenu des qualifications telles que techniciens en radiologie ou infirmiers diplômés d'État.

7.1.3 Étude d'impact sectoriel sur l'emploi et évaluation randomisée du projet QUEST. L'organisme qui a mené la dernière étude, Public/Private Ventures,

a entrepris au milieu des années 2000 des évaluations randomisées de trois programmes sectoriels matures. P/PV a également lancé un projet pilote, qui a été mené à terme par un autre groupe, Economic Mobility Corporation (« Mobility »). Ces quatre programmes sectoriels sont aussi diversifiés que possible (Maguire et al., 2010, p. ii ; Roder et Elliott, 2018, p. 26-27).

- Projet QUEST, à San Antonio.

Le Wisconsin Regional Training Partnership (WRTP), une association de syndicats et d'employeurs de Milwaukee, a organisé des formations de 2 à 8 semaines à la demande d'employeurs. Ces derniers se sont engagés à embaucher les participants ayant suivi la formation.

- Le Jewish Vocational Service de Boston (JVS-Boston) a mené des formations intensives, Des formations de cinq mois et demi préparant les participants à travailler dans la facturation et la comptabilité médicales, entre autres activités.

- Per Scholas à New York était une « entreprise sociale » qui remettait à neuf de vieux ordinateurs, Ils les ont distribués à des personnes à faibles revenus et ont utilisé ces activités comme vecteur de formation.

Ces programmes ont permis d'augmenter les revenus de 2 000 à 5 000 \$ par an (Maguire et al., 2010, tableaux 11, 15 et 17 ; Roder et Elliott, 2018, figure 6). Ces résultats, remarquables (quatre sur quatre), ont été immédiatement reconnus comme tels.

7.1.4 Year Up : Premiers résultats randomisés. En 2008, Year Up, une

organisation à but non lucratif qui accompagne les jeunes adultes et noue des partenariats étroits avec de grandes entreprises, a mandaté Mobility pour mener une évaluation préliminaire de ses activités à Boston, New York et Providence. Les premiers résultats en termes de bénéfices étaient tout aussi impressionnants. Cela semble avoir incité Mobility à publier les résultats de l'étude pilote et à obtenir des financements pour un suivi plus long des participants (Roder et Elliott, 2014, p. 1).

7.1.5 Parcours pour l'avancement des carrières et de l'éducation (PACE)

Dans le cadre de l'intérêt fédéral porté aux programmes de parcours professionnels, le Département américain de la Santé et des Services sociaux a lancé une série d'expérimentations appelées PACE (Parcours pour l'avancement des carrières et de l'éducation). Rappelons que les programmes de parcours professionnels accompagnent les participants tout au long de leur scolarité, de leur formation et de leur insertion professionnelle, par étapes de difficulté croissante. Bien que cette progression ne doive pas nécessairement se faire au détriment de la collaboration avec les employeurs – un des volets de la stratégie sectorielle –, dans la pratique, c'est souvent le cas (King et Prince, 2015, p. 202).

Le programme PACE a évalué plusieurs initiatives qui, à juste titre, pouvaient prétendre aux deux appellations, notamment la Valley Initiative for Development and Advancement (VIDA) dans la vallée inférieure du Rio Grande au Texas, inspirée du projet QUEST, et Year Up, cette fois-ci avec un échantillon plus large réparti dans huit villes. Une fois de plus, le programme ayant déjà fait ses preuves a obtenu de loin les meilleurs résultats : Year Up a permis une augmentation exceptionnelle des revenus de 8 000 \$ par an, un impact qui s'est maintenu pendant sept ans (Fein et Dastrup, 2022, annexe 2-1), malgré des coûts particulièrement élevés d'environ 30 000 \$ par participant. Les huit autres évaluations du programme PACE n'ont pas mis en évidence d'impact significatif sur l'emploi ou les revenus.

7.1.6 WorkAdvance

À partir de 2011, MDRC a mené quatre évaluations sous le nom de WorkAdvance. Le terme « avancée » souligne l'importance accordée au soutien et à l'accompagnement des personnes ayant trouvé un meilleur emploi, afin de favoriser leur progression professionnelle après leur entrée dans la stabilité. Une fois de plus, le programme Per Scholas (voir ci-dessus), déjà performant, s'est distingué. Il a incontestablement permis une augmentation des salaires. Les trois autres programmes (Towards Employment, St. Nicks Alliance et Madison Strategies Group ; Schaberg et Greenberg 2020, tableaux 2.1 à 2.4) n'ont pas eu d'effet immédiat. Cependant, 7 à 10 ans après la randomisation, l'impact de St. Nicks a atteint 8 000 \$ par an, avec une signification statistique (Yusim et al. 2025, tableau 3).

7.1.7 Katz et coll. (2022) : Reconnaissance académique

En 2020, les économistes universitaires Lawrence Katz et Jonathan Roth ont collaboré avec Richard Hendra et Kelsey Schaberg du MDRC pour analyser les raisons de l'efficacité de certains programmes sectoriels.

Cet article rassemble les données probantes concernant les trois programmes évalués par P/PV, ainsi que par Project QUEST, Year Up et WorkAdvance. L'étude de Katz et al. (2022) constitue la synthèse de référence sur l'efficacité des programmes sectoriels.

7.2 Quelques points de vue sur l'histoire de la recherche sur les programmes sectoriels

Nous résumons cette histoire en quelques observations :

- La diversité, au sein d'un vaste pays, des acteurs de la mise en œuvre et des bailleurs de fonds ; la flexibilité des organisations privées pour expérimenter ; les mécanismes communautaires pour partager les espoirs et les découvertes ; et la collaboration entre les acteurs, les bailleurs de fonds et les chercheurs ont conduit à un apprentissage collectif lent.

Le placement des études randomisées n'est en réalité pas aléatoire. Plusieurs programmes ont été évalués à plusieurs reprises. De toute évidence, leurs responsables étaient suffisamment attachés à la méthode scientifique pour collaborer avec des chercheurs désireux de vérifier des résultats de recherche surprenants. Dans le cas de Year Up, les résultats d'une étude pilote se sont avérés suffisamment impressionnants pour justifier leur publication et le lancement d'une étude de plus grande envergure. Dans le cas du projet QUEST, des résultats probants ont permis d'obtenir des financements pour poursuivre le suivi des impacts (Roder et Elliott, 2021).

Malgré les mécanismes de sélection et d'auto-sélection, il est clair que : a) certains programmes ont un réel impact sur leurs participants ; et b) ils partagent des caractéristiques communes, notamment un engagement à recueillir et à intégrer des données sur les besoins en compétences des employeurs locaux. L'impact de certains d'entre eux sur les revenus est illustré à la figure 1.

Du point de vue des évaluateurs, les programmes performants se découvrent plutôt qu'ils ne se créent. Cela s'explique notamment par le temps nécessaire à leur maturation. Dix des quatorze cas de réplcation du CET ont été classés comme étant de fidélité faible ou moyenne. Les programmes plus récents de WorkAdvance n'ont pas obtenu d'aussi bons résultats que Per Scholas, du moins pendant les cinq premières années. Toutefois, cette généralisation est à nuancer. Même les réplcations de haute fidélité du programme de San Jose par le CET n'ont pas permis d'augmenter les revenus à long terme. La réplcation la plus réussie mise en évidence par la recherche est l'expansion de Year Up en dehors de Boston (Fein et Dastrup 2022, annexe 2-9).

- Il est pertinent de formuler des hypothèses a posteriori sur les raisons du succès de certains programmes. Cependant, l'exercice est délicat. L'échantillon d'études est restreint. La tentation est grande d'expliquer les échecs en les attribuant à des programmes sectoriels non pertinents. Les comparaisons entre essais doivent être nuancées, car corrélation n'implique pas causalité.

7.3 Réexamen de l'échantillon d'études de Katz et al. (2022)

L'étude de Katz et al. (2022) n'est ni une méta-analyse ni une revue systématique. Son objectif est d'utiliser les résultats de certaines évaluations encourageantes pour tester des théories économiques expliquant leur efficacité.

Néanmoins, étant donné que l'étude de Katz et al. est devenue une preuve concrète de l'efficacité des programmes sectoriels, il convient d'examiner de manière critique dans quelle mesure ses 9 études portant sur 8 programmes représentent l'ensemble des données probantes sur les programmes sectoriels.

Reconsidérer le choix des études soulève une question de définition : qu'est-ce qu'un programme sectoriel ?

Le CET est-il valable même s'il est antérieur à la définition du terme ? Qu'en est-il de ses répliques de faible fidélité ?

Le concept de « programme sectoriel » se révèle être un ensemble d'attributs potentiels mal définis. L'effort le plus systématique pour concrétiser ce concept est présenté dans la méta-analyse de Peck et al. (2021), examinée au paragraphe 6.4. Parmi les plus de 75 caractéristiques de programme recensées par cette analyse, une quinzaine concernent la collaboration avec les employeurs. Par exemple : la mise en place de conseils consultatifs pour recueillir des avis de haut niveau sur la conception du programme ; ou encore l'engagement des entreprises à embaucher les diplômés ou à fournir un mentorat. Ces caractéristiques sont listées à gauche du tableau 4 ci-dessous ; ce tableau présente également la notation attribuée par Peck et al. aux huit programmes étudiés par Katz et al.

Trait	Emploi sectoriel								
	RYTHME	Étude d'impact			Avance de travail				
	Projet QUÊTE	Année En haut	JVS- Boston	Par Scholas	Wisconsin Régional Entraînement	Madison Stratégies Groupe	Vers Emploi	Par Scholas	Saint Nicks Alliance
Les efforts du programme pour impliquer les employeurs									
Convoquer un conseil consultatif des employeurs									
Des programmes d'études adaptés aux besoins des employeurs									
Partenaire officiel									
Invitation à des événements									
Offrir une formation aux employés en poste									
Autre									
Cotisations de l'employeur									
Contribution à l'élaboration des programmes d'études									
Informations sur le type de candidats									
Instruction ou instructeurs									
Services d'orientation professionnelle									
Formation rémunérée en milieu de travail									
Apprentissage non rémunéré en milieu de travail									
Mentors									
Aide financière aux étudiants									
Informations sur la demande du marché du travail									
S'engager à embaucher des									

diplômés Tableau 4. Codage des caractéristiques d'engagement des employeurs par Peck et al. (2021) des programmes dans Katz et al. (2022)

Remarques : Peck et al. (2021) codent les programmes dans 46 études portant sur plus de 75 caractéristiques. Les codages relatifs au rôle des employeurs sont présentés ici.

Scholas a été évalué à deux reprises et a reçu des codages légèrement différents à chaque fois. Les efforts du programme visent à garantir que les « programmes d'études adaptés à

En pratique, il est difficile de distinguer les « besoins des employeurs » et leurs contributions à l'« élaboration des programmes ou des cursus ».

Les études présentées sont celles décrites dans Katz et al. (2022). PACE est le programme d'évaluation des parcours de carrière et d'éducation. ⁵²

⁵² Source : ensemble de données à usage public accompagnant Peck et al. (2021) (Département du Travail des États-Unis, Bureau en chef de l'évaluation, s.d.).

En matière de collaboration avec les employeurs, Year Up est de loin le programme le plus complet. Il fait partie des deux seuls programmes où les entreprises partenaires s'engagent à embaucher les diplômés, ou plus précisément, à leur proposer un stage de six mois à l'issue de la formation (Fein et Dastrup, 2022, p. 4). Seul Year Up fait en sorte que ses partenaires fournissent des mentors.

Lorsqu'on examine le tableau par lignes plutôt que par colonnes, deux caractéristiques quasi synonymes se dégagent comme étant universelles : « des programmes d'études adaptés aux besoins des employeurs » et la participation des employeurs à l'élaboration des programmes. Nous assimilons donc provisoirement le terme « programme sectoriel » à l'implication des employeurs dans les programmes d'études. Cela rejoint l'idée centrale de créer un lien fort entre la pédagogie et les besoins du monde du travail. Comme mentionné au paragraphe 6.4, Peck et al. constatent que cette caractéristique est l'un des facteurs les plus fortement corrélés à un ⁵³ (Dans notre méta-analyse du §6.5, ce impact positif sur les salaires. (Ce facteur explicatif est toutefois moins important que l'engagement des employeurs à embaucher des diplômés.)

Forts de cette définition opérationnelle d'un programme sectoriel, nous nous sommes interrogés sur les programmes sectoriels ayant fait l'objet d'études randomisées et non couverts par Katz et al. Le tableau 5 répond à cette question. Il est construit à partir de la liste des programmes évalués par randomisation que Peck et al. considèrent comme présentant l'une ou l'autre de ces deux caractéristiques quasi identiques. Trois programmes ont ensuite été exclus, soit faute de preuves textuelles solides démontrant leur implication significative des employeurs dans l'élaboration des programmes d'études, soit en raison de l'absence d'estimations d'impact à moyen ou long terme. Par ailleurs, le CET et ses ⁵⁴ De l'autre côté répliqués, exclus de l'étude de Peck et al. car ne relevant pas de la catégorie des « parcours de carrière », ont été ajoutés.

⁵³Techniquement, cette affirmation se rapporte au second critère, à savoir la participation des employeurs. Cependant, les deux sont corrélés à 0,58 (Peck et al., 2007, annexe E-9), ce qui rend difficile la distinction de leurs liens avec les résultats. Les programmes ⁵⁴Green Jobs and Health Care à Minneapolis, IBEST et YouthBuild sont exclus sur le premier critère et Career Pathways en Louisiane sur le second.

Programme	Des preuves de l'implication de l'employeur ont-elles été obtenues ?	Dernier suivi	Source de résultats (taux de réponse)	À Katz et al.
Centre pour l'emploi et Formation (San José)	Points forts : « Les programmes de formation continue impliquent les employeurs dans la conception de leurs programmes et en tant qu'évaluateurs. » programmes de formation	Q7–10 (adultes), Q9–16 (jeunes) Année 5	Auto-déclaration (80 %, 84 %)	Non
Répliques CET	Résultats mitigés : 6 programmes ont obtenu une note élevée, 5 une note moyenne et 1 une note faible concernant « l'implication des employeurs dans la conception et l'entraînement »		Auto-déclaration (77 %)	Non
Étude d'impact sectoriel sur l'emploi				
WRTP	Points forts : Des comités employeurs-syndicats élaborent des formations ; les employeurs ont « commandé » des formations	Q5–8	Auto-déclaration (87 %)	Oui
JVS–Boston	Points forts : Les comités consultatifs des employeurs « ont contribué à l'élaboration du programme de formation aux compétences professionnelles ».	Q5–8	Auto-déclaration (73 %)	Oui
Per Scholas	Points forts : Les employeurs « donnent des conseils sur le programme d'études »	Q5–8	Auto-déclaration (78 %)	Oui
Emplois verts et soins de santé				
Bakersfield, Californie	Points forts : Les employeurs ont fourni des orientations sur le contenu des cours ; les programmes ont été adaptés à leurs besoins (Q5-6). Points		Officiel	Non
Gainesville, Texas	faibles : Pour certains programmes, les formateurs ont consulté quelques employeurs afin d'harmoniser les programmes (Q5-6). Points modérés : Les		Officiel	Non
Grand Rapids, Michigan	employeurs ont initialement élaboré les programmes, mais leur rôle s'est réduit.	Q5–6	Officiel	Non
Projet QUEST	Fort : Des membres du personnel détachés auprès de l'université repensent le programme de mécanique diesel « en consultation avec « employeurs locaux » ; le programme de formation des infirmières auxiliaires a été modifié pour inclure les soins gériatriques en fonction de la demande des employeurs.	Année 11	Officiel	Oui
Formation accélérée pour Fabrication en Illinois	Résultats mitigés : « Certaines régions ont réussi à obtenir la contribution des employeurs. »	Q5–8	Officiel	Non
Accélérer les connexions à l'emploi	Points forts : Conseils aux employeurs sur la conception des programmes et des projets ; contenu adapté aux besoins des employeurs (Q5-8)		Officiel	Non
Avance de travail				
Per Scholas	Points forts : Les employeurs « donnent des conseils sur le programme d'études »	Année 6	Officiel	Oui
Alliance de St. Nicks	Points faibles : contribution limitée à la préparation à l'emploi ; personnel « jamais vraiment engagé au service des employeurs ».	Année 6	Officiel	Oui
Stratégies de Madison	Points forts : Groupes consultatifs ; citation d'un employeur sur l'élaboration du programme ; modifications apportées aux cours suite aux retours d'information.	Année 6	Officiel	Oui
Vers l'emploi	Niveau modéré : Programme adapté aux besoins des employeurs, mais principalement par le biais d'intermédiaires.	Année 6	Officiel	Oui
RYTHME				
Year Up (8 villes)	Points forts : Les employeurs façonnent directement les programmes de formation professionnelle ; les entreprises d'accueil de stage fournissent des retours d'information.	Année 7	Officiel	Oui
Carreras en Salud (Chicago)	Modéré : Le conseil semestriel discute des « modifications potentielles du programme », mais la formation reste standard. programme de soins infirmiers	Année 6	Officiel	Nouveau
Formation de la main-d'œuvre Académie (Des Moines)	Modéré : Ce sont les employeurs, et non le contenu des programmes, qui ont déterminé les certifications à proposer.	Année 6	Officiel	Nouveau
VIDA (Rio Grande)	Point faible : Ce sont les employeurs qui déterminent les programmes universitaires à financer, et non les programmes d'études.	Année 7	Officiel	Nouveau

Tableau 5. Force des preuves en faveur de l'implication des employeurs dans l'élaboration des programmes d'études des secteurs sujets à débat.

Remarques : Q = trimestre. Hormis le CET et ses répliques, tous les programmes sont identifiés par Peck et al. (2021) comme impliquant les employeurs dans l'élaboration des programmes d'études. Trois Les mentions « Green Jobs/AIOIC », « PACE/I-BEST » et « PACE/PCPP » sont omises faute de contexte textuel. Sources : Burghardt et al. (1992), Rangarajan, Burghardt. et Gordon (1992), Cave et al. (1993), Miller et coll. (2005), Maguire et coll. (2010), Martinson et coll. (2016), Copson et coll. (2016), Rademacher, Bear et Conway (2001), Roder et Elliott (2021), Betesh et al. (2015), Modicamore et al. (2017), Schaberg et Greenberg (2020), Gardiner et Juras (2019), Copson, Martinson, et Gardiner (2014), Gardiner et Grittner (2022), Rolston, Copson et Gardiner (2017), Farrell et Martinson (2017).

Cet exercice nous confronte à une distinction entre intention et exécution. Les évaluateurs

Il a été jugé que seulement 6 des 14 sites de réplique du CET ont atteint un niveau de fidélité élevé en ce qui concerne « l'implication des employeurs dans la conception et la formation » (Miller et al., 2005, tableau 1.1). (Par ailleurs, comme indiqué précédemment, seuls 4 sites ont été jugés comme ayant répliqué le CET avec une fidélité globale élevée.) On peut donc affirmer que les sites ayant obtenu un faible niveau de fidélité devraient être exclus du corpus de données probantes relatif à la stratégie sectorielle. En d'autres termes, ils ne sont pas pertinents car ils n'ont pas mis en œuvre le dispositif, ou ne l'ont pas fait correctement.⁵⁵ Cela se justifie lorsque l'objectif est de tirer des leçons du succès.

Mais il est parfois plus pertinent de généraliser à partir de ce qui était prévu plutôt que de ce qui a été réalisé.

Pour un décideur qui doit décider d'approuver ou non un projet visant à étendre la formation sectorielle à un plus grand nombre de personnes, la question n'est pas tant de savoir s'il faut le faire, mais plutôt s'il faut soutenir les personnes qui le portent. Le groupe de référence approprié pour prévoir les impacts est alors l'ensemble des projets antérieurs ayant des objectifs similaires.

Puisqu'il est pertinent d'analyser à la fois les programmes sectoriels envisagés et les programmes sectoriels mis en œuvre, le tableau 5 répertorie les programmes correspondant à la définition la plus inclusive (intention) tout en fournissant des informations sur la définition plus exigeante (mise en œuvre). De plus, ce tableau situe l'étude de Katz et al. dans le contexte plus large des données probantes sur les programmes sectoriels. Il précise la durée du suivi des participants, car Katz et al. n'ont inclus que les études présentant des estimations d'impact à moyen terme (portant sur deux ans ou plus après la randomisation). Enfin, il indique si les données de résultats proviennent d'auto-déclarations ou de sources officielles et, dans ce dernier cas, le taux de réponse.

Globalement, étant donné notre intérêt à généraliser à partir de l'ensemble des données probantes sur les programmes sectoriels (potentiels), il est pertinent d'ajouter 10 évaluations randomisées de programmes sectoriels aux 9 (sur 8 programmes) de Katz et al. Il s'agit du CET et de ses répliques ; de trois des expériences Green Jobs and Health Care (Martinson et al. 2016) ; du programme Accelerated Training for Illinois Manufacturing (Betesh et al. 2015) ; des programmes Accelerating Connections to Employment dans le Maryland et le Texas (Modicamore et al. 2017) ; et de trois programmes PACE en plus de Year Up, dont VIDA dans la région du Rio Grande au Texas.

Il est certes discutable que le critère que nous avons retenu comme déterminant – l'implication des employeurs dans les programmes d'études – le soit véritablement. Néanmoins, nous pensons que cet exercice élargit la catégorie de référence des programmes sectoriels d'une manière pertinente et peu arbitraire.

⁵⁵Il est tout à fait possible que certaines déviations par rapport au modèle CET aient rendu les répliques plus efficaces.

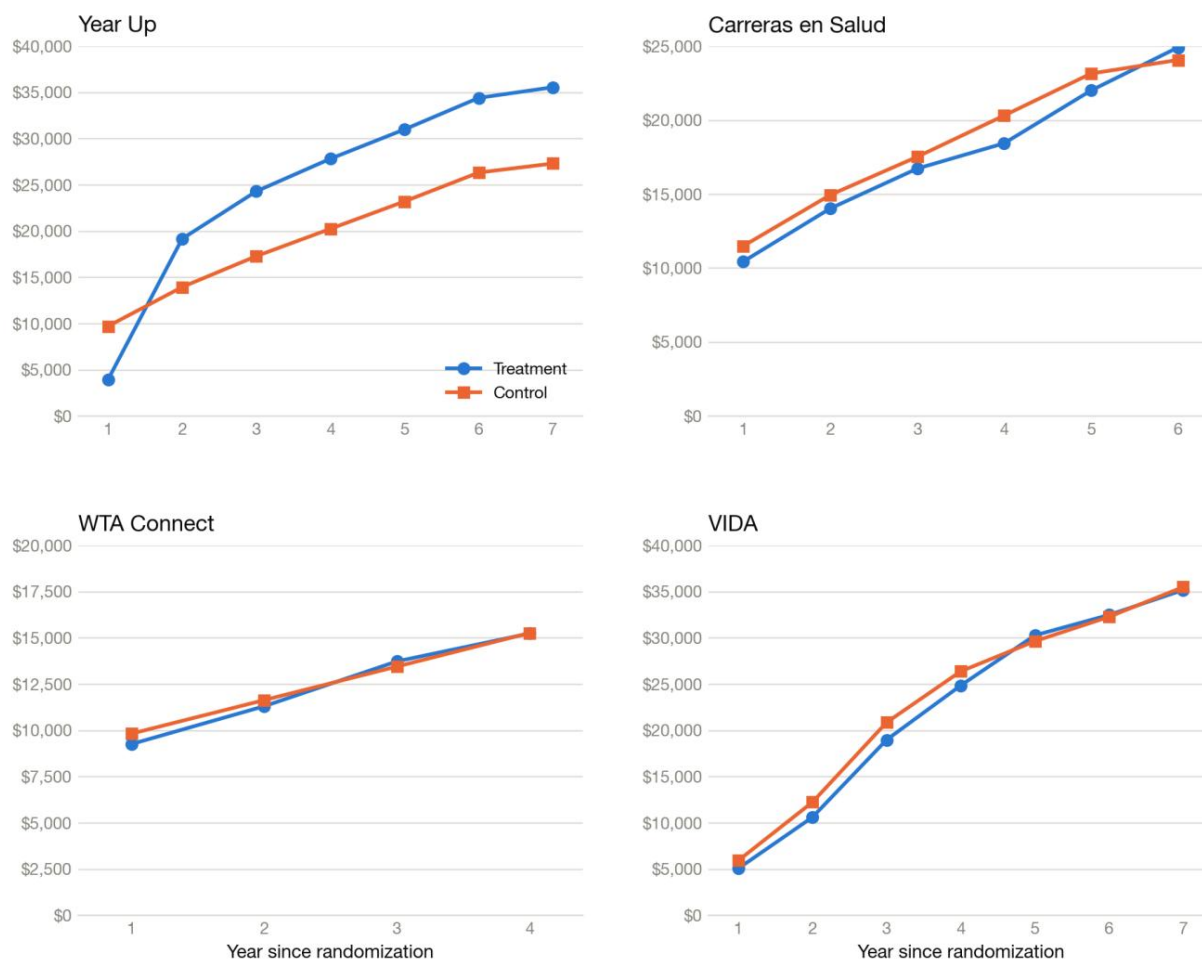


Figure 12. Revenus annuels des participants à certains programmes évalués par PACE. Remarques : Adapté de Juras et al. (2022). Sont présentés les résultats des quatre programmes de l'évaluation Pathways for Advancing Careers and Education (PACE) qui correspondent le mieux à l'étiquette « stratégie sectorielle ». La randomisation a eu lieu au trimestre 0.

Pour illustrer l'éventail des performances au sein de ce groupe plus large, la figure 12 présente les revenus des groupes expérimental et témoin des quatre programmes évalués par PACE. L'un des graphiques se distingue des autres. Les étudiants du programme « Year Up » gagnent moins durant leur première année de cours et de stage, mais leurs revenus dépassent ensuite ceux de leurs pairs du groupe témoin. Par la suite, l'écart entre les deux groupes se stabilise autour de 2 000 \$ par trimestre et par étudiant du groupe expérimental. Ce programme permet (en moyenne) aux étudiants d'accéder à des carrières mieux rémunérées. Les trois autres programmes PACE n'atteignent pas un tel résultat. succès.

Les estimations d'impact méta-analytiques pour la famille plus large de programmes sectoriels sont rapportées dans le tableau 6. Pour contextualiser, les figures 13 et 14 regroupent les estimations de l'ITT issues des différentes études. Sur ces graphiques en forêt, les estimations centrales sont représentées par des points et les intervalles de confiance à 95 % par des barres d'erreur ; les résultats sont également présentés sous forme numérique à droite. Les impacts sur les revenus annuels sont ajustés de l'inflation en dollars de 2025. Une moyenne à effets aléatoires figure au bas de chaque graphique.

	Contrôle signifier	Impact	N
Taux d'adoption du programme (%)	13,5 (6,5)	55,5 (8,6)	17
Taux de participation à la formation (%)	53,8 (4,6)	24,5 (4,2)	8
Emploi, court terme (%)	67,5 (6,0)	3,7 (1,4)	12
Emploi, moyen terme (%)	73,2 (4,8)	2,9 (0,9)	10
Emploi, à long terme (%)	80,3 (0,6)	1,3 (0,8)	6
Revenus à court terme (en dollars de 2025)	16 119 (1 048)	1 169 (1 092)	13
Revenus à moyen terme (en dollars de 2025)	19 447 (938)	3 149 (1 260)	12
Revenus à long terme (en dollars de 2025)	25 137 (1 664)	3 703 (1 545)	9

Tableau 6. Moyennes du groupe témoin et estimations d'impact dans les expérimentations sectorielles. Remarques : Les impacts sont estimés par méta-analyse du maximum de vraisemblance restreint (REML). Les statistiques du groupe témoin sont calculées à l'aide des pondérations REML correspondantes. Les chiffres entre parenthèses sous chaque estimation représentent les écarts-types des valeurs du groupe témoin et les erreurs-types des estimations d'impact. La taille des échantillons varie en raison des données manquantes. Les résultats à court terme sont suivis pendant environ un an, ceux à moyen terme pendant environ deux ans et ceux à long terme pendant environ trois à cinq ans.

Les graphiques en forêt présentent des résultats mitigés, avec une prédominance de bonnes nouvelles. Dans les quatre graphiques, les estimations ponctuelles varient de valeurs négatives à positives. Cependant, les moyennes à effets aléatoires sont positives, avec une significativité statistique à $p < 0,05$ dans tous les cas, sauf le quatrième (impact à long terme sur l'emploi). Sous réserve du fait que l'échantillon d'études varie d'un graphique à l'autre, les programmes sectoriels entraînent en moyenne une hausse de l'emploi, partiellement transitoire (environ 3 points), et une augmentation plus durable du salaire brut (avant impôt), de 3 000 à 4 000 dollars par an en dollars de 2025.

La prise en compte des répercussions sur le taux de participation conduit à des taux d'engagement (LATE) environ deux fois plus élevés. L'expérimentation sectorielle moyenne a généré un écart de 56 points dans la participation au programme évalué, comme le montre le tableau 6 ; cela signifie que chaque programme LATE représente 1,79 fois le taux d'engagement initial (ITT) correspondant. Comme pour les programmes de formation en général, l'impact sur le taux de participation à toute formation est d'environ 25 %. Toutefois, si les programmes de formation alternatifs sont nettement moins efficaces que les programmes sectoriels testés, alors la formation LATE, à 4 fois la formation initiale, ne doit pas être interprétée comme une estimation de l'impact des programmes sectoriels.

Les moyennes des groupes témoins présentées dans le tableau 6 reflètent également les différences entre les populations ciblées, témoignant de la sélectivité des programmes sectoriels. Pour les expériences intégrant une formation, le taux d'emploi à long terme des groupes témoins est de 60 %, contre 80 % pour les programmes sectoriels. Les groupes témoins issus de ces programmes perçoivent également des revenus supérieurs de 50 à 75 % selon l'horizon temporel. Les programmes sectoriels recrutent parmi des populations mieux rémunérées même sans formation.

Bien que les avantages des programmes sectoriels dépassent largement ceux d'une intervention de formation primaire moyenne, leurs coûts, eux, ne suivent pas cette tendance : 11 602 \$ par membre du groupe de traitement au lieu de 13 598 \$. Par conséquent, nous estimons le ratio avantages-coûts des programmes sectoriels à 7,18, soit près de cinq fois le ratio de 1,44 du groupe principal. (Voir le tableau G2 en annexe.) Si le gouvernement finançait un programme sectoriel statistiquement moyen, il serait largement rentabilisé. De ce fait, la valeur maximale du programme (VMP) serait infinie, ce qui signifie qu'un bienfait serait produit à un coût budgétaire négatif.

En résumé, cet élargissement stratégique du corpus d'études présenté dans Katz et al. renforce fondamentalement l'optimisme quant aux programmes sectoriels. Même en prenant en compte les programmes qui n'ont pas réussi à collaborer étroitement avec les employeurs, l'impact sur les revenus est 3 à 5 fois supérieur à celui d'un programme de formation professionnelle classique.

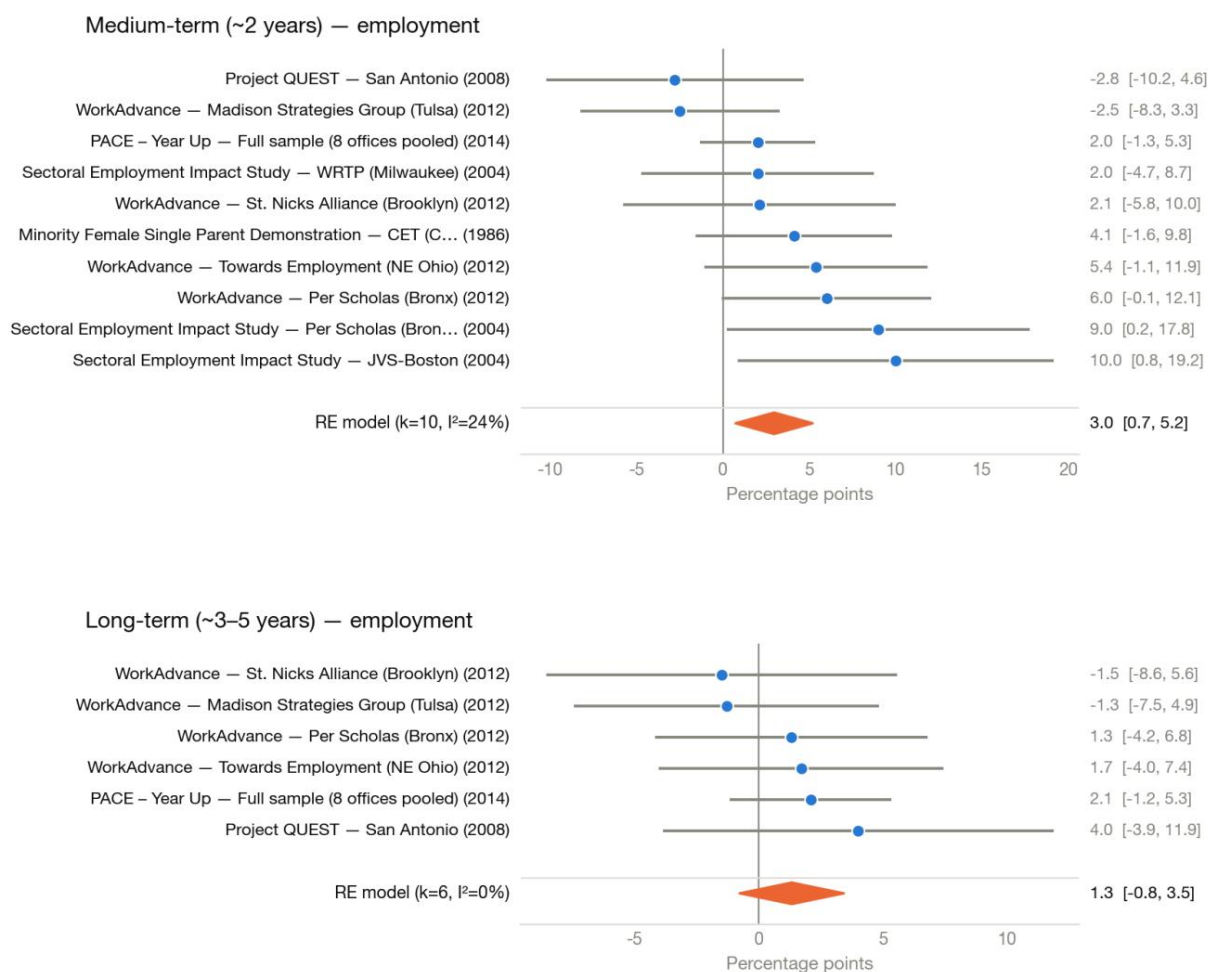


Figure 13. Impact des programmes sectoriels sur l'emploi.

Remarques : « RE » désigne l'effet moyen issu d'une méta-analyse à effets aléatoires. La longueur des barres d'erreur et des losanges représente les intervalles de confiance à 95 %, également indiqués numériquement à droite. Tous les résultats proviennent des sources citées dans le tableau 5. Il s'agit d'impacts en intention de traiter (ITT).

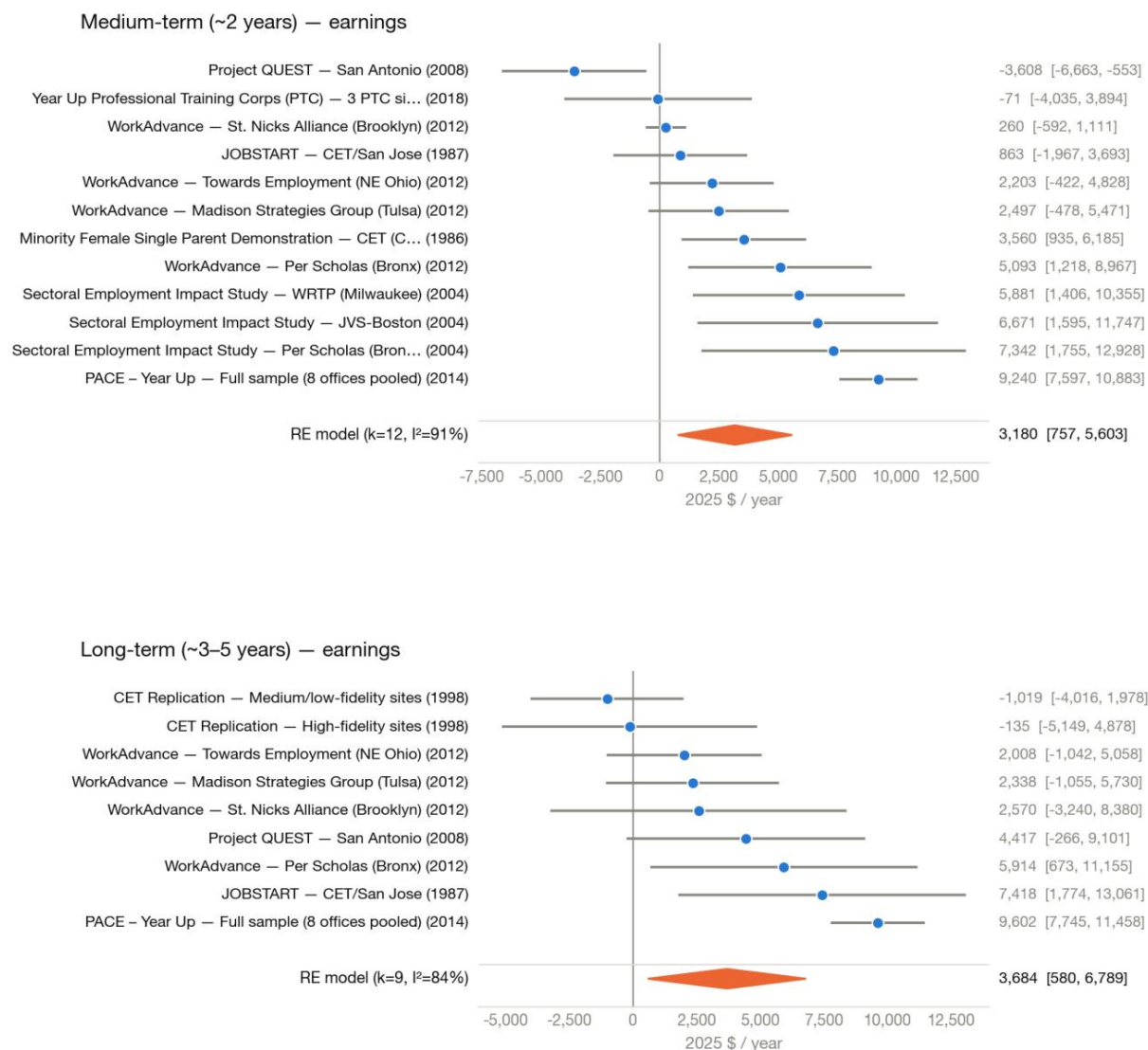


Figure 14. Impact des programmes sectoriels sur les revenus (en \$2025 par an)

Remarques : Voir les notes relatives à la figure précédente.

7.4 Programmes sectoriels : Résumé

Pourquoi certains programmes sectoriels, comme Project QUEST, Per Scholas et Year Up, se distinguent-ils dans les études ? Les programmes sectoriels performants possèdent généralement les composantes suivantes :

- Sélectivité. Les quatre programmes WorkAdvance ont procédé à un « dépistage intensif » de
Les candidats étaient sélectionnés sur la base de l'évaluation par le personnel de qualités subtiles telles que l'engagement et la motivation. Environ 20 % des candidatures ont été retenues (Hendra et al., 2016, p. 29). En 2025, 70 000 personnes ont postulé pour 5 000 places chez Per Scholas. L'organisation utilisait des critères comportementaux.

des évaluations et autres tests, y compris un travail préparatoire, pour sélectionner les candidats et tester leur engagement à travers la charge du processus lui-même.

56

- Formation portant sur les « compétences générales » (comportement lors des entretiens et au travail) et les compétences professionnelles, comme l'installation de panneaux solaires.
- Un accompagnement individualisé pour soutenir les élèves et les aider à progresser.
- L'insertion professionnelle après l'obtention du diplôme. Year Up tisse des liens étroits avec les employeurs : ils prennent en charge un Per Scholas prend en charge une part importante des coûts de formation et s'engage à embaucher ses diplômés. L'entreprise consacre du personnel au maintien de relations avec les principaux employeurs et au suivi des tendances régionales en matière de demande de compétences. Au fil du temps, les programmes gagnent la confiance des employeurs quant à l'aptitude de leurs diplômés à intégrer le marché du travail.

L'importance exacte de chacun de ces éléments reste floue et mérite d'être approfondie. Les deuxième et troisième sont les moins distinctifs. Toutefois, l'implication étroite des acteurs des deux côtés du marché du travail semble essentielle. Pour que cette stratégie ait un impact significatif, les employeurs doivent recruter des personnes qu'ils auraient normalement écartées pour certains postes, et qui, de leur côté, n'auraient peut-être jamais envisagé ces emplois comme accessibles. Ce rôle de passerelle permet aux stagiaires d'accéder à des parcours professionnels nettement supérieurs.

Plus largement, nous tirons de cette histoire plusieurs principes qui ne reposent pas sur les données quantitatives en tant que telles, mais qui éclairent nos extrapolations :

Transmettre aux individus des compétences et les préparer à de nouveaux contextes de travail relève d'un art subtil et complexe. Sur le plan administratif, ces tâches sont bien plus ardues que le versement des allocations chômage. Ce sont les petites organisations privées, agiles et performantes, qui ont obtenu les meilleurs résultats. Si l'accès à diverses formations est d'intérêt public, leur mise en œuvre efficace s'avère plus difficile pour les institutions publiques, notamment en raison de leur difficulté à exercer une sélection rigoureuse. De plus, des incitations institutionnelles perverses, telles que les critères de « performance » du JTPA, peuvent compromettre l'équilibre fragile d'un programme de formation professionnelle efficace.

- Bien que les programmes sectoriels n'aient certainement pas atteint leurs limites, il est tout aussi certain que ces limites existent. L'une des raisons pour lesquelles Year Up s'est étendu de Boston à d'autres villes, plutôt que de se concentrer sur sa ville d'origine, tient au fait que son approche repose sur des partenaires corporatifs aux moyens financiers importants. Ces derniers financent une grande partie de la formation et s'engagent à embaucher les diplômés. Dans chaque ville, le nombre de tels partenaires est limité. Les stages, d'une durée relativement⁵⁷ En effet, le programme Year Up dure un an, moitié entraînement, moitié... longue comparée aux programmes sectoriels, constituent un autre aspect important. Ainsi, d'après les données disponibles, les programmes sectoriels efficaces se limitent aux compétences pouvant être acquises en quelques mois.

56Conversation avec Tamara Johnson et Sang Lee, Per Scholas, 28 avril 2026.

57Conversation avec Garret Warfield, ancien directeur de la recherche, Year Up United, 24 avril 2026.

8 Des preuves du passé à un avenir bouleversé par l'IA

Nous ne disposons d'aucune donnée de recherche sur l'efficacité de la formation professionnelle pour faciliter l'adaptation des individus lorsque de grands modèles linguistiques entrent en fonction. Afin d'apporter un éclairage partiel, ce rapport examine néanmoins les données disponibles sur les programmes de formation professionnelle.

Presque tous les programmes faisant l'objet de recherches ont été subventionnés par des gouvernements ou des fondations. La plupart de ces formations s'adressaient aux personnes à faibles revenus, tandis qu'une minorité était destinée aux personnes ayant perdu leur emploi ou présentant un handicap intellectuel. À l'exception possible des formations dispensées par le programme d'aide à la reconversion professionnelle (TAA) aux États-Unis, presque toutes les formations étaient de courte durée, c'est-à-dire bien inférieures à un an.

L'une des principales limites de ce rapport est qu'il ne couvre que les programmes de formation professionnelle autonomes. Ce ne sont là que quelques-uns des moyens d'acquérir des compétences professionnelles. Notre étude ne s'arrête pas là : les formations professionnelles au lycée, l'apprentissage, les IUT, les universités, les écoles professionnelles et la formation en ligne. De nombreuses études estiment dans quelle mesure d'autres formes de formation augmentent les revenus, certaines s'appuyant sur des méthodologies prometteuses comme les loteries (par exemple, Grosz, 2020). Aux États-Unis, les programmes de formation professionnelle se prêtent particulièrement bien aux expérimentations randomisées, car bien que financés par des fonds publics, ils ne sont pas considérés comme un droit fondamental. De ce fait, ils sont soumis à la pression de prouver leur efficacité, tout en conservant une certaine latitude pour permettre – à des fins de recherche – de refuser ou de reporter l'accès de manière aléatoire.

Principaux résultats :

- La quasi-totalité des preuves hautement crédibles proviennent d'essais randomisés menés aux États-Unis.

Les évaluations du programme national de démonstration de travail soutenu mené dans plusieurs villes dans les années 1970 et du Job Training Partnership Act dans les années 1980 ont mis en évidence des bénéfices pour les adultes, mais pas pour les jeunes. Le Job Corps, un autre programme destiné aux jeunes, n'a également montré aucun impact, selon les données officielles, sur l'emploi et les revenus. L'évaluation la plus récente du principal programme national, l'étude WIA Gold Standard, n'a donné aucun résultat. Cela pourrait s'expliquer en grande partie par un manque de puissance statistique : de nombreuses personnes ayant reçu une offre de formation ne l'ont pas suivie, et beaucoup d'autres, n'en ayant pas reçu, l'ont trouvée ailleurs.

- La poignée d'études randomisées et basées sur la discontinuité réalisées en Europe n'ont pas été très informatives quant aux impacts de la formation (voir annexes C et D).

Les méta-analyses s'accordent à constater de faibles effets positifs moyens. Il existe peu de consensus sur les corrélats d'impact, tels que le ciblage des femmes.

- Une nouvelle méta-analyse d'études randomisées menées aux États-Unis indique que l'impact à long terme (analyse en intention de traiter) de la formation professionnelle se traduit en moyenne par une augmentation de l'emploi de 1,7 à 1,8 point et un gain de revenu annuel de 700 à 800 dollars. Nous estimons que l'impact sur le revenu avant impôt dépasse largement le coût des programmes.

- Dans cette même analyse, l'impact moyen sur le taux de participation à toute formation — y compris les formations hors du cadre de l'expérience — n'est que de 28 points de pourcentage. Avec des hypothèses supplémentaires, cela signifie que l'effet moyen de la formation sur les participants marginaux est près de quatre fois supérieur.

Deux études sophistiquées de randomisation par un juge, menées aux États-Unis et au Danemark, montrent que l'aide apportée aux personnes ayant perdu leur emploi contribue à leur adaptation, notamment en les aidant à se reconvertir professionnellement, voire à déménager. Bien que moins rigoureuses que les essais randomisés, ces études constituent des arguments solides en faveur d'une identification causale. Elles portent par ailleurs sur des programmes pour lesquels nous disposons de peu ou pas de données randomisées.

Bien que non mentionnées précédemment, la méta-analyse inclut deux études de programmes pilotes pour les travailleurs licenciés : le projet de démonstration de réinsertion professionnelle des bénéficiaires de l'assurance chômage du New Jersey et le projet de démonstration d'adaptation des travailleurs du Texas. Au cours des douze mois de suivi, l'aide à la recherche d'emploi et la formation ne donnent pas de meilleurs résultats que l'aide seule (Corson et al., 1989, tableau VI.1 ; Bloom et al., 1990, tableau 7.5).

- Il y a de bonnes raisons de croire que certains « programmes sectoriels » accomplissent un certain nombre de Leur impact est nettement supérieur à celui des programmes de formation initiale moyens. Il semble toutefois que ce résultat soit dû en partie à une sélection très rigoureuse. D'autres facteurs peuvent limiter leur portée, comme la disponibilité d'employeurs disposant de moyens financiers importants et capables de s'engager à embaucher leurs diplômés, ainsi que le nombre de candidats immédiatement opérationnels.

Que nous apprennent ces résultats sur la manière dont la société devrait aider les travailleurs déplacés par l'IA ?

Un article de Manning et Aguirre (2026) présente le contexte de ce défi. À partir de sept ensembles de données sur les travailleurs et les professions aux États-Unis, le projet estime : 1) le nombre de personnes exerçant actuellement 356 professions ; 2) un indice d'exposition à l'IA mesurant la capacité de l'IA à réaliser chacune de ces tâches ; 3) un indice de « capacité d'adaptation » pour les personnes en emploi, prenant en compte leur patrimoine, leur âge, la densité d'emplois par kilomètre carré dans leur zone de travail et le nombre d'emplois dans des professions connexes exigeant des compétences similaires. La figure 15, extraite de l'étude, présente les principaux résultats.

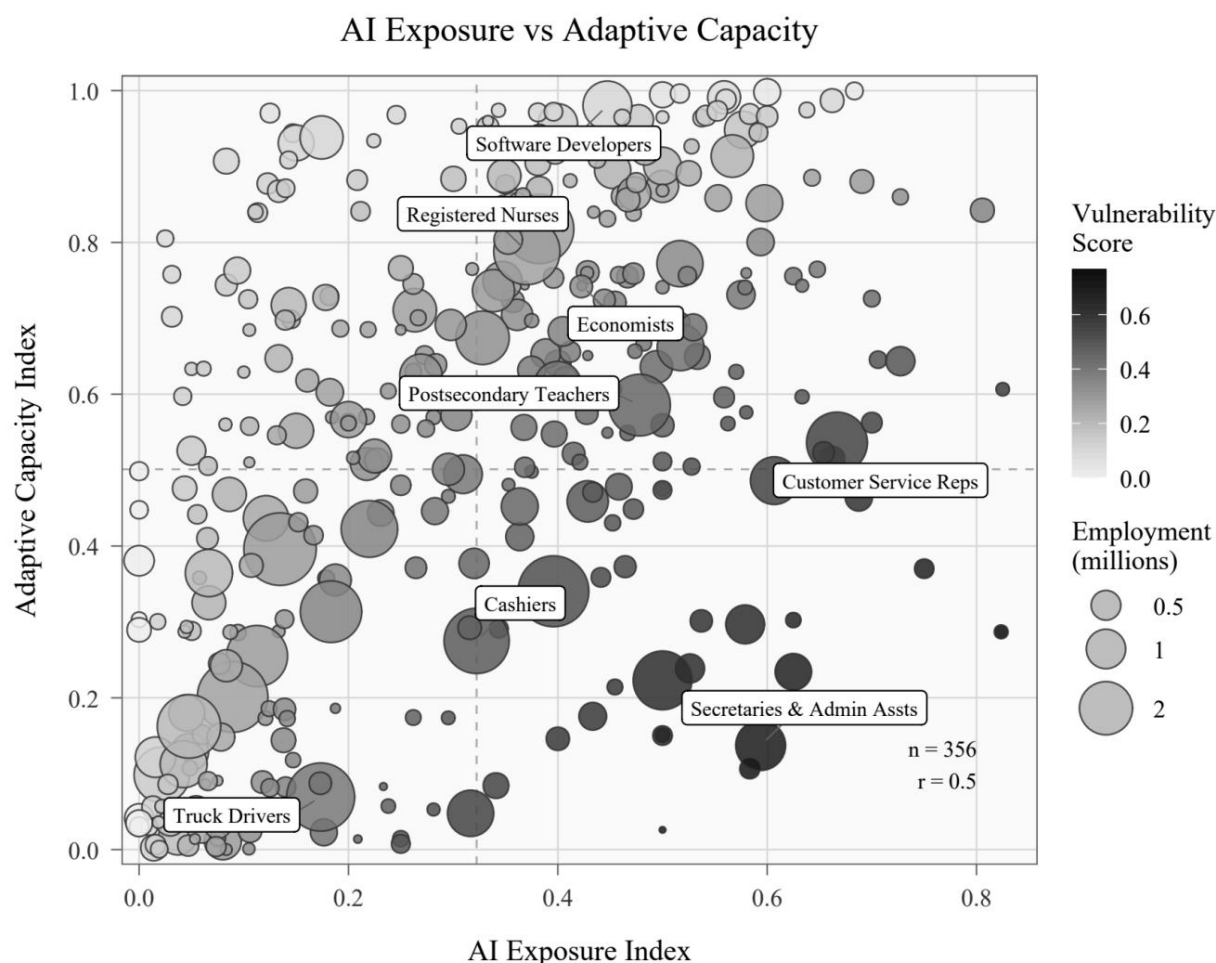


Figure 15. Capacité d'adaptation vs. exposition à l'IA pour 356 professions aux États-Unis, d'après Manning et Aguirre (2026)

En moyenne, les travailleurs considérés comme plus exposés à l'IA ont également plus de moyens pour s'adapter.

C'est plutôt rassurant. Il existe néanmoins d'importantes exceptions. Par exemple, 1,7 million de personnes travaillent comme « secrétaires et assistants administratifs, à l'exception des secteurs juridique, médical et de la direction » ; leur niveau d'exposition à l'IA est élevé (59 %) et leur capacité d'adaptation faible (14 %), ce qui les place en bas à droite du graphique. Parmi les professions similaires, on trouve les « agents de paie et de gestion du temps » et les « employés de bureau » – en général, des cols blancs peu qualifiés. Les méthodes d'analyse montrent que les chauffeurs routiers sont relativement peu⁵⁸ Manning et Aguirre exposés à l'IA (voir en bas à gauche du graphique), bien que les tendances plus générales en matière d'automatisation – qu'on les considère ou non comme de l'IA – puissent certainement affecter ces emplois. Cela augmenterait encore la taille de la catégorie des travailleurs fortement exposés et peu adaptables.

Parallèlement, même les professionnels hautement qualifiés, qui ont une plus grande capacité d'adaptation, peuvent devoir franchir des fossés professionnels plus importants pour retrouver leur ancien niveau de revenus.⁵⁹ Si un avocat en milieu de carrière avec des enfants,

⁵⁸Sur les défis auxquels ces travailleurs sont confrontés, voir Molly Kinder, « La perturbation invisible », publié sur Substack le 9 juin 2026.

⁵⁹Conversation avec Molly Kinder, Brookings Institution, 21 mai 2026.

Si, en raison d'un prêt hypothécaire et de prêts contractés pour leurs études de droit, ils sont contraints de quitter leur profession, ils peuvent être confrontés à un choix : soit une baisse de salaire importante et permanente, soit un retour aux études pour des années de reconversion.

Le problème étant ainsi posé, nous proposons quelques enseignements tirés des éléments de preuve que nous avons examinés.

8.1 Le type de programmes de courte durée qui sont mis en avant dans cette étude pourrait aider certains travailleurs « peu qualifiés » déplacés par l'IA.

Parmi les professions potentiellement vulnérables, on trouve les conducteurs – de taxis, de poids lourds ou même de véhicules de livraison. Ces emplois n'ont pour la plupart pas disparu car, d'un point de vue technique, ils requièrent des compétences pointues et d'excellentes capacités cognitives. En revanche, d'un point de vue économique, ils sont considérés comme peu qualifiés, car l'offre de conducteurs qualifiés est importante et leur coût faible.

L'automatisation et la disparition des emplois à bas salaire ne sont pas un phénomène nouveau. Depuis des décennies, les robots et la concurrence étrangère contribuent au déclin de l'industrie industrielle du Sud des États-Unis. L'informatisation a également supprimé de nombreux emplois peu rémunérés, comme ceux de coursier et de dactylographe. Nos estimations de l'impact des programmes sectoriels, et des programmes de formation en général, s'inscrivent dans ce contexte historique.

À cet égard, l'histoire pourrait servir de guide. Des programmes sectoriels permettant d'acquérir de nouvelles compétences et certifications en moins d'un an, conçus par des organismes attentifs aux besoins locaux en matière de compétences, pourraient changer la vie de nombreux travailleurs.

Il y a lieu d'espérer que les performances passées sous-estiment les performances futures. Pour les personnes ayant un bon parcours professionnel, les programmes sectoriels pourraient être moins sélectifs. De plus, comme expliqué au paragraphe 7.3, l'impact moyen d'un programme sectoriel par participant (LATE) est deux fois supérieur à son impact par personne à qui une place est proposée dans une expérimentation (ITT). Si la demande de formation augmente fortement par rapport à l'offre, davantage de personnes pourraient accepter les places proposées, ce qui accroîtrait l'impact moyen par rapport aux ITT obtenus lors des expérimentations.

8.2 Les programmes de courte durée ne correspondent pas aux besoins des travailleurs plus qualifiés.

Le déplacement massif de comptables, d'avocats, de programmeurs et d'autres professions libérales serait sans précédent, tant en termes de classes sociales concernées que, peut-être, en termes de rapidité.

Les programmes de courte durée ne répondraient généralement pas à leurs besoins, pour plusieurs raisons :

1. La formation subventionnée aide les personnes suffisamment motivées et autonomes pour maîtriser
Il ne s'agit pas d'offrir une opportunité, mais pas au point qu'ils n'en aient pas besoin. Proposer une aide particulière à quelqu'un qui n'en a pas besoin n'a aucun impact.
2. La reconversion professionnelle en profondeur prend des années. Humlum, Munch et Plato (2025) suivent des travailleurs danois victimes d'accidents du travail graves, les obligeant à quitter leur métier qualifié. L'étude exploite le fait que certaines certifications professionnelles au Danemark permettent d'accéder à l'enseignement supérieur, contrairement à d'autres. Un groupe de travailleurs blessés ayant cette possibilité l'a saisie et a finalement vu son salaire augmenter de 25 %. Ils y sont parvenus après quatre années d'études supérieures classiques.

3. Les longues périodes de formation coûtent non seulement beaucoup plus cher, mais elles rompent également le modèle de programme sectoriel.

Si un programme sectoriel détecte une demande régionale croissante de techniciens débutants dans les centres de données, il peut être en mesure de former davantage de diplômés qualifiés pour ce poste en 6 à 12 mois.

Mais si la demande de compétences commence à évoluer rapidement et de manière imprévisible — s'il est difficile de savoir quels seront les métiers en croissance dans quatre ans —, il sera difficile pour tout organisme de formation d'adapter un programme de formation sur plusieurs années aux besoins des employeurs.

8.3 L'approche d'assistance à l'ajustement commercial (TAA) mérite d'être envisagée pour les personnes qui sont manifestement licenciées en raison de l'IA.

Les données probantes sur l'efficacité du TAA sont limitées. Il n'en demeure pas moins que le temps constitue une ressource essentielle pour s'adapter aux perturbations. Gagner du temps, c'est offrir aux individus un filet de sécurité, afin que la perte d'un emploi, voire d'une profession, n'entraîne pas une catastrophe encore plus grave. Aux États-Unis, le programme TAA prolonge l'assurance chômage de six mois à trois ans, tout en proposant des aides à la reconversion professionnelle.

Bien entendu, l'approche conditionnelle du TAA ne fonctionne que pour les personnes pouvant prouver avoir perdu leur emploi à cause d'un facteur de rupture tel que l'IA. Lorsque le marché du travail se transforme non pas par des licenciements, mais par le refus des entreprises d'embaucher pour certains postes suite au départ de leurs employés ou à leur croissance, cette aide conditionnelle ne sera pas efficace.

8.4 Les bailleurs de fonds publics et privés devraient soutenir activement les démonstrations, les évaluations et la mise à l'échelle des programmes sectoriels.

Le modèle de programme sectoriel, étayé par de solides preuves, a émergé d'un processus lent et empirique. Compte tenu de l'urgence potentielle de la situation et du potentiel des programmes sectoriels, les bailleurs de fonds devraient investir davantage dans l'identification des composantes essentielles du modèle et dans la recherche des meilleures façons de le reproduire dans d'autres régions et pour un plus grand nombre de travailleurs. Une option envisageable serait un exercice de simulation rapide visant à déployer à grande échelle l'un de ces programmes auprès d'une population ciblée.

9 Discussion

Au fil des décennies, la composition professionnelle aux États-Unis a connu des transformations spectaculaires (Autor, 2015). De nombreux emplois ont disparu. Ces mutations ne sont pas rares et ne s'accompagnent pas toujours de graves difficultés économiques. Cavounidis et al. (2023) montrent, par exemple, que les salaires des conducteurs de chariots hippomobiles ont augmenté malgré le déclin de l'emploi dans ce secteur. En revanche, le déploiement des camions a été progressif. L'une des principales préoccupations économiques liées à l'IA réside dans la combinaison de ses immenses capacités et de son déploiement rapide.

Il est difficile de trouver à ce jour des preuves d'un impact significatif de l'IA sur l'emploi. Cependant, une automatisation rapide n'est pas à exclure. Les ascensoristes, les projectionnistes et les standardistes téléphoniques ont connu des périodes de chômage de courte durée.

Bien que de nombreux observateurs ne prévoient pas de hausse du chômage à court terme, l'incertitude est suffisamment grande pour justifier une analyse approfondie des options politiques. Ce rapport examine les connaissances actuelles sur l'une des options les plus courantes : la reconversion professionnelle.

Nous constatons qu'en moyenne, les programmes ont des impacts positifs, mais modestes. Notre ensemble actuel de programmes de formation peut constituer une dépense publique justifiée, dans la mesure où les investissements finissent par couvrir leur coût initial au cours de la vie active du bénéficiaire. Cependant, leurs impacts sont suffisamment faibles pour ne pas avoir d'incidence significative sur un taux de chômage toujours élevé. Nous entrevoyons un potentiel prometteur dans le regroupement de programmes sectoriels, bien que ceux-ci n'aient pas encore été testés à grande échelle.

La meilleure réponse serait probablement une combinaison de solutions partielles. Dans cet esprit, et compte tenu des besoins potentiellement importants, nous pensons qu'il est préférable que les gouvernements et autres bailleurs de fonds explorent et développent activement et simultanément plusieurs pistes de solution. Cela inclut la formation professionnelle, et en particulier les programmes sectoriels. Certaines de ces recherches sont d'actualité : elles fournissent des données utiles quel que soit l'impact de l'IA sur le chômage. Que l'IA bouleverse ou non le marché du travail, nous n'aurons aucun regret à avoir appris comment aider au mieux les travailleurs à s'adapter.

Citation suggérée

Roodman, D., & Massenkoff, M. (2026). Analyse des données probantes sur la reconversion professionnelle des travailleurs. The Anthropic Institut, Document de travail 2026-01.

www.anthropic.com/research/reviewing-the-evidence-on-worker-retraining-programs

BibTeX

```
@techreport{roodman2026training,  
  auteur      = {Roodman, David et Massenkoff, Maxim},  
  titre       = {Une analyse des données probantes sur la reconversion des travailleurs},  
  institution = {L'Institut anthropique},  
  type        = {Document de travail},  
  nombre      = {2026-01},  
  URL de      = {2026},  
  l'année     =  
{www.anthropic.com/research/reviewing-the-evidence-on-worker-retraining-programs}  
}
```

Références

Abner, Erin, Frederick A Schmitt, Marta S Mendiondo, Jennifer L Marcum et Richard J Kryscio. 2011. « Vitamine E et mortalité toutes causes confondues : une méta-analyse. » *Current Aging Science* 4 (2) : 158–70.

Aizer, Anna et Joseph J. Doyle. 2015. « Incarcération des mineurs, capital humain et criminalité future : données issues de juges désignés aléatoirement ». *Quarterly Journal of Economics* 130 (2) : 759-803. doi.org/10.1093/qje/qjv003.

Albrecht, James, Gerard J. Van den Berg et Susan Vroman. 2009. « Les effets agrégés du programme suédois d'élévation du savoir sur le marché du travail ». *Review of Economic Dynamics* 12 (1) : 129–46.

Allcott, Hunt. « Biais de sélection de site dans l'évaluation de programmes. » *The Quarterly journal of economics* 130, no. 3 (2015) : 1117-1165.

Anderson, Patricia, Walter Corson et Paul Decker. 1990. Le projet de démonstration de réinsertion professionnelle de l'assurance chômage du New Jersey : rapport de suivi. Mathematica Policy Research. doh.gov/sites/dohgov/files/ETA/publications/op_01-91.pdf.

Angrist, Joshua D., et Jörn-Steffen Pischke. 2010. « La révolution de la crédibilité en économie empirique : comment une meilleure conception de la recherche élimine les arnaques en économétrie. » *Journal of Economic Perspectives* 24 (2) : 3–30.

Anthropic. 2026. « Résultats du premier registre public anthropique. » anthropic.com/news/anthropic-public-record.

Ashenfelter, Orley. 1978. « Estimation de l'effet des programmes de formation sur les revenus. » *Revue d'économie et de statistique* 60 (1) : 47–57.

Ashenfelter, Orley C., et David Card. 1985. « Utilisation de la structure longitudinale des revenus pour estimer l'effet des programmes de formation. » *Revue d'économie et de statistique*.

Auletta, Ken. 1981a. « La sous-classe – Moi ». *New Yorker*. 16 novembre. [newyorker.com/magazine/1981/11/16/i-the-underclass](https://www.newyorker.com/magazine/1981/11/16/i-the-underclass)

Auletta, Ken. 1981b. « La sous-classe – II ». *New Yorker*. 23 novembre. [newyorker.com/magazine/1981/11/23/ii-the-underclass](https://www.newyorker.com/magazine/1981/11/23/ii-the-underclass)

Auletta, Ken. 1981c. « La classe inférieure – III. » *New-Yorkais*. 30 novembre. [newyorker.com/magazine/1981/11/30/iii-the-underclass](https://www.newyorker.com/magazine/1981/11/30/iii-the-underclass)

Auletta, Ken. 1999 (1982). La sous-classe. Open Road Integrated Media.

Autor, David H. 2015. « Pourquoi y a-t-il encore autant d'emplois ? L'histoire et l'avenir de l'automatisation du lieu de travail. » *Journal of Economic Perspectives* 29 (3) : 3–30.

Battistin, Erich et Enrico Rettore. 2002. « Test des effets de programme dans un modèle de régression par discontinuité avec observance imparfaite ». *Journal of the Royal Statistical Society. Série A* 165 (1) : 39–57. doi.org/10.1111/1467-985x.0asp6.

Betesh, Hannah, Deborah Kogan, Laura Aron, David Mitnick, Brandon Nicholson, Castle Sinicrope. 2015. Conception et mise en œuvre d'initiatives de parcours professionnels sectoriels : enseignements tirés de la mise en œuvre du programme de formation avancée pour le secteur manufacturier de l'Illinois (ATIM). Programme. Associés de recherche en politique sociale.

Bhatt, Monica P., Jonathan Guryan, Jens Ludwig et Anuj K. Shah. 2021. Défis liés à l'impact social. N° w28406. Bureau national de la recherche économique.

Bhuller, Manudeep, Gordon B. Dahl, Katrine V. Løken et Magne Mogstad. 2020. « Incarcération, récidive et emploi. » *The Journal of Political Economy* 128 (4) : 1269–1324. doi.org/10.1086/705330.

Blank, Dianne, Laura Heald et Cynthia Fagnoni. 2011. « Un aperçu de WIA ». Dans Douglas J. Besharov et Phoebe H. Cottingham, dir. *The Workforce Investment Act: Implementation Experiences and Evaluation Findings*. WE Upjohn Institute for Employment Research. research.upjohn.org/cgi/viewcontent.cgi?filename=1&article=1227&context=up_press&type=additional.

Bloom, Howard S. 1990. Retour au travail : tester les services de réemploi pour les travailleurs déplacés. Institut WE Upjohn pour la recherche sur l'emploi. doi.org/10.17848/9780585283319.

Bloom, Howard S., Larry L. Orr, George Cave, Stephen H. Bell et Fred Doolittle. 1993. Étude nationale JTPA : Impacts du titre II-A sur les revenus et l'emploi après 18 mois. Abt Associates. dol.gov/sites/dolgov/files/ETA/publications/1993_23.pdf.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu et Justin Sandefur. 2018. « Preuves expérimentales de l'extension des réformes éducatives au Kenya ». *Journal of Public Economics* 168 : 1–20. doi.org/10.1016/j.jpubeco.2018.08.007.

Borenstein, Michael, Larry V. Hedges, Julian PT Higgins et Hannah R. Rothstein. 2010. « Introduction de base aux modèles à effets fixes et à effets aléatoires pour la méta-analyse ». *Research Synthesis Methods* 1 (2) : 97–111. doi.org/10.1002/jrsm.12.

Bradley, David H. 2013. La loi sur l'investissement dans la main-d'œuvre et le système de guichet unique. Service de recherche du Congrès.
www2.law.umaryland.edu/marshall/crsreports/crsdocuments/R41135_04052013.pdf.

Brenninkmeijer, Veerle et Roland WB Blonk. 2012. « L'efficacité du programme JOBS auprès des chômeurs de longue durée : une expérience randomisée aux Pays-Bas. » *Health Promotion International* 27 (2) : 220-229. doi.org/10.1093/heapro/dar033.

Burghardt, John, Anu Rangarajan, Anne Gordon et Ellen Kisker. 1992. Évaluation du projet pilote auprès des femmes mères célibataires issues de minorités : Volume I : Rapport de synthèse. Mathematica Policy Research.

web.archive.org/web/20200530032506if_/https://www.mathematica.org/-/media/publications/pdfs/family_support/mfsp_vol1.pdf.

Caplan, Bryan. 2018. Plaidoyer contre l'éducation : pourquoi le système éducatif est une perte de temps et d'argent. Princeton University Press.

Card, David, Jochen Kluge et Andrea Weber. 2010. « Évaluations des politiques actives du marché du travail : une méta-analyse ». *Economic Journal* 120 (548) : F452-77. doi.org/10.1111/j.1468-0297.2010.02387.x.

Card, David, Jochen Kluge et Andrea Weber. 2018. « Qu'est-ce qui fonctionne ? Une méta-analyse des évaluations récentes des programmes actifs du marché du travail. » *Journal of the European Economic Association* 16 (3) : 894-931. doi.org/10.1093/jeaa/jvx028.

Cave, George, Hans Bos, Fred Doolittle, Cyril Toussaint. 1993. JOBSTART : Rapport final sur un programme pour les décrocheurs scolaires. MDRP. mdrc.org/sites/default/files/full_416.pdf.

Cavounidis, Costas, Qingyuan Chai, Kevin Lang et Raghav Malhotra. 2023. Rentes d'obsolescence : Teamsters, Truckers et innovations imminentes. Document de travail w31743. Bureau national de la recherche économique.

Cederlöf, Jonas, Martin Söderström et Johan Vikström. 2025. « Le rôle des conseillers en emploi : recherche d'emploi, qualité de l'emploi et déterminants de la valeur ajoutée ». *Journal of the European Economic Association*. doi.org/10.1093/jeaa/jvaf055.

Cheung, Maria, Johan Egebark, Anders Forslund, Lisa Laun, Magnus Rödin et Johan Vikström. 2025. « L'aide à la recherche d'emploi réduit-elle le chômage ? Preuves sur les effets et les mécanismes de déplacement. » *Journal de l'économie du travail*. 43 (1) : 47-81. [est ce que je.org/10.1086/726384](https://est.ce.queje.org/10.1086/726384).

Christensen, Garret et Edward Miguel. 2018. « Transparence, reproductibilité et crédibilité de la recherche en économie ». *Journal of Economic Literature* 56 (3) : 920-980. doi.org/10.1257/jel.20171350.

Collins, Benjamin. 2014. « Aide à l'ajustement commercial pour les travailleurs ». Service de recherche du Congrès.

web.archive.org/web/20150603092058if_/https://greenbook.waysandmeans.house.gov/sites/greenbook.waysandmeans.house.gov/files/R42012_gb_0.pdf.

Conway, Maureen. 2014. « Brève histoire des stratégies sectorielles ». Dans Maureen Conway et Robert P. Giloth (dir.), Connecting People to Work: Workforce Intermediaries and Sector Strategies. The American Assembly. aspeninstitute.org/publications/connecting-people-work.

Conway, Maureen et Robert P. Giloth, 2014. « Introduction ». Dans Maureen Conway et Robert P. Giloth (dir.), Connecting People to Work: Workforce Intermediaries and Sector Strategies. The American Assembly. aspeninstitute.org/publications/connecting-people-work.

Copson, Elizabeth, Karin Martinson et Karen Gardiner. 2014. Profil du programme « Parcours de carrière et d'éducation » : Programme Carreras en Salud de l'Instituto del Progreso Latino. Bureau de la planification, de la recherche et de l'évaluation, ACF, HHS. acf.gov/sites/default/files/documents/opre/pace_final_report_instituto_9_2_2015_b508_2.pdf.

Copson, Elizabeth, Karin Martinson, Valerie Benson, Michael Benson, Annalisa Mastri, Karen Needels et Julie Williams. 2016. Évaluation de l'impact des emplois verts et des soins de santé : résultats de l'étude de mise en œuvre de quatre programmes de formation pour les travailleurs sans emploi et défavorisés. Abt Associates. abtgloba.com/sites/default/files/2024-09/f661309f-e287-4705-b1d8-08392cd43ce2.pdf.

Corson, Walter, Paul T. Decker, Shari Miller Dunstan et Anne R. Gordon. 1989. Le projet de démonstration de réinsertion professionnelle de l'assurance chômage du New Jersey : rapport d'évaluation final. Mathematica Policy Research. web.archive.org/web/20041025070201/https://workforcesecurity.doleta.gov/dmstree/op/op89/op_03-89.pdf.

Couch, Kenneth A. 1992. « Nouvelles preuves sur les effets à long terme des programmes de formation à l'emploi. » Journal of Labor Economics 10 (4) : 380–88. doi.org/10.1086/298292.

Conseil des conseillers économiques (CEA). 2019. Programmes gouvernementaux d'emploi et de formation : évaluation des données probantes sur leur rendement.

Courty, Pascal et Gerald Marschke. 2011. « Le système d'incitation JTPA : mise en œuvre de la mesure de la performance et du financement ». Dans James J. Heckman, Carolyn J. Heinrich, Pascal Courty, Gerald Marschke et Jeffrey Smith (dir.), La performance des normes de performance. Institut WE Upjohn pour la recherche sur l'emploi. doi.org/10.17848/9780880993982.ch4.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot et Philippe Zamora. 2013. « Les politiques du marché du travail ont-elles des effets de déplacement ? Résultats d'une expérience randomisée en grappes. » *The Quarterly Journal of Economics* 128 (2) : 531-580. doi.org/10.1093/qje/qjt001.

Crépon, B., & Van Den Berg, GJ (2016). Politiques actives du marché du travail. *Annual Review of Economics*, 8(1), 521-546.

Decker, Paul T., et Craig V. Thornton. 1995. « Les effets à long terme des services d'emploi transitoires ». *Bulletin de la sécurité sociale*, 58(4), 71–81. ssa.gov/policy/docs/ssb/v58n4/v58n4p71.pdf.

Dedrick, Fred, 2014. « Partenariats industriels : théorie et pratique ». Dans Maureen Conway et Robert P. Giloth (dir.), *Connecting People to Work: Workforce Intermediaries and Sector Strategies*. The American Assembly. aspeninstitute.org/publications/connecting-people-work.

Dehejia, Rajeev H., et Sadek Wahba. 1999. « Effets causaux dans les études non expérimentales : réévaluation de l'évaluation des programmes de formation ». *Journal of the American Statistical Association* 94 (448) : 1053-1062. doi.org/10.1080/01621459.1999.10473858.

Dehejia, Rajeev H., et Sadek Wahba. 2002. « Méthodes d'appariement par score de propension pour les études causales non expérimentales ». *Review of Economics and Statistics* 84 (1) : 151–161. doi.org/10.1162/003465302317331982.

Doolittle, Fred et Linda Traeger. 1990. Mise en œuvre de l'étude nationale JTPA. MDRC. files.eric.ed.gov/fulltext/ED320010.pdf.

Dorn, David. 2009. « Essais sur l'inégalité, l'interaction spatiale et la demande de compétences. » Dissertation. Université de Saint-Gall n° 3613. Voir ddorn.net/data.htm#Local%20Labor%20Market%20Geography.

Dutta-Gupta, Indi, Kali Grant, Matthew Eckel et Peter Edelman. 2016. Leçons tirées de 40 ans de programmes d'emploi subventionné : cadre d'analyse, examen des modèles et recommandations pour aider les travailleurs défavorisés. Centre sur la pauvreté et les inégalités. Faculté de droit de Georgetown. georgetownpoverty.org/wp-content/uploads/2016/07/GCPI-Subsidized-Employment-Paper-20160413.pdf.

ETA. 2025. « Programmes d'apprentissage ; réduction du fardeau réglementaire ». 90 Fed. Reg. 28947 (proposition du 2 juillet 2025). federalregister.gov/d/2025-12317.

Farrell, Mary et Karin Martinson. 2017. Programme « Pont vers l'emploi dans le secteur de la santé » du comté de San Diego : rapport de mise en œuvre et d'impact initial. Abt Associates. acf.gov/sites/default/files/documents/opre/bridge_to_employment_implementation_and_early_impact_report_final_pdf.pdf.

Fein, David. 2012. Les parcours professionnels comme cadre de conception et d'évaluation de programmes. Rapport OPRE 2012-30. Bureau de la planification, de la recherche et de l'évaluation, Administration pour les enfants et les familles, Département américain de la santé et des services sociaux. [foc-network.org/files/Career-Pathways-as-a-Framework-by-David-Fein.pdf](https://www.foc-network.org/files/Career-Pathways-as-a-Framework-by-David-Fein.pdf).

Fein, David et Samuel Dastrup. 2022. Des avantages durables : impact à long terme et conclusions sur le rapport coût-bénéfice du programme Year Up. Abt Associates. [acf.gov/sites/default/files/documents/opre/year%20up%20long-term%20impact%20report_apr2022.pdf](https://www.acf.gov/sites/default/files/documents/opre/year%20up%20long-term%20impact%20report_apr2022.pdf).

Fortson, Kenneth, Dana Rotz, Paul Burkander, Annalisa Matri, Peter Schochet, Linda Rosenberg et Sheena McConnell. 2017. « Fournir des services publics d'emploi aux demandeurs d'emploi : résultats d'une étude d'impact sur 30 mois des programmes WIA pour adultes et travailleurs déplacés ». Mathematica Policy Research. [web.archive.org/web/20190919051434/https://www.dol.gov/sites/dolgov/files/OASP/legacy/files/WIA-30mo-main-rpt.pdf](https://www.web.archive.org/web/20190919051434/https://www.dol.gov/sites/dolgov/files/OASP/legacy/files/WIA-30mo-main-rpt.pdf).

Fraker, Thomas et Rebecca Maynard. 1984. Évaluation de méthodologies alternatives de groupes de comparaison pour l'évaluation des programmes d'emploi et de formation : rapport final. Recherche en politique Mathematica. n2t.net/ark:/13960/s2hhz59kbqm.

Gardiner, Karen et Amanda Grittnier. 2022. Programme Carreras en Salud de l'Instituto del Progreso Latino : rapport d'impact sur six ans. Abt Associés. [acf.gov/sites/default/files/documents/opre/opre-carreras%206-year%20impact%20report%20january%202022.pdf](https://www.acf.gov/sites/default/files/documents/opre/opre-carreras%206-year%20impact%20report%20january%202022.pdf).

Gardiner, Karen et Randall Juras. 2019. Pathways for Advancing Careers and Education (PACE) : Cross-Program Implementation and Impact Study Findings. Abt Associates. [acf.gov/sites/default/files/documents/opre/pace_cross_program_implementation_and_impact_study_findings_final.pdf](https://www.acf.gov/sites/default/files/documents/opre/pace_cross_program_implementation_and_impact_study_findings_final.pdf).

Gautier, Pieter, Paul Muller, Bas van der Klaauw, Michael Rosholm et Michael Svarer. 2018. "Estimation des effets d'équilibre de l'aide à la recherche d'emploi." Journal of Labor Economics 36 (4): 1073–1125. doi.org/10.1086/697513.

Bureau général de la comptabilité (GAO). 1996. Loi sur le partenariat pour la formation professionnelle : revenus à long terme et résultats en matière d'emploi. [gao.gov/assets/hehs-96-40.pdf](https://www.gao.gov/assets/hehs-96-40.pdf).

Bureau de la comptabilité générale (GAO). 2019. Programmes d'emploi et de formation : le ministère du Travail devrait évaluer les efforts déployés pour coordonner les services entre les programmes. [gao.gov/assets/d19200.pdf](https://www.gao.gov/assets/d19200.pdf).

Giloth, Robert. 2010. « Les programmes de modélisation par réplication : une attraction fatale ? » Dans Robert Giloth et Colin Austin (dir.), *Mistakes to Success: Learning and Adapting When Things Go Wrong*. iUniverse.

Ginzberg, Eli, Richard Nathan et Robert Solow. 1984. « Les leçons tirées du projet pilote de travail accompagné ». Dans Robinson G. Hollister, Jr., Peter Kemper et Rebecca Maynard (dir.), *Le projet pilote national de travail accompagné*. Presses universitaires du Wisconsin.

Gordon, Brett R., Robert Moakler et Florian Zettelmeyer. 2023. « Suffisamment proche ? Une exploration à grande échelle des approches non expérimentales de la mesure publicitaire. » *Marketing Science* 42 (4) : 768–93. doi.org/10.1287/mksc.2022.1413.

Green, Donald P., et Daniel Winik. 2010. « Utilisation de l'attribution aléatoire des juges pour estimer les effets de l'incarcération et de la probation sur la récidive chez les délinquants liés à la drogue. » *Criminology* 48(2) : 357–87. doi.org/10.1111/j.1745-9125.2010.00189.x.

Greenberg, David et Mark Shroder. 2004. *Le recueil des expériences sociales*. 3 Urban Institute Press. rd éd. Le

Greenberg, David H., Charles Michalopoulos et Philip K. Robins. 2006. « Les évaluations expérimentales et non expérimentales donnent-elles des réponses différentes quant à l'efficacité des programmes de formation financés par le gouvernement ? » *Journal of Policy Analysis and Management* 25 (3) : 523-52. doi.org/10.1002/pam.20190.

Griffen, Andrew S., et Petra E. Todd. 2017. « Évaluation de la performance des estimateurs non expérimentaux pour l'évaluation du programme Head Start ». *Journal of Labor Economics* 35 (S1) : S7–63. doi.org/10.1086/691726.

Grosz, Michel. 2020. « Les rendements d'un programme de collège communautaire de grande envergure : données issues des loteries d'admission. » *American Economic Journal : Economic Policy* 12 (1) : 226–53. doi.org/10.1257/pol.20170506.

Haelermans, Carla et Lex Borghans. 2012. « Effets salariaux de la formation en cours d'emploi : une méta-analyse ». *British Journal of Industrial Relations* 50 (3) : 502-528. doi.org/10.1111/j.1467-8543.2012.00890.x.

Hanushek, Eric A., Guido Schwerdt, Ludger Woessmann et Lei Zhang. 2017. « Formation générale, formation professionnelle et résultats sur le marché du travail tout au long de la vie. » *Journal of Human Resources* 52 (1) : 48–87. doi.org/10.3368/jhr.52.1.0415-7074r.

Heckman, James, Neil Hohmann, Jeffrey Smith et Michael Khoo. 2000. « Substitution et biais d'abandon dans les expériences sociales : une étude d'une expérience sociale influente ». *Quarterly Journal of Economics* 115 (2) : 651-94. doi.org/10.1162/003355300554764.

Heckman, James J., et V. Joseph Hotz. 1989. « Choisir parmi différentes méthodes non expérimentales pour estimer l'impact des programmes sociaux : le cas de la formation de la main-d'œuvre. »
Journal of the American Statistical Association 84 (408): 862. doi.org/10.2307/2290059.

Heckman, James J., V. Joseph Hotz et Marcelo Dabos. 1987. « Avons-nous besoin de données expérimentales pour évaluer l'impact de la formation de la main-d'œuvre sur les revenus ? » Evaluation Review 11 (4) : 395–427. doi.org/10.1177/0193841x8701100402.

Heckman, James J., Robert J. LaLonde et Jeffrey A. Smith. 1999. « L'économie et l'économétrie des programmes actifs du marché du travail ». Dans Orley Ashenfelter et David Card, éd. Manuel d'économie du travail 3. Elsevier Science.

Heckman, James J., et Jeffrey Smith. 2000. « La sensibilité des estimations d'impact expérimentales (données issues de l'étude nationale JTPA) ». Dans David G. Blanchflower et Richard B. Freeman (dir.), Emploi des jeunes et chômage dans les pays avancés. Presses de l'Université de Chicago. nber.org/system/files/chapters/c6810/c6810.pdf.

Heinrich, Carolyn J. 2023. « Les services de développement de la main-d'œuvre aux États-Unis ». Dans Douglas J. Besharov et Douglas M. Call (dir.), Travail et filet de sécurité sociale. Oxford University Press. doi.org/10.1093/oso/9780190241599.003.0010.

Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan et Harold A. Pollack. 2017. « Thinking, fast and slow? Some field experiments to reduce crime and dropoutout in Chicago ». Quarterly Journal of Economics 132 (1) : 1–54. doi.org/10.1093/qje/qjw033.

Hendra, Richard, David H. Greenberg, Gayle Hamilton, Ari Oppenheim, Alexandra Pennington, Kelsey Schaberg et Betsy L. Tessler. 2016. Des données encourageantes sur une stratégie de promotion sectorielle : impacts sur deux ans du projet de démonstration WorkAdvance. MDRC. mdrc.org/sites/default/files/2016_Workadvance_Final_Web.pdf.

Hendra, Richard et Gayle Hamilton. 2015. « Améliorer l'efficacité des programmes d'éducation et de formation pour les personnes à faible revenu : tirer des connaissances de trois décennies d'expérimentations rigoureuses. » Dans Carl Van Horn, Tammy Edwards et Todd Greene, éd. Transformer les politiques américaines de développement de la main-d'œuvre pour le XXI^e siècle. Institut WE Upjohn de recherche sur l'emploi. cew.georgetown.edu/wp-content/uploads/Transforming-Workforce-Development-Policies-for-the-21st-Century-.pdf.

Hendra, Richard, James A. Riccio, Richard Dorsett, David H. Greenberg, Genevieve Knight, Joan Phillips, Philip K. Robins, Sandra Vegeris et Johanna Walter. 2011. Rompre le cycle des bas salaires et des absences de rémunération : résultats finaux du projet pilote britannique sur le maintien et la promotion de l'emploi (ERA). Ministère du Travail et des Pensions. mdrc.org/sites/default/files/breaking-low-pay-no-pay-cycle_fr.pdf.

Hendren, Nathaniel et Ben Sprung-Keyser. 2020. « Une analyse unifiée du bien-être des politiques gouvernementales ». *Quarterly Journal of Economics* 135 (3) : 1209–1318. doi.org/10.1093/qje/qjaa006.

Humlum, Anders, Jakob R. Munch et Pernille Plato. 2025. « Changer de cap : l'investissement en capital humain après une perte de capacité ». *American Economic Review* 115 (5) : 1520-1554. doi.org/10.1257/aer.20231067.

Humlum, Anders, Jakob R. Munch et Mette Rasmussen. 2025. « Qu'est-ce qui fonctionne pour les chômeurs ? Données issues d'affectations quasi aléatoires de travailleurs sociaux. » nber.org/papers/w33807.

Hyman, Benjamin G. 2018. « La main-d'œuvre déplacée peut-elle être reconvertie ? Preuves tirées d'une affectation quasi aléatoire à une aide à l'ajustement commercial. » Document de travail.

Imbens, G., et Angrist, J. 1994. « Identification et estimation des effets moyens locaux du traitement ». *Econometrica* 62 (2) : 467–75. doi.org/10.2307/2951620.

Imbens, Guido W., et Yiqing Xu. 2025. « Comparaison des méthodes expérimentales et non expérimentales : quelles leçons avons-nous apprises quatre décennies après LaLonde (1986) ? » *Journal of Economic Perspectives* 39 (4) : 173–201. doi.org/10.1257/jep.20251440.

Groupe consultatif de recherche sur l'enquête longitudinale sur la formation professionnelle (JTLSRAP). 1985. *Recommandations du comité consultatif de recherche sur l'enquête longitudinale sur la formation professionnelle*. Département du Travail des États-Unis. n2t.net/ark:/13960/s2fkijmsc8z.

Juras, Randall, Karen Gardiner, Laura Peck et Larry Buron. 2022. *Résumé et conclusions du suivi à long terme de dix évaluations de formation professionnelle PACE et HPOG 1.0 : rapport intersites sur six ans*. Département de la Santé et des Services sociaux des États-Unis. acf.gov/sites/default/files/documents/opre/hpog%20pace%206-yr%20cross%20site%20report_august%202022_0.pdf.

Karahan, Fatih et Serdar Ozkan. 2013. « Sur la persistance des chocs de revenus au cours du cycle de vie : données probantes, théorie et implications ». *Review of Economic Dynamics* 16 (3) : 452–76. doi.org/10.1016/j.red.2012.08.001.

Karger, Ezra, Otto Kuusela, Jason Abaluck, Kevin A. Bryan, Basil Halperin, Todd R. Jones, Connacher Murphy et al. 2026. *Prévision des effets économiques de l'IA*. N° w35046. National Bureau of Economic Research, Document de travail.

Katz, Lawrence F., Jonathan Roth, Richard Hendra et Kelsey Schaberg. 2022. « Pourquoi les programmes d'emploi sectoriels fonctionnent-ils ? Leçons tirées de WorkAdvance. » *Journal of Labor Economics* 40 (S1) : S249–91. doi.org/10.1086/717932.

Kerachsky, Stuart, Craig Thornton, Anne Bloomenthal, Rebecca Maynard et Susan Stephens.

1985. Impacts de l'emploi transitoire pour les jeunes adultes ayant une déficience intellectuelle — Résultats du projet pilote STETS. MDRC. mdrc.org/sites/default/files/impacts_of_transitional_employment_fr.pdf.

King, Christopher. 2014. « Main-d'œuvre sectorielle et stratégies connexes : ce que nous savons... et ce que nous devons savoir ». Dans Maureen Conway et Robert P. Giloth (dir.), Connecting People to Work: Workforce Intermediaries and Sector Strategies. The American Assembly. aspeninstitute.org/wp-content/uploads/2014/06/Final-Version-Connecting-People-to-Work.pdf.

King, Christopher T., et Heath J. Prince. 2015. « Faire passer les programmes sectoriels et de parcours professionnels de la promesse à l'échelle. » Dans Carl Van Horn, Tammy Edwards et Todd Greene, éd. Transformer les politiques américaines de développement de la main-d'œuvre pour le XXI^e siècle. Institut WE Upjohn de recherche sur l'emploi. cew.georgetown.edu/wp-content/uploads/Transforming-Workforce-Development-Policies-for-the-21st-Century-.pdf.

Kline, Patrick et Christopher R. Walters. 2016. « Évaluation des programmes publics avec des substituts proches : le cas de Head Start. » Quarterly Journal of Economics 131 (4) : 1795–1848. doi.org/10.1093/qje/qjw027.

Kling, Jeffrey R. 2006. « Durée d'incarcération, emploi et revenus ». American Economic Review 96 (3) : 863–76. doi.org/10.1257/aer.96.3.863.

Krugman, Paul. 1994. « L'Europe sans emploi, l'Amérique sans le sou ? » Foreign Policy 95 : 19–34. doi.org/10.2307/1149421.

LaLonde, Robert J. 1986. « Évaluation des évaluations économétriques des programmes de formation à l'aide de données expérimentales. » American Economic Review 76 (4) : 604–20. doi.org/10.2307/1806062.

LaLonde, Robert et Rebecca Maynard. 1987. « Quelle est la précision des évaluations des programmes d'emploi et de formation ? Données issues d'une expérience sur le terrain. » Evaluation Review 11 (4) : 428–51. doi.org/10.1177/0193841x8701100403.

Lee, I-Min, Nancy R. Cook, J. Michael Gaziano, David Gordon, Paul M. Ridker, JoAnn E.

Manson, Charles H. Hennekens et Julie E. Buring. 2005. « La vitamine E dans la prévention primaire des maladies cardiovasculaires et du cancer : l'étude Women's Health Study : un essai contrôlé randomisé ». Journal of the American Medical Association 294 (1) : 56–65. doi.org/10.1001/jama.294.1.56.

Leuven, Edwin et Hessel Oosterbeek. 2004. « Évaluation de l'effet des déductions fiscales sur la formation ». Journal of Labor Economics 22 (2) : 461–88. doi.org/10.1086/381257.

Loeffler, Charles E. 2013. « L'emprisonnement modifie-t-il le parcours de vie ? Preuves sur la criminalité et l'emploi à partir d'une expérience naturelle. » *Criminology* 51(1) : 137–66. doi.org/10.1111/1745-9125.12000.

Maguire, Sheila, Joshua Freely, Carol Clymer, Maureen Conway et Deena Schwartz. 2010. À l'écoute des marchés du travail locaux : conclusions de l'étude d'impact sectoriel sur l'emploi. Partenariats public/privé. ppv.issuelab.org/resources/5101/5101.pdf.

Mangat, Ravinder. 2007. Aperçus sectoriels : Profil des initiatives sectorielles, 2007. Centre national de développement économique et de droit. policyarchive.org/download/8785.

Manning, Sam et Tomás Aguirre. 2026. « Dans quelle mesure les travailleurs américains sont-ils adaptables aux suppressions d'emplois induites par l'IA ? » Dans Ajay K. Agrawal, Erik Brynjolfsson et Anton Korinek (dir.), * The Economics of Transformative AI*. University of Chicago Press. nber.org/books-and-chapters/economics-transformative-ai.

Manpower Demonstration Research Corporation. 1980. Résumé et conclusions du projet national de démonstration de travail soutenu. mdrc.org/sites/default/files/full_249.pdf.

Manpower Demonstration Research Corporation. 2009. Étude nationale d'évaluation du travail soutenu, 1975-1979 : Fichiers à usage public. ICPSR07865-v2. Consortium interuniversitaire pour la recherche politique et sociale [distributeur]. doi.org/10.3886/ICPSR07865.v2.

Martin, SE, S. Annan et B. Forst. 1993. « Les effets dissuasifs particuliers d'une sanction d'emprisonnement sur les conducteurs ivres pour la première fois : une étude quasi-expérimentale ». *Accident Analysis and Prevention* 25(5) : 561–68. [doi.org/10.1016/0001-4575\(93\)90008-k](https://doi.org/10.1016/0001-4575(93)90008-k).

Martinson, Karin, Julie Williams et al. 2016. Évaluation de l'impact des emplois verts et des soins de santé : résultats de l'étude d'impact de quatre programmes de formation pour les travailleurs sans emploi et défavorisés. Mathematica Policy Research. dol.gov/sites/dolgov/files/ETA/publications/ETAOP-2017-07%20Findings%20from%20the%20Impact%20Study.pdf.

Massenkoff, Maxim et Peter McCrory. 2026. « Impacts de l'IA sur le marché du travail : une nouvelle mesure et des résultats préliminaires ». *Anthropic*. cdn.sanity.io/files/4zrzovbb/website/2b5bbaf2c1eb81dbf6e6fb813c1a24e35a64d376.pdf.

Mathematica Policy Research, Inc. (MPR) et Social and Scientific Systems, Inc. (SSS). 2004. Étude nationale d'évaluation du travail soutenu, 1975-1979 : Guide de l'utilisateur. doi.org/10.3886/ICPSR07865.v2.

Maynard, Rebecca, Randall Brown et Jennifer Schore. 1979. Le projet pilote de travail accompagné : effets au cours des 18 premiers mois suivant l'inscription. Mathematica Policy Research. ojp.gov/pdffiles1/Digitization/59202NCJRS.pdf.

McConnell, Sheena et Steven Glazerman. 2001. Étude nationale sur le Job Corps : avantages et coûts du Job Corps. Mathematica Policy Research. dol.gov/sites/dolgov/files/ETA/publications/01-jcbenefit.pdf.

Miller, Cynthia, Johannes M. Bos, Kristin E. Porter, Fannie M. Tseng et Yasuyo Abe. 2005. Le défi de reproduire le succès dans un monde en mutation : rapport final sur les sites de réplication du Centre de formation à l'emploi. MDRC. dol.gov/sites/dolgov/files/ETA/research/pdfs/cet.pdf.

Miller, Sarah, Elizabeth Rhodes, Alexander W. Bartik, David E. Broockman, Patrick K. Krause et Eva Vivalt. 2024. « Le revenu a-t-il une incidence sur la santé ? Résultats d'un essai contrôlé randomisé sur un revenu garanti. » Document de travail w32711. Bureau national de la recherche économique. doi.org/10.3386/w32711.

Modicamore, Dominic, Yvette Lamb, Jeffrey Taylor, Ama Takyi-Laryea, Kathy Karageorge et Enzo Ferroggiaro. 2017. Accélérer l'accès à l'emploi : Volume I : Rapport d'évaluation final. ICF. files.eric.ed.gov/fulltext/ED618503.pdf.

Mountjoy, Jack. 2026. « Rendements marginaux des universités publiques ». Quarterly Journal of Economics 141 (1) : 429–97. doi.org/10.1093/qje/qjaf055.

Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Winston Lin, George Cave et Fred Doolittle. 1994. Étude nationale JTPA : impact, avantages et coûts du titre II-A. Abt Associates. upjohn.org/sites/default/files/2019-02/jtpa_national_evaluation.zip.

Peck, Laura, Deena Schwartz, Julia Strawn, Christopher C. Weiss, et al. 2021. Une méta-analyse de 46 évaluations d'impact des parcours professionnels. Abt Associates. dol.gov/sites/dolgov/files/ETA/publications/ETAOP2022-04_A%20Meta-Analysis%20of%2046%20Career%20Pathways%20Impact%20Evaluations_final%20report.pdf.

Porter, Bruce. 1974. « La voie du chat sauvage ». New York Times Magazine. 28 avril. nyti.ms/4ILWqrO.

Raaum, Oddbjørn et Hege Torp. 2002. « La formation sur le marché du travail en Norvège : effet sur les revenus ». Labour Economics 9 (2) : 207–47. [doi.org/10.1016/S0927-5371\(02\)00041-6](https://doi.org/10.1016/S0927-5371(02)00041-6).

Rademacher, Ida, Marshall Bear et Maureen Conway. 2001. Projet QUEST : Étude de cas d'une approche sectorielle de développement de l'emploi. Aspen Institute. d3n8a8pro7vhmx.cloudfront.net/interfaitheducationfund/pages/1443/attachments/original/15555_07182/Project_Quest_-_Aspen_Study_2001.PDF.

Rangarajan, Anu, John Burghardt et Anne Gordon. 1992. Démonstration sur les mères célibataires issues de minorités : Volume II : Supplément technique à l'analyse des impacts économiques. Recherche en politiques mathématiques.

web.archive.org/web/20200529190732/https://mathematica.org/-/media/publications/pdfs/family_support/mfsp_vol2.pdf.

Richmond, AM (2026). Cadre de transition vers l'emploi à l'ère de l'IA : cartographie de l'impact à court terme de l'IA sur l'emploi. OpenAI. https://cdn.openai.com/pdf/the-ai-jobs-transition-framework_report.pdf

Roder, Anne. 2008. Cibler les industries, former les travailleurs et améliorer les perspectives : Rapport final de l'Initiative sectorielle pour l'emploi. Partenariats public-privé. economicmobilitycorp.org/wp-content/uploads/2018/01/Targeting-Industries-Training-Workers-and-Improving-Opportunities.pdf.

Roder, Anne et Mark Elliott. 2011. Un début prometteur : les premiers impacts de Year Up sur les carrières des jeunes adultes à faible revenu. Economic Mobility Corporation. clear.dol.gov/sites/default/files/OYRoder2014.pdf.

Roder, Anne et Mark Elliott. 2014. Des gains durables : l'impact continu de Year Up sur les revenus des jeunes adultes. Economic Mobility Corporation. economicmobilitycorp.org/wp-content/uploads/2018/01/A-Promising-Start.pdf.

Roder, Anne et Mark Elliott. 2021. Gains sur onze ans : l'investissement du projet QUEST continue de porter ses fruits. Economic Mobility Corporation. dol.gov/sites/dolgov/files/ETA/publications/ETAOP-2017-07%20Findings%20from%20the%20Impact%20Study.pdf.

Rolston, Howard, Elizabeth Copson et Karen Gardiner. 2017. Initiative de la vallée pour le développement et la promotion : rapport de mise en œuvre et d'impact initial. Abt Associates. abtglobal.com/sites/default/files/migrated_files/5378864c-5b8a-4a6c-a94c-9cb7084c0fd4_0.pdf.

Roodman, David. 2009. « Note sur le thème du trop grand nombre d'instruments ». Oxford Bulletin of Economics and Statistics 71 (1) : 135–58. doi.org/10.1111/j.1468-0084.2008.00542.x.

Roodman, David. 2011. « Ajustement de modèles de processus mixtes récurrents entièrement observés avec cmp. » Journal Stata. 11 (2) : 159-206. est-ce-que-je.org/10.1177/1536867X1101100202.

Roodman, David. 2014. « Les impacts économiques nationaux de l'immigration ». SSRN. doi.org/10.2139/ssrn.3635860.

Rosholm, Michael et Lars Skipper. 2009. « La formation professionnelle est-elle un fléau pour les chômeurs ? Résultats d'une expérience sociale. » Journal of Applied Econometrics 24 (2) : 338-365. doi.org/10.1002/jae.1048.

Rothstein, Jesse, Robert Santillano, Till von Wachter, Wahid Khan et Mary Yang. 2022. CAAL-Skills : Étude des programmes de formation de la main-d'œuvre en Californie. Laboratoire de politiques publiques de Californie.

capolicylab.org/wp-content/uploads/2022/09/CAAL-Skills-Study-of-Workforce-Training-Programs-in-California.pdf.

Rotz, Dana, Paul Burkander, Mary Grider, Kenneth Fortson, Linda Molinari, Elias Sanchez-Eppler et Lindsay Cattell. 2017. « Fournir des services publics d'emploi aux demandeurs d'emploi : résultats d'une étude d'impact sur 30 mois des programmes WIA pour adultes et travailleurs déplacés : supplément technique ». Mathematica Policy Research. doi.gov/sites/dolgov/files/ETA/publications/ETAOP_2018-04_2-WIA-30m-TechSupplement_508_compliant.pdf.

Rubin, Donald B. 1980. « Analyse de randomisation des données expérimentales : commentaire sur le test de randomisation de Fisher. » *Journal of the American Statistical Association* 75 (371) : 591. doi.org/10.2307/2287653.

Şahin, Ayşegül, Joseph Song, Giorgio Topa et Giovanni L. Violante. 2014. « Inadéquation du chômage ». *Revue économique américaine* 104 (11) : 3529-64. [est ce que je.org/10.1257/aer.104.11.3529](https://est.ce.queje.org/10.1257/aer.104.11.3529).

Sarna, Maureen et Julie Strawn. 2018. Synthèse de la mise en œuvre des parcours professionnels : étude de conception des parcours professionnels. Abt Associates. doi.gov/sites/dolgov/files/OASP/legacy/files/3-Career-Pathways-Implementation-Synthesis.pdf.

Schaberg, Kelsey et David H. Greenberg. 2020. Effets à long terme d'une stratégie de développement sectoriel : coûts, avantages et impacts du projet de démonstration WorkAdvance. MDRC. mdrc.org/sites/default/files/WorkAdvance_5-Year_Report-Final.pdf.

Schochet, Peter Z. 2009. « Une approche pour aborder le problème des tests multiples dans les évaluations d'impact des politiques sociales. » *Evaluation Review* 33 (6) : 539–67. doi.org/10.1177/0193841X09350590.

Schochet, Peter Z. 2021. « Effets à long terme du programme Job Corps sur le marché du travail : données issues d'une expérience représentative à l'échelle nationale. » *Journal of Policy Analysis and Management* 40 (1) : 128–57. doi.org/10.1002/pam.22233.

Schochet, Peter Z., John Burghardt et Sheena McConnell. 2008. « Le programme Job Corps est-il efficace ? » Résultats d'impact de l'étude nationale sur le Job Corps. *American Economic Review* 98 (5) : 1864–86. doi.org/10.1257/aer.98.5.1864.

Simonsohn, Uri, Leif D. Nelson et Joseph P. Simmons. 2014. « La courbe P : une clé pour le tiroir à dossiers ». *Journal of Experimental Psychology: General* 143 (2) : 534. doi.org/10.1037/a0033242.

Smedslund, Geir, Kåre Birger Hagen, Asbjørn Steiro, Torill Johme, Therese Kristine Dalsbø et Mons Georg Rud. 2006. « Programmes de travail pour les bénéficiaires de l'aide sociale », *Campbell Systematic Reviews* 2 (1) : 1-122. [est ce que je.org/10.4073/csr.2006.9](https://est.ce.queje.org/10.4073/csr.2006.9).

Smith, Jeffrey A., et Petra E. Todd. 2005. « L'appariement permet-il de surmonter la critique de LaLonde à l'égard des estimateurs non expérimentaux ? » *Journal of Econometrics* 125 : 305–53. doi.org/10.1016/j.jeconom.2004.04.011.

Solon, Gary, Steven J. Haider et Jeffrey M. Wooldridge. 2015. « Pour quoi pondérons-nous ? » *Journal of Human Resources* 50 (2): 301–16. doi.org/10.3368/jhr.50.2.301.

Sterling, Theodore D. 1959. « Décisions de publication et leurs effets possibles sur les inférences tirées des tests de signification – ou vice versa. » *Journal of the American Statistical Association* 54 (285) : 30–34. doi.org/10.1080/01621459.1959.10501497.

Straight Talk on Evidence. 2020. « Une étude récemment publiée sur le programme fédéral Job Corps affirme à tort démontrer des effets positifs à long terme. » *Arnold Ventures*. 21 juillet. straighttalkonevidence.org/2020/07/21/newly-published-study-of-federal-job-corps-program-inaccurately-claims-to-demonstrate-long-term-positive-effects.

Stromsdorfer, Ernst W. 1987. « Évaluation économique de la loi sur l'emploi et la formation complets : un aperçu des résultats récents et des progrès des méthodes d'évaluation. » *Evaluation Review* 11 (4) : 387–94. doi.org/10.1177/0193841x8701100401.

Torp, Hege, Oddbjørn Raaum, Erik Hernes et Harald Goldstein. 1993. «La première expérience norvégienne». Dans Karsten Jensen et Per Kongshoj Madsen, éd. *Mesurer les mesures du marché du travail : évaluer les effets des initiatives politiques actives sur le marché du travail*. Ministère du Travail, Danemark.

Bureau des statistiques du travail des États-Unis. « Taux de chômage – Noirs ou Afro-Américains [LNS14000006] ». Extrait de FRED, Banque fédérale de réserve de Saint-Louis. Consulté le 10 août 2026. fred.stlouisfed.org/series/LNS14000006.

Département du Travail des États-Unis, Bureau principal de l'évaluation. s.d. « Données utilisées pour la méta-analyse de 46 évaluations d'impact des parcours professionnels ». Jeu de données à usage public. Archivé le 5 janvier 2025 sur Internet Archive.

web.archive.org/web/20250105070056if_/https://www.dol.gov/sites/dolgov/files/OASP/evaluation/pdf/cp_meta_publicdataset.xlsx.

Commission du travail et des ressources humaines du Sénat américain. 1980. *Emploi des jeunes et réforme de l'aide sociale : Auditions devant la sous-commission sur l'emploi, la pauvreté et le travail migrant*. 96e Congrès, 2e session, 5-6 et 12-13 mars. ERIC ED198345. files.eric.ed.gov/fulltext/ED198345.pdf.

Vera Institute of Justice. 1976. *Profils d'emploi des anciens toxicomanes : impact du travail accompagné sur l'emploi à long terme*. vera-institute.files.svdcdn.com/production/downloads/publications/2023.pdf.

Vivalt, Eva. 2020. « Dans quelle mesure pouvons-nous généraliser à partir des évaluations d'impact ? »
Journal of the European Economic Association 18 (6) : 3045–89. doi.org/10.1093/jeaa/jvaa019.

Yusim, Aleksandra, Kelsey Schaberg, Betsy Tessler et Alexandra Ubalijoro. 2025. Effets de la formation
sectorielle après 10 ans : résultats de l'évaluation WorkAdvance. MDRC. [mdrc.org/sites/default/files/
WorkAdvance_10-Year_Report_MDRC.pdf](https://mdrc.org/sites/default/files/WorkAdvance_10-Year_Report_MDRC.pdf).

Zandniapour, Lily et Maureen Conway. 2002. Gaining Ground : The Labor Market Progress of Participants of
Sectoral Employment Development Programs. Aspen Institute. [aspeninstitute.org/wp-
content/uploads/2025/05/Gaining-Ground.pdf](https://aspeninstitute.org/wp-content/uploads/2025/05/Gaining-Ground.pdf).

Appendice

Annexe A : Autres thèmes

A.1 Historique de la législation américaine sur le travail

Aux États-Unis, l'histoire des programmes fédéraux de formation professionnelle remonte à la loi de 1962 sur le développement et la formation de la main-d'œuvre (MDTA), et se poursuit avec les programmes d'emploi du New Deal de Roosevelt. Les motivations de l'intervention fédérale dans le développement de la main-d'œuvre ont toujours été multiples : renforcer la compétitivité économique, lutter contre les injustices raciales, réduire la pauvreté et aider les travailleurs à s'adapter à la concurrence étrangère exacerbée par l'ouverture au commerce. Comme c'est souvent le cas dans la politique américaine, les tendances à la fragmentation et à l'unification se sont opposées au fil des décennies, le Congrès créant de nombreux programmes et s'efforçant parfois d'en regrouper plusieurs sous une même égide.

Bien que l'historique législatif constitue un élément de contexte important pour la plupart des recherches analysées ici, son examen complet dépasse le cadre du présent rapport. Le tableau A.1 présente un aperçu des principales lois.

Act	Year	What it did
Manpower Development and Training Act (MDTA)	1962	Officially established the federal public employment and training system in the US
Comprehensive Employment and Training Act (CETA)	1973	Extended the WPA approach by providing work for the long-term unemployed and low-income adults, plus summer jobs for low-income youth, while ceding more administrative control to state governments.
Job Training Partnership Act (JTPA)	1982	Decentralized administration and expanded the private sector's role in delivering services; eliminated CETA's public service employment and stipends; and introduced the first national outcomes-based performance standards system for federal workforce programs.
Workforce Investment Act (WIA)	1998	Kept authority largely with states and localities, required State and Local Workforce Investment Boards led by business and other stakeholders, and mandated one-stop career centers in every local area.
Workforce Innovation and Opportunity Act (WIOA)	2014	Strengthened coordination and co-location across workforce programs, gave states more flexibility to shift funds between Adult and Dislocated Worker programs, required at least 20% of Title I youth funds to go to work-based learning, and established Pay for Performance as an eligible use of formula funds.

Tableau A.1. Historique des lois américaines sur le développement de la main-d'œuvre Remarques : Basé sur Heinrich (2023), pp. 243–44.

A.2 Le déclin de l'étude d'Ashenfelter et la révolution de la crédibilité. L'expérience des

50 dernières années a permis de dégager un consensus : les études non randomisées sont bien moins fiables que les études randomisées. Parmi les exemples les plus significatifs, on peut citer la formation professionnelle. Ashenfelter (1978) étudie l'impact de la loi sur le développement et la formation de la main-d'œuvre (Manpower Development and Training Act), principal programme fédéral de formation des années 1960. L'accès à ce programme n'étant pas randomisé, Ashenfelter n'a eu d'autre choix que de comparer les personnes ayant bénéficié de la formation à un groupe témoin n'en ayant pas bénéficié.

L'étude révèle une tendance particulière, désormais appelée « effet Ashenfelter ». Les personnes ayant intégré la formation MDTA en 1964 avaient déjà subi une baisse de leurs revenus en 1962-1963 (Ashenfelter 1978, tableau 1). Durant ces mêmes années précédant le traitement, les revenus des non-participants avaient augmenté. De toute évidence, même avant le traitement, les deux groupes présentaient une différence systématique. Et si la différence de revenus observée entre le groupe traité et le groupe témoin avant le traitement ne pouvait être attribuée à la formation, comment attribuer avec certitude une quelconque différence observée après le traitement ? Il est possible que toute hausse constatée après le traitement dans le groupe traité n'ait été qu'un rebond suite à une période difficile, c'est-à-dire un retour à la moyenne. C'est probablement l'une des raisons.

Orley Ashenfelter et le futur lauréat du prix Nobel David Card, qui étudiaient également la formation professionnelle, ont conclu que « des essais randomisés... sont nécessaires pour déterminer de manière fiable les effets du programme » (Ashenfelter et Card 1985).

Une autre raison de leur valorisation de la randomisation résidait sans doute dans une étude d'une importance capitale pour une génération, que Robert LaLonde, étudiant d'Ashenfelter, venait de terminer. LaLonde (1986) démontre que différentes méthodes raisonnables de constitution d'un groupe témoin (non randomisé) et d'analyse quantitative aboutissent à des estimations très différentes de l'impact du programme. Cela remet en question toutes les évaluations dont les groupes témoins sont constitués sans randomisation. Gordon et al. (2023) constituent une démonstration plus récente de ce même phénomène : les interventions non randomisées, même avec les techniques modernes d'apprentissage automatique permettant de générer le meilleur groupe contrefactuel possible, produisent des estimations biaisées.

En partie à cause des travaux de LaLonde (1986), une « révolution de la crédibilité » (Angrist et Pischke, 2010) a bouleversé l'économie. Ce phénomène, à la fois culturel et méthodologique, privilégie les recherches qui exploitent la randomisation, ainsi que certaines autres techniques permettant de simuler des expériences naturelles presque aussi convaincantes que les expériences réelles. La méthode de la discontinuité de régression (RDD) en est un exemple. Souvent, les programmes imposent des seuils d'admissibilité, tels qu'un revenu maximal ou un score minimal à un test. Les personnes dont les revenus de l'année précédente se situent juste de part et d'autre de ce seuil – 49 999 \$ ou 50 001 \$ respectivement – mènent probablement des vies statistiquement similaires jusqu'au jour où un groupe peut accéder à un programme et l'autre non.

La révolution de la crédibilité influence fortement ce rapport. Nous consacrons l'essentiel de notre analyse aux études randomisées. Ce choix est facilité par le fait que des expériences randomisées de formation professionnelle sont menées aux États-Unis depuis 50 ans. Nous aborderons également deux évaluations de la méthode RDD (Randomized Due Dirigation), dans l'annexe D, ainsi que deux études de « randomisation par juge », au paragraphe 5.

A.3 Effets de l'observateur et de la démonstration

L'expérimentation modifie l'objet de l'expérimentation.

Par exemple, l'« effet Hawthorne » peut se produire : les entraîneurs peuvent adapter leur entraînement car ils savent que leurs performances sont évaluées. Il existe aussi l'« effet John Henry » : les membres du groupe témoin peuvent redoubler d'efforts pour compenser leur désavantage perçu.

De même, de nombreuses évaluations portent sur des « démonstrations », mises en place expressément pour évaluer des méthodes de formation spécifiques. Ces démonstrations nécessitent généralement un financement particulier de la part d'organismes gouvernementaux ou de fondations. Ce soutien, conjugué à la volonté de rendre justice aux stratégies testées, peut se traduire par un financement plus important pour la démonstration qu'un programme en activité, et par un suivi plus rigoureux de sa fidélité au modèle. La question du financement est particulièrement cruciale pour la formation, car elle coûte plus cher que l'aide à la rédaction de CV et à la recherche d'emploi. Si une stratégie était déployée à l'échelle nationale suite à une démonstration concluante, mais sans budget adéquat, le volet formation pourrait être fortement compromis, creusant un fossé entre théorie et pratique, entre l'original et la réplique.

A.4 Externalités et effets d'équilibre général. La randomisation permet

d'isoler les impacts causaux, mais de nombreux effets restent incommensurables. Si une personne trouve un emploi grâce à une formation, elle est moins susceptible de commettre un crime, protégeant ainsi d'éventuelles victimes. Cependant, cette personne, nouvellement embauchée grâce à la formation, peut avoir pris la place d'autres candidats, les privant des mêmes avantages. Ces effets sont importants pour connaître le bénéfice net pour la société, même lorsqu'ils sont difficiles à observer et à quantifier.

Un effet externe malheureux est l'effet de déplacement ou d'équilibre général (Heckman et al. (1999 ; Crépon & Van Den Berg 2016). Si une personne du groupe expérimental est formée à la soudure et décroche un excellent emploi en usine, ne prend-elle pas simplement ce poste à quelqu'un d'autre ? Cette personne ne participerait probablement pas à l'étude, et son sort passerait donc inaperçu lors du calcul des effets du programme. Pire encore, si elle faisait partie du groupe témoin, les bénéfices de la formation seraient comptabilisés deux fois, puisque le salaire aurait augmenté dans le groupe expérimental et diminué dans le groupe témoin.

60

Un élément de preuve important concernant le déplacement de la population active provient de l'étude de Crépon et al. (2013), qui présente les résultats d'une expérience nationale menée en France. Le traitement consistait en une orientation vers un organisme de placement privé pour un accompagnement intensif pouvant aller jusqu'à un an. Dans chaque localité, la proportion de personnes devant bénéficier de ce traitement a d'abord été fixée aléatoirement à 0 %, 25 %, 50 %, 75 % ou 100 %. Ensuite, la proportion requise a été sélectionnée aléatoirement et contactée. L'étude montre que le traitement a temporairement augmenté l'emploi parmi les personnes traitées – l'effet se manifestant environ huit mois plus tard, mais s'estompant au bout de douze mois – mais que même cet avantage temporaire a été contrebalancé par une baisse de l'emploi chez les personnes non traitées, en particulier dans les régions où le marché du travail était atone. Gautier et al. (2018) et Cheung et al. (2025) ont mis en évidence des effets de déplacement de la population active similaires au Danemark et en Suède.

Des données microéconomiques sur le programme, ne concernant que les participants, auraient été encourageantes. Mais en raison de ces effets de déplacement, concluent les auteurs, « le programme semble avoir eu très peu d'avantages nets ». Ce résultat suggère que les études surestiment les bénéfices de l'aide à la recherche d'emploi, car elles ne prennent pas en compte les effets secondaires sur les personnes extérieures à l'étude.

Cela s'applique-t-il à la reconversion professionnelle ? Nous pensons que les effets de substitution sont probablement moindres pour les programmes de formation qui développent les compétences, surtout à long terme. Ces programmes accroissent la capacité productive de la société, et une économie dynamique tend à tirer parti des ressources disponibles. (Pour une version formelle de cet argument, voir Albrecht et al. (2009).) De plus, les programmes les plus efficaces que nous avons recensés dans cette étude ciblent les professions en forte demande, c'est-à-dire celles où la concurrence pour les emplois disponibles est apparemment moindre.

Pour mieux comprendre, notons que ce même problème se pose en matière d'immigration (Roodman 2014). Les immigrants prennent-ils le travail des natifs ? C'est plausible. En revanche, si le marché du travail fonctionnait selon un principe de somme nulle, la multiplication par dix de la population américaine depuis 1860 aurait entraîné un taux de chômage de 90 %. La solution à ce paradoxe réside dans la prise en compte de...

60 Dans le langage de Rubin (1980), les effets de déplacement violent l'hypothèse de valeur de traitement unitaire stable (SUTVA).

Le rôle du temps. À très court terme, le nombre d'offres d'emploi est fixe. À plus long terme, qui peut être relativement rapide, le stock de capital (usines et bâtiments) réagit aux opportunités créées par l'augmentation de l'offre de main-d'œuvre et se développe également. Par exemple, l'afflux, pendant des décennies, d'immigrants latino-américains aux États-Unis – des immigrants prêts à travailler dur pour de bas salaires – a probablement amélioré la rentabilité du secteur de la restauration. Outre la stimulation de la demande de main-d'œuvre locale pour le contact avec la clientèle, cet afflux a permis à l'économie d'intégrer efficacement les immigrants. La rapidité avec laquelle ce long terme se concrétise – la rapidité avec laquelle la disponibilité accrue de main-d'œuvre entraîne la création d'emplois – dépend des travailleurs concernés et du contexte économique.

A.5 Le placement des études randomisées n'est pas aléatoire

Bien que la randomisation soit un outil puissant pour l'inférence causale, même les études randomisées doivent être lues et interprétées avec un regard critique. Les programmes dont les évaluations randomisées nous parviennent ne le font qu'après avoir été soumis à des biais de sélection non aléatoire et d'auto-sélection (Allcott, 2015). Par conséquent, les évaluations randomisées, prises dans leur ensemble, ne sont pas strictement représentatives des programmes de formation professionnelle dans leur ensemble. Il ne s'agit pas d'une critique, mais d'une simple mise en garde quant à la généralisation des résultats.

La répartition des études randomisées est influencée par de nombreux facteurs. Elles sont souvent menées pour tester des approches nouvelles plutôt que des approches éprouvées. Les expériences sociales sont contraignantes et coûteuses.

Le personnel des programmes pourrait se montrer réticent face à un dispositif l'obligeant à passer outre son jugement professionnel et à se priver du service qui donne du sens à son travail. Si les responsables de programmes plus confiants dans l'efficacité de leur organisation sont plus enclins à participer à une expérimentation, et s'ils ont souvent raison, alors les meilleurs programmes seront davantage évalués. Les dirigeants qui détournent des fonds ne se porteront probablement pas volontaires. Les gouvernements et les fondations pourraient être plus disposés à financer le suivi à long terme de programmes ayant obtenu de bons résultats à court terme, ce qui biaiserait les données à long terme dont nous disposons.

En revanche, les programmes de formation destinés aux personnes les plus difficiles à aider et les moins influentes politiquement risquent de faire l'objet de davantage de remises en question quant à leur efficacité. Cela inverse le biais : les programmes aux missions les plus ardues bénéficieraient d'une plus grande couverture médiatique.

A.6 Signalisation vs. capital humain

Grâce notamment aux travaux de Gary Becker, les économistes ont traditionnellement considéré l'investissement dans l'éducation comme l'équivalent cognitif de l'achat d'un tracteur : un investissement initial dans le capital de compétences qui s'autofinance en multipliant la valeur du travail. L'éducation et la formation accroissent le capital humain.

Tout en reconnaissant que les étudiants en ingénierie apprennent réellement l'ingénierie et que les étudiants en médecine apprennent réellement la médecine, les sceptiques affirment que peu de ce qui passe aujourd'hui pour de l'éducation enseigne quoi que ce soit d'utile sur le plan économique. Une partie de ce qui est « enseigné » n'est en réalité pas apprise. Une grande partie de ce qui est appris est rapidement oubliée. Et une grande partie de ces connaissances — le français, l'histoire médiévale, les tables de multiplication — ne rend pas les travailleurs plus productifs.

Pourquoi, dès lors, l'obtention d'un diplôme d'études secondaires, voire d'un diplôme universitaire, semble-t-elle augmenter autant les revenus ? Dans son ouvrage **The Case Against Education**, l'économiste Bryan Caplan soutient que la principale fonction économique d'un diplôme est de signaler la valeur de son titulaire. Obtenir un diplôme exige « intelligence, conscience professionnelle et conformisme » (Caplan, 2018, p. 37). Les bons élèves font de bons employés. L'effet « peluche » constitue un argument en faveur de cette théorie : les personnes qui ont failli obtenir leur diplôme, mais qui l'abandonnent, gagnent nettement moins que celles qui l'ont obtenu. Pourtant, ces personnes, bien que n'ayant pas ce précieux diplôme, sont sans doute presque aussi compétentes. Elles ont suivi un nombre de cours presque équivalent.

Si les programmes de formation professionnelle n'aident les individus qu'en certifiant leurs capacités intellectuelles, leur persévérance et leur conformisme, cela renforce l'idée que les gains sont à somme nulle : la certification permet simplement aux individus de se démarquer dans la concurrence pour un nombre limité d'emplois. Dans cette optique, l'argent investi dans la formation est entièrement gaspillé en raison des effets indirects.

Il est pourtant bien établi que les agents économiques fonctionnent grâce à l'information, notamment sur la quantité et la qualité de leurs intrants et sur la demande pour leurs extrants, et que l'acquisition de cette information a un coût. Lorsqu'une personne finance une formation, même si celle-ci se limite à la certification de compétences préexistantes, elle génère des informations qui permettent une meilleure combinaison des facteurs de production. Dans le cadre de la formation professionnelle des personnes défavorisées, la réussite d'un cursus, qu'il s'agisse de formation en salle de classe ou en entreprise, atteste de leur aptitude à intégrer le monde du travail. Un programme de formation qui diffuse ces connaissances sur le marché du travail accroît de fait l'offre de main-d'œuvre disponible pour les employeurs.

En réalité, Caplan, critique virulent de l'éducation, soutient la formation professionnelle, une catégorie qui englobe les programmes que nous étudions. Caplan (2018, p. 303) estime que si 80 % de la valeur d'un diplôme universitaire relève de la signalisation, ce chiffre atteint 40 % pour un certificat professionnel. Les données sur lesquelles s'appuie Caplan (2018, chapitre 8, notes 9-13) fonde cette conclusion favorable ne comprennent aucune étude randomisée et, à la rigueur, une seule quasi-expérience potentiellement convaincante. Malgré cela, il est tout à fait plausible que la formation professionnelle soit productive, même dans un contexte où la signalisation scolaire joue un rôle prépondérant.

⁶¹ À tout moment

⁶¹Une partie de Hanushek et al. (2017) exploite la variation des fermetures d'usines en Autriche.

Annexe B : Figures supplémentaires

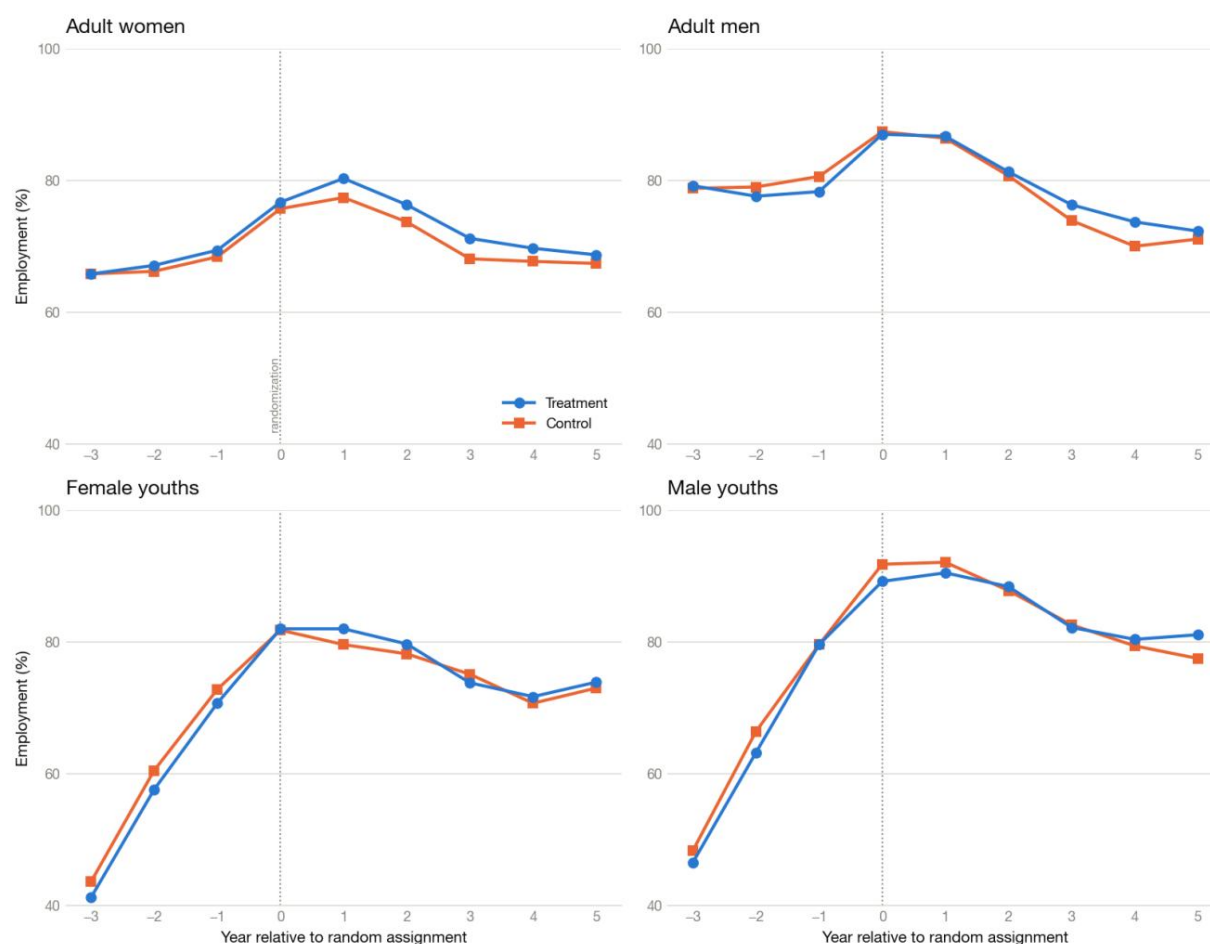


Figure B1. Salaires perçus dans l'étude nationale JTPA par année avant/depuis la randomisation Remarques : Source : GAO (1996), tableaux II.1–II.4.

Annexe C : Expériences en Europe

Les données probantes de haute qualité sur l'impact des formations sont beaucoup plus rares hors des États-Unis. En Europe, si les expériences sociales sont rarement menées, c'est peut-être parce que des prestations comme l'aide individualisée à l'emploi sont perçues comme des droits fondamentaux qui ne devraient pas être arbitrairement refusés à des fins de recherche.

⁶² Nous passons ici en revue tout ce que nous avons trouvé concernant la formation : trois efforts qui, en fin de compte, aboutissent à ne pas générer d'expériences solides et rigoureuses, et une évaluation de style américain au Royaume-Uni dans laquelle la formation était une composante secondaire.

⁶² Sur le cas du Danemark, voir Humlum, Munch et Rasmussen (2025), annexe B.

Raaum et Torp (2002), « La formation sur le marché du travail en Norvège – effet sur les revenus »

Influencés par le débat des années 1980 aux États-Unis sur la pertinence des évaluations non randomisées, deux chercheurs norvégiens ont entrepris ce qui semble être la première expérience randomisée de formation professionnelle en Europe. Eux aussi ont utilisé les résultats de cette randomisation pour évaluer les méthodes non randomisées qu'ils auraient autrement appliquées.

Raaum et Torp (2002) étudient le programme de formation professionnelle du gouvernement norvégien (LMT). Ce programme proposait des formations professionnelles de plusieurs mois dans des domaines tels que le travail de bureau, les transports et l'industrie. La demande pour les services du LMT a fortement augmenté avec la hausse du chômage, passé de 3,2 % à 5,5 % entre 1988 et 1994 (Raaum et Torp, 2002, p. 210). Le gouvernement a dû rationner les places de formation, ce qui a conduit à la mise en place d'un système de tirage au sort. L'échantillon expérimental de Raaum et Torp est composé de 770 personnes ayant postulé à la LMT en août-septembre 1991 dans 6 des 19 comtés de Norvège. La plupart des sujets ont terminé leur formation début 1992 (Raaum et Torp 2002, p. 221).

Le personnel de base du LMT avait des réserves quant à la randomisation de leur jugement professionnel (Torp et al., 1993, p. 101). Les chercheurs ont donc accepté d'exclure certains cours de l'expérience. « Il s'agissait notamment des cours pour lesquels les employeurs partenaires insistaient sur le respect des procédures d'affectation habituelles, des cours destinés à des groupes cibles spécifiques et des cours de suivi. »

La pratique a également dévié du plan initial : après la randomisation, certains membres du personnel ont « désactivé » les candidatures de personnes qu'ils jugeaient moins aptes à suivre certains cours. Cela a permis de libérer des places pour les personnes figurant sur la liste d'attente (aléatoire) (Raaum et Torp, 2002).

Le texte ne précise pas comment ces écarts par rapport à l'arbitraire affectent les comparaisons entre le traitement et le groupe témoin. Si les comparaisons portent sur les groupes de traitement et de contrôle initiaux, définis par randomisation, les estimations devraient rester non biaisées pour les effets de l'expérience telle qu'elle a été réellement menée – « l'expérience » incluant le filtrage post-randomisation effectué par l'équipe de recherche.

Cependant, les deux groupes présentent des différences statistiquement significatives au départ, ce qui est inhabituel et préoccupant. Le groupe traité compte davantage de femmes que le groupe témoin (62 % contre 53 %), une plus grande proportion de personnes nées en Norvège (96 % contre 93 %) et un niveau d'éducation plus élevé (38 % contre 30 % ont terminé leurs études secondaires, 10 % contre 7 % ont fait des études supérieures ; Raaum et Torp, 2002, tableau 1). L'article ne présente pas de test global d'équilibre statistique. Nous constatons que 13 des 37 tests individuels d'égalité des caractéristiques individuelles – soit plus d'un tiers – sont significatifs ($p < 0,1$). Il est plausible qu'un si grand nombre de différences significatives soient dues au hasard si les différentes caractéristiques démographiques sont fortement corrélées, car dans ce cas, les 37 tests ne sont pas indépendants. En l'absence d'informations sur les corrélations, nous supposons par défaut que la randomisation a connu un dysfonctionnement.

Au cours des deux années civiles suivant la formation, 1993 et 1994, le groupe expérimental a gagné respectivement 11 000 et 14 000 couronnes de plus que le groupe témoin, soit environ 17 % de plus (Raaum et Torp 2002, tableau).

1) ⁶³ Une évaluation naïve non randomisée, qui se compare à un échantillon de personnes qui ne s'applique pas au LMT, génère des nombres plus importants : 20 000 et 22 000 (Raaum et Torp 2002, Tableau 2, premier panneau).

Raaum et Torp (2002) apportent des preuves supplémentaires que l'évaluation randomisée peut produire des résultats systématiquement différents de ceux de l'évaluation non randomisée, en l'occurrence en réduisant l'estimation de l'impact de plus de 40 %. Il est possible que, dans des conditions expérimentales plus rigoureuses, les estimations d'impact randomisées auraient été encore plus faibles.

Rosholm et Skipper (2009), « La formation sur le marché du travail est-elle un fléau pour les chômeurs ? Preuves tirées d'une expérience sociale »

Une autre étude se déroule au milieu des années 1990 au Danemark, un pays nordique. Elle aussi randomise un programme national de formation professionnelle. Et elle aussi rencontre des difficultés pour mener à bien l'expérience.

Au milieu des années 1990, les programmes de formation professionnelle danois (AMU) proposaient des formations courtes pour aider les chômeurs et les personnes en emploi à se reconvertir. Environ 600 formations étaient offertes, d'une durée de 1 à 7 semaines (2 semaines en moyenne). Dans des conditions normales, hors contexte expérimental, les personnes en emploi étaient prioritaires, ce qui laisse supposer que ces formations, outre leur brièveté, n'étaient pas spécifiquement conçues pour les personnes sans emploi.

Dans l'expérience rapportée par Rosholm et Skipper (2009), cependant, la priorité a été inversée, afin que les résultats puissent témoigner de la capacité des programmes à remettre les gens au travail (Rosholm et Skipper 2009, pp. 341–42).

Bien que les groupes de traitement et de contrôle définis par randomisation semblent ici statistiquement équilibrés, 48 % des participants du ⁶⁴ groupe de traitement ont interrompu le traitement, tandis que 22 % des participants du groupe de contrôle ont, d'une manière ou d'une autre, ... Compris (Rosholm et Skipper 2009, p. 338-339). Une différence de 30 ⁶⁵ Comme dans les études de la WIA (§4.4), le points dans le taux d'adoption entre les groupes réduit la puissance statistique.

Les participants ont commencé leur formation en mai et juin 1994 et la plupart l'ont terminée avant la fin de l'année. Rosholm et Skipper (2009, tableaux IV à VI, colonne 1) relèvent des indices d'impacts négatifs sur l'emploi et les revenus en 1995-1996. Si ces effets sont avérés, les auteurs avancent l'hypothèse d'un effet de blocage temporaire. Les personnes n'ayant pas suivi de cours au cours du second semestre 1994 ont consacré davantage de temps à la recherche d'un emploi, ce qui s'est avéré payant en moyenne en 1995-1996.

⁶³ En contrôlant toutes les caractéristiques de base, y compris les revenus avant traitement, on obtient 10 000 et 14 000 (Raaum et Torp 2002, tableau 2, premier panneau).

⁶⁴ Bien que cela ne soit pas immédiatement évident dans l'article (Rosholm et Skipper 2009, tableau II), 4 des 34 Les comparaisons intergroupes sur des variables prédéterminées sont significatives ($p < 0,1$). Deux d'entre elles concernent le niveau d'études et ne sont donc pas indépendantes. Cette proportion correspond approximativement à celle attendue par hasard dans des conditions d'équilibre parfait.

⁶⁵ « Concernant le degré élevé de chevauchement, aucune explication claire n'est facilement disponible. » (Rosholm et al.) Skipper 2009).

Quoi qu'il en soit, les auteurs soulignent que deux semaines de formation conçues pour les personnes en emploi ne devraient avoir qu'un faible impact sur les chômeurs. La taille modérée de l'échantillon (810) et l'impact limité sur la participation rendent impossible la détection fiable de faibles effets.

Brenninkmeijer et Blonk (2012), « L'efficacité du programme JOBS auprès des chômeurs de longue durée : une expérience randomisée dans le Pays-Bas »

Il s'agit d'une étude psychologique dont le protocole randomisé est entaché de défauts et dont l'échantillon est petit.

À Lelystad, aux Pays-Bas, 160 personnes bénéficiant d'allocations chômage ont été réparties aléatoirement en trois groupes : un groupe témoin ; un groupe recevant des bons d'achat pour financer des formations, notamment de conduite ; et un groupe suivant le programme « JOBS », un ensemble de cinq demi-journées de formation visant à renforcer la confiance en soi et la motivation dans la recherche d'emploi. Six mois après le programme, 26 % des participants du groupe « JOBS » avaient trouvé un emploi, contre 9 % dans le groupe recevant les bons d'achat et 11 % dans le groupe témoin (Brenninkmeijer et Blonk, 2012, p. 225).

Ces résultats proviennent toutefois d'un échantillon réduit, présentant un risque important d'attrition et de biais de sélection. 35 participants, répartis de manière assez homogène entre les groupes, ont abandonné l'étude. Sept autres ont apparemment trouvé un emploi avant le traitement, et 21 autres, initialement non affectés au groupe « Emploi », s'y sont inscrits volontairement. L'étude exclut de son analyse ces deux derniers groupes.

(Brenninkmeijer et Blonk 2012, p. 223.) Face à ces abandons et à ces changements de groupe, les études plus rigoureuses conservent tous les sujets et continuent de les classer selon leur affectation aléatoire. Seule cette méthode permet d'obtenir des comparaisons non biaisées.

Hendra et al. (2011), Rompre le cycle des bas salaires et de l'absence de salaire : données finales de l'enquête britannique sur le maintien et la promotion de l'emploi (ERA) Démonstration

Il s'agit d'une rare évaluation à l'américaine d'un programme européen, c'est-à-dire une évaluation dans laquelle l'organisme qui met en œuvre l'intervention initie et adhère pleinement à l'expérimentation randomisée, et fournit le financement substantiel et durable nécessaire à une évaluation professionnelle.

En effet, c'est une organisation américaine qui a mené l'évaluation de ce programme de formation professionnelle : MDRC, créée par la Fondation Ford dans les années 1970 précisément pour ce type de travaux.

L'étude apporte finalement peu d'informations sur notre sujet d'intérêt, la formation professionnelle. Comme son nom l'indique, le programme britannique de maintien et de promotion de l'emploi (ERA) visait à aller au-delà des programmes gouvernementaux existants qui, en pratique, s'arrêtent à la porte de l'employeur – des programmes qui aident les personnes à trouver un emploi puis à le quitter. L'ERA proposait un ensemble de services et d'incitations pour aider les personnes à se maintenir en emploi et à progresser dans leur carrière. Le financement de la formation constituait l'un de ces éléments.

Une fois embauchés, les participants au programme ERA pourraient bénéficier d'au moins deux ans de conseils et d'assistance de la part d'un conseiller en emploi pour les aider à poursuivre leur travail et

L'avancement professionnel était possible. Les employés travaillant à temps plein de façon constante pouvaient recevoir d'importantes primes, appelées « primes de fidélisation ». Les participants pouvaient également bénéficier d'une aide financière pour leurs frais de scolarité et de récompenses pour la réussite de formations. (Hendra et al., 2011)

La manifestation ciblait trois groupes de personnes définis par la structure existante des programmes d'aide sociale :

- Des parents célibataires sans emploi avec de jeunes enfants, qui s'étaient volontairement inscrits à un programme d'aide sociale à l'emploi appelé New Deal for Lone Parents (NDLP).

Le crédit d'impôt pour les travailleurs (WTC), qui était analogue au crédit d'impôt sur le revenu gagné aux États-Unis.
- Des parents célibataires d'enfants d'âge scolaire, qui travaillaient à temps partiel et recevaient des allocations.

L'allocation de recherche d'emploi est une prestation financière destinée aux personnes recherchant activement un emploi. Les personnes relevant de cette catégorie étaient déjà tenues de participer à un programme appelé New Deal 25 Plus, ce qui a donné son nom au groupe : ND25+. (Hendra et al., 2011, p. 1.)

Les participants aux programmes NDLP et WTC étaient majoritairement des femmes, tandis que ceux du groupe ND25+ étaient majoritairement des hommes. Ces derniers présentaient également une situation socio-économique nettement plus défavorable. 36 % d'entre eux n'avaient aucun diplôme, contre 23 % et 12 % respectivement pour les groupes NDLP et WTC. « Les problèmes de santé, les antécédents de toxicomanie et les démêlés avec la justice étaient fréquents. [Le groupe ND25+] était généralement perçu comme difficile à aider. » (Hendra et al., 2011, p. 8)

Étonnamment, l'ERA a eu l'impact le plus important sur les personnes considérées comme les plus difficiles à aider. Les résultats concernant le groupe NDLP (les personnes ayant de jeunes enfants) indiquent une légère hausse initiale, mais transitoire, des revenus, de 308 £ au cours du premier exercice fiscal de suivi. L'effet sur la durée d'emploi est resté quasiment nul. (Hendra et al., 2011, tableau 4.1). Les résultats sont tout aussi faibles pour le groupe WTC (avec des enfants plus âgés ; Hendra et al., 2011, tableau 4.2). En revanche, dans le groupe ND25+, l'ERA a augmenté l'emploi de 0,2 mois par an, soit un effet global sur l'emploi d'environ 0,2 / 12 = 2 points de pourcentage. De plus, elle a permis d'accroître les revenus de 296 £ par an sur cinq ans, sans signe de ralentissement à la fin. (Hendra et al., 2011, tableau 6.1)

Le programme ERA encourageait la formation professionnelle de plusieurs manières. Les travailleurs sociaux incitaient les participants à l'envisager. Le programme remboursait jusqu'à 1 000 £ de frais de formation et offrait une prime de 1 000 £ à la fin de chaque formation, ainsi que 8 £ par heure de formation. Pourtant, seule une minorité de participants a profité de ces incitations : environ 15 % dans le groupe NDLP, 33 % dans le groupe WTC et 10 % dans le groupe ND25+ (Hendra et al., 2011, tableau 3.5). Le faible taux de participation du groupe ND25+ s'inscrivait dans une tendance plus générale. Les membres de ce groupe avaient passé la majeure partie de leur vie adulte en marge des systèmes sociaux, se contentaient souvent d'un emploi stable et étaient pessimistes quant à leurs perspectives d'avancement (Hendra et al., 2011, p. 188).

Au sein des trois groupes cibles, la formation était donc inversement corrélée à ses effets sur l'emploi et les revenus. Ce résultat peut s'expliquer de diverses manières, et nous ne le considérons donc pas comme une preuve que la formation est néfaste. Toutefois, l'évaluation du programme ERA apporte peu d'éléments indiquant que ces offres de formation gouvernementales ont été bénéfiques. Un problème possible réside dans le fait que les travailleurs sociaux qui conseillaient les clients sur les formations à suivre connaissaient mal les besoins en recrutement des employeurs locaux (Hendra et al., 2011, p. 10). 71-72 ; Hendra et Hamilton 2015, p. 421).

Annexe D : Études basées sur la discontinuité

Les partisans de la révolution de la crédibilité en économie ne sont pas si étroits d'esprit qu'ils ne croient qu'aux essais randomisés. D'autres stratégies analytiques (plans de recherche) permettent de concevoir une expérience naturelle de telle sorte que les groupes de traitement et de contrôle, construits statistiquement, soient comparables et que la répartition du traitement entre eux soit, de fait, arbitraire. Cette section traite des études qui prétendent, de manière plus ou moins convaincante, étudier une telle variation aléatoire « quasi » du traitement.

L'une des stratégies potentiellement intéressantes est la méthode de régression par discontinuité (RDD). Nous avons trouvé deux études de formation basées sur la RDD, que nous présentons ici.

Battistin et Rettore (2002), « Test des effets de programme dans un modèle de discontinuité de régression avec conformité imparfaite »

Entre octobre 1995 et juin 1996, la ville de Turin a organisé des formations de 600 heures à l'utilisation de logiciels de comptabilité, de traitement de texte et autres. Ces formations étaient destinées aux personnes sans emploi. Le nombre de candidatures étant deux fois supérieur au nombre de places disponibles, les responsables ont refusé les candidats ayant obtenu un score inférieur à 55 à un test d'aptitudes (Battistin et Rettore, 2002).

Pour assurer le suivi des personnes se situant de part et d'autre de ce seuil, Battistin et Rettore ont cherché à les interviewer par téléphone 17 mois après la fin des cours. Ils ont obtenu un taux de réponse de 89 % (Battistin et Rettore, 2002, p. 45) et un échantillon de 212 personnes. Étant donné que les méthodes décrites dans l'article accordent une importance maximale aux observations les plus proches du score requis et aucune importance à celles qui s'en écartent de plus de 13, 15 ou 17 points, la taille effective des échantillons est probablement bien inférieure à 212. L'article effectue les analyses successivement avec ces trois intervalles de confiance afin d'observer leur influence sur les résultats. La valeur privilégiée, 15, est sélectionnée grâce à un algorithme de validation croisée, qui détermine l'intervalle de confiance permettant d'obtenir les prédictions les plus précises des résultats d'emploi en dehors du seuil.

L'estimation d'impact privilégiée par l'étude est de 12 points d'emploi, ce qui correspond à une valeur p de 0,3 (Battistin et Rettore 2002, tableau 2, ligne 2). Un impact positif est plausible, mais, probablement en raison de la petite taille de l'échantillon, il n'est pas clairement mis en évidence.

Louvain et Oosterbeek (2004), « Évaluation de l'effet des déductions fiscales sur la formation »

Bien que cette étude se déroule dans un contexte différent, avec des variables d'un autre ordre, son raisonnement économétrique abstrait rappelle celui présenté précédemment. En 1996, le gouvernement néerlandais a instauré plusieurs crédits d'impôt pour inciter les entreprises à former leurs employés. L'un d'eux a introduit une discontinuité de qualification : un crédit de 40 % pour les dépenses de formation des employés de plus de 40 ans. Leuven et Oosterbeek (2004) soulignent que les entreprises pouvaient réagir à cette incitation de deux manières : soit en formant davantage les plus de 40 ans, soit en attendant que les moins de 40 ans atteignent l'âge requis pour les former ultérieurement. Cette dernière option permet de bénéficier du crédit d'impôt sans pour autant augmenter le volume total de formation.

Quoi qu'il en soit, si les personnes de 41 ans ont bénéficié d'une formation nettement supérieure à celle des personnes de 39 ans, cela constitue une expérience naturelle. Afin d'évaluer les impacts, ces auteurs ont également mené une enquête téléphonique en 1999. Ils ont demandé aux personnes interrogées si elles avaient suivi une formation au cours des 12 derniers mois et quel était leur salaire actuel (Leuven et Oosterbeek, 2004, p. 471). Les questions portant sur une période de 0 à 12 mois, un délai moyen de 6 mois s'est écoulé entre la formation et le suivi.

Leuven et Oosterbeek (2004, tableau 1, colonne 4) démontrent que cette quasi-expérience a eu lieu : les personnes de 41 ans avaient 17 points de plus de chances de s'être entraînées au cours de l'année précédente que celles de 39 ans, et ce, avec une signification statistique élevée, même dans un échantillon de 149 personnes. (Les personnes âgées de 40 ans au moment de l'enquête ont été exclues car elles se seraient situées de part et d'autre du seuil au cours de l'année précédente.) En élargissant les tranches d'âge de part et d'autre (38-39 ans et 41-42 ans, ou 37-39 ans et 41-43 ans, etc.), la différence estimée diminue progressivement, et les erreurs standard se réduisent. Ceci illustre ^{une} anniversaire. À l'anniversaire. comment l'élargissement de la plage d'âge accroît la précision de la mesure, même s'il modifie l'objet de la mesure.

L'impact sur la rémunération est quasiment indiscernable de 0, quelle que soit la bande passante utilisée (Leuven et Oosterbeek 2004, Tableau 4, premier panneau).

La question de recherche de cet article est la suivante : les entreprises qui ont formé des employés en réponse au crédit d'impôt les ont-elles mieux rémunérés environ six mois plus tard ? La réponse : pas vraiment, d'après nos observations. On ignore si, à plus long terme, ces employés ont bénéficié d'augmentations de salaire ou ont trouvé des emplois mieux rémunérés ailleurs.

Annexe E : Concepts préliminaires à la méta-analyse

effets fixes et aléatoires

Lorsqu'on compare les résultats de plusieurs études, une méthode consiste à fusionner mathématiquement les échantillons de chaque étude. C'est ce qu'on appelle la méthode à effets fixes ou à effets communs. Une étude portant sur 1 000 personnes aura un poids dix fois supérieur à celui d'une étude portant sur 100 personnes. La méta-analyse à effets communs repose sur l'hypothèse que toutes les études regroupées estiment la même chose, par exemple l'impact du TAA sur les jeunes Blancs de moins de 30 ans licenciés des usines automobiles de l'Ohio en 2002.

Il est souvent plus réaliste de supposer que différentes études mesurent des choses quelque peu différentes – dans notre cas, l'impact de différents types de formation sur différents types de personnes dans différents contextes. L'approche à effets aléatoires reconnaît cette complexité, mais la maîtrise (généralement) en supposant que, dans les différentes études, l'impact moyen est distribué selon une courbe normale ou « en cloche ». En pratique, la méta-analyse à effets aléatoires surpondère les ⁶⁶ Le centre de cette courbe correspond à ce que nous appellerons la moyenne des effets aléatoires. ⁶⁷ Dans petites études. L'idée est que toutes les études sont informatives quant à la façon dont l'impact varie selon les contextes, de sorte que les petites études sont précieuses de manière disproportionnée par rapport à leur taille (Borenstein et al., 2010). L'étude portant sur 1 000 personnes a un poids inférieur à 10 fois celui de l'étude portant sur 100 personnes – bien qu'elle ait tout de même un poids plus important.

biais de publication

Le biais de publication survient lorsque les études produisant les estimations d'impact les plus importantes (ou, dans certains cas, les plus faibles) sont plus susceptibles d'être publiées et donc d'être incluses dans une méta-analyse.

Une méthode courante pour détecter un biais de publication dans une méta-analyse consiste à vérifier la corrélation entre l'estimation moyenne de l'impact d'une étude et sa précision, laquelle dépend notamment de la taille de l'échantillon. Le raisonnement est le suivant : imaginons un vaste ensemble de sondages d'opinion sur la popularité présidentielle, avec des échantillons allant de 20 à 2 000 répondants. Il s'agit d'un scénario à effets communs : tous les sondages posent la même question à la même population, au même moment. Supposons que le taux d'approbation réel soit de 50 %. En l'absence de biais de publication, on s'attendrait à ce qu'un nombre à peu près égal de sondages, quelle que soit leur taille, surestiment ce taux que le nombre de sous-estimateurs. De plus, en raison de leur imprécision, les résultats des petits sondages seront plus dispersés autour de la valeur réelle. Si un biais de publication était présent – par exemple, si les résultats inférieurs à 45 % étaient exclus – cela éliminerait davantage de petits sondages. Cela aurait pour effet de rehausser la moyenne des petits sondages ayant atteint le public. Pour l'ensemble des sondages publiés, le résultat serait inversement corrélé à la précision.

Les chercheurs peuvent vérifier cette corrélation après avoir rassemblé un ensemble d'études, ce qui est impossible avec une seule étude. Ils peuvent effectuer cette vérification graphiquement, à l'aide d'un diagramme en entonnoir, ou numériquement par différentes méthodes.

Le biais de publication est particulièrement pernicieux dans les méta-analyses à effets aléatoires. En effet, ce biais se manifeste surtout parmi les petits sondages, auxquels les méthodes à effets aléatoires accordent un poids disproportionné.

Comparaisons intergroupes

Dans le cadre de l'analyse des méta-analyses, nous procéderons à des comparaisons entre groupes. Les programmes privés sont-ils plus efficaces que les programmes publics ? Sont-ils plus performants dans le Midwest ? Même lorsque les études individuelles sont convaincantes quant à la causalité – grâce, par exemple, à la randomisation –, les comparaisons entre elles ne le sont pas. Si, par hasard, des approches efficaces ont été évaluées à New York et non dans le New Jersey, cela ne prouve pas que la formation est plus efficace à New York. Pire encore, le nombre d'études incluses dans la plupart des méta-analyses est faible d'un point de vue statistique. Une méta-analyse pourrait inclure trois études menées à New York et deux dans le New Jersey.

⁶⁶Toutes les méthodes à effets aléatoires ne supposent pas la normalité.

⁶⁷Les noms dans la littérature incluent « effet global », « effet combiné », « effet groupé » et « effet moyen ».

Pour ces raisons, les comparaisons intergroupes sont pertinentes et doivent être effectuées avec soin.

Annexe F : Méta-analyse avec terme de persistance

Ce tableau présente les résultats des méta-régressions « survivant » lorsque l'impact de la période précédente

L'estimation est prise en compte. Les coefficients du nouveau terme de « persistance » apparaissent dans la première ligne.

L'horizon précédent pour les méta-régressions à moyen terme (année 2 après la randomisation) est le

à court terme (année 1) ; à long terme (années 3 à 5), il s'agit de moyen terme.

	Emploi				Gains			
	moyen terme		à long terme		moyen terme		à long terme	
Persistance								
Impact à l'horizon précédent	0,18*	0,15**	0,48***	0,32** (0,10)	0,45***	0,39***	0,52***	0,25*
	(0,10)	(0,08)	(0,15)		(0,10)	(0,09)	(0,12)	(0,14)
Cadre (omis = Sud, urbain)								
Année de randomisation	-0,09*	-0,06	-0,06*	-0,03				
	(0,05)	(0,04)	(0,03)	(0,04)				
Nord-Est	3,69**	3,31***						
	(1,31)	(1,16)						
Midwest	3,73***	3,07***	-0,07	-0,02				
	(1,13)	(0,99)	(0,74)	(0,69)				
Ouest	1,95*	2,04*						
	(0,98)	(1,08)						
Taux de chômage après suivi (%)	0,80**	0,85**			180	166		
	(0,38)	(0,36)			(156)	(144)		
Population cible (omise = adultes bénéficiant de l'aide sociale ou à faible revenu)								
abus de substances			6,77***	7,24*** (1,05)			3315***	3931***
			(0,77)				(271)	(346)
impliqué dans le système judiciaire							1584***	2162***
							(263)	(356)
travailleurs déplacés	-3,25**	-2,49***	-1,27*	-1,52*				
	(1,21)	(0,93)	(0,71)	(0,80)				
Jeunesse	-2,54	-3,40***						
	(1,48)	(1,32)						
Invalidité	5,43**	5,58*	3,56***	5,20***				
	(2,45)	(2,87)	(1,15)	(1,53)				
Données démographiques								
% mâle								1
								(4)
Traitement (omis = obligatoire)								
Formation en salle de classe	0,71	0,41	0,18					
	(1,02)	(1,03)	(0,82)					
Formation en cours d'emploi						263		
						(283)		
Volontaire		-1,01						
		(0,94)						
Coût par personne traitée (en milliers de dollars)		0,00*						
		(0,00)						

Engagement des employeurs									
L'employeur s'engage à embaucher				4693*** 5074*** (1349)			4651*** 6854***		
				(1214)			(1108) (1204)		
Caractéristiques de l'étude (données omises = source des résultats : documents officiels)									
Source des données sur les résultats : enquête							692***		26
							(173)		(358)
imputation multiple									
R²		0,54		1,00		0,87		0,98 0,91	
Nombre d'expériences		73 0,67 87		51 0,56 64		100 0,85 121		75 81	
Remarques : Erreurs standard entre									
* p < 0,10, ** p < 0,05, *** p < 0,01. Estimations REML. Colonnes MI () regroupées sur 10 imputations via les règles de Rubin ;									
parenthèses. Les autres colonnes utilisent la casse entière.									

Annexe G : Analyse coûts-avantages

Dans les sections 6.5.3 et 7.3, nous comparons les coûts et les avantages de la formation professionnelle subventionnée aux États-Unis. à partir des moyennes méta-analytiques du coût brut de la formation et de la valeur actuelle nette (VAN) des impacts sur les salaires avant impôt, les deux étant exprimés en dollars de 2025 par membre du groupe de traitement groupe. Nous expliquons ici la méthodologie.

Comme c'est généralement le cas, nous distinguons trois perspectives : celle des membres du groupe de traitement ; celle des gouvernements, qui, dans cet exercice, sont supposés financer la formation ; et celle de la société dans son ensemble. La perspective sociale est la somme algébrique des deux autres. Ainsi, dans la perspective sociétale Dans cette perspective, les transferts liés à la fiscalité et aux programmes de protection sociale s'annulent. Par simplicité, les gouvernements fédéral, étatiques et locaux sont regroupés.

Nous effectuons l'analyse pour les programmes de formation initiale et pour l'ensemble des programmes sectoriels. identifié dans le tableau 5. Nous suivons de près la méthodologie coût-bénéfice de Schaberg et L'évaluation de WorkAdvance par Greenberg (2020), sauf en ce qui concerne l'estimation des impacts sur la durée de vie, revenus avant impôt.

Un élément clé pour comprendre les avantages est : combien de temps durent les gains ? Seulement trois Les expériences de notre méta-analyse ont initialement produit des impacts positifs sur les bénéfices, puis Ce suivi s'est étendu sur plus de 6 ans, ce qui rend les preuves concernant la persistance à long terme assez minces. 68 Nous estimons les impacts à long terme en nous appuyant sur des recherches concernant la persistance des chocs de revenus. Karahan et Ozkan (2013) étudient la persistance des chocs de revenus dans l'étude de panel de Dynamique des revenus (PSID). Karahan et Ozkan (2013, tableau 1, partie gauche) estiment qu'après En contrôlant l'âge, le niveau d'éducation et les effets fixes individuels, la corrélation sérielle dans le résidu La composante logarithmique des revenus commence à 0,7 à 24 ans, passe à 0,98 à 45 ans, puis diminue. légèrement jusqu'à l'âge de 60 ans, le tout modélisé par une courbe cubique. 69 Pour appliquer ces multiplicateurs, nous

68Le groupe féminin du projet de démonstration national de travail soutenu (Couch 1992), projet QUEST (Roder et Elliott 2021) et Per Scholas dans l'évaluation WorkAdvance (Yusim et al. 2025).
69Karahan et Ozkan (2013, tableau 1, panneau de gauche) présentent des estimations de coefficients pour un modèle cubique pour le Paramètre de corrélation sérielle spécifique à l'âge ρ . Le meilleur ajustement est $\rho = 0,7003 + 0,2974h - 0,0978h^2 + 0,0095h^3$, dans dont $h = (\text{âge} - 24)/10$. Le modèle est ajusté sur des données jusqu'à l'âge de 60 ans, mais nous étendons la spline jusqu'à l'âge de 65 ans.

On suppose que les participants commencent leur formation à 27 ans, l'âge moyen de l'échantillon, et travaillent jusqu'à 65 ans. L'impact sur les revenus des années 3 à 5, exprimé en pourcentage des revenus du groupe témoin, constitue le choc qui se propage. Nous nous inspirons des travaux de Hendren et Sprung-Keyser (2020) et de Mountjoy (2026) pour modéliser le profil de revenus du groupe témoin au cours de son cycle de vie, à partir d'une série de données sur les revenus moyens par âge issues de l'American Community Survey. Les moyennes de l'échantillon transversal de 2015 sont ajustées à la hausse pour tenir compte d'une croissance annuelle des salaires réels de 0,5 %. L'ensemble de la série est ensuite rééchantillonné afin que la valeur à 31 ans – le milieu de la période des années 3 à 5 par rapport à l'âge de départ de 27 ans – corresponde à la moyenne du groupe témoin pour la même période, issue de la méta-analyse.

70

Nous neutralisons les impacts futurs en tenant compte du coût réel à long terme des fonds du gouvernement américain, estimé par le rendement du marché d'une obligation TIPS à 30 ans, actuellement de 2,82 %.

71

En suivant généralement Schaberg et Greenberg (2020), nous comptons ces éléments :

- Le coût du programme correspond à la différence moyenne entre les dépenses du programme par rapport au groupe témoin. Pour les études fournissant cette information, le coût par membre du groupe de traitement est indiqué. Seules les dépenses liées à l'expérience sont prises en compte. Certaines sources fournissent les coûts par participant ; ceux-ci sont convertis en coûts par membre du groupe de traitement à l'aide des taux de participation déclarés.

Réduction des dépenses liées à d'autres formations. Les formations dispensées dans le cadre d'une expérimentation peuvent parfois se substituer à celles proposées par d'autres organismes, y compris publics, réduisant ainsi les dépenses totales de formation. Dans notre ensemble de données, pour les études primaires sur la formation fournissant les deux chiffres, cette définition plus large du coût net est en moyenne inférieure de 20 % à la définition plus restrictive mentionnée précédemment. Pour les seuls programmes sectoriels, l'écart est de 11 %. Ce poste budgétaire tient compte de cet ajustement en appliquant l'une de ces réductions.

- Le revenu avant impôt correspond à la VAN des impacts sur les revenus jusqu'à l'âge de 65 ans, comme décrit précédemment.

- Avantages sociaux, y compris les cotisations sociales patronales. Toutes les études de notre Cette méta-analyse estime les impacts sur les salaires avant impôt. Elle exclut les impacts sur les avantages sociaux, qui font partie intégrante de la rémunération du travail et constituent une source de recettes fiscales pour l'État, sous forme de cotisations patronales. Le ratio salaires avant impôt/rémunération totale varie selon l'année et la profession. D'après les données du Bureau of Labor Statistics (BLS) sur les coûts de l'employeur pour la rémunération des employés (ESEC), les salaires avant impôt représentaient 72,4 % de la rémunération totale en 2025 pour le groupe professionnel « emplois de bureau et de soutien administratif », une catégorie professionnelle représentative des postes occupés par les jeunes diplômés. Nous estimons l'impact sur la VAN des avantages sociaux à $(1 - 0,724)/0,724$ fois la VAN de ⁷² Nous prenons ce niveau d'entrée l'impact sur les revenus avant impôt.

70 Si $Y0^{*t}$ représente le revenu moyen du groupe témoin à l'âge t , et r l'impact proportionnel au cours des années 3 à 5, le bénéfice estimé est $Y0^{*t}[(1+r)\tilde{p}-1]$ où $\tilde{p} = \Pi_t = 32 \text{ à } 65$ et p_t est modélisé comme décrit dans la note 10. [71 fred.stlouisfed.org/series/DFII30](https://fred.stlouisfed.org/series/DFII30), Données du 18 mai 2026. [72 bls.gov/web/ecec/ecec-civilian-dataset.xlsx](https://www.bls.gov/web/ecec/ecec-civilian-dataset.xlsx) consulté le 25 mai 2026.

Les charges sociales (à la charge des employés et de l'employeur) sont estimées à 7,65 % du salaire brut, soit le taux standard payé par les employés. Ce calcul suppose qu'aucun salarié ne dépasse le plafond de revenus pour le calcul des cotisations de sécurité sociale, actuellement fixé à 184 500 \$ par an. Les charges patronales, y compris l'assurance chômage, sont estimées à 6,85 % de la rémunération totale, toujours selon les données de l'ECEC pour les « professions de bureau et de soutien administratif » de 2025. Il est important de noter que ce calcul considère les charges sociales comme de simples transferts à l'État. Il ne prend pas en compte les prestations d'assurance et de retraite auxquelles elles donnent droit aux salariés.

Impôts sur le revenu. Schaberg et Greenberg (2020) estiment le taux marginal d'imposition sur le revenu des programmes Per Scholas et St. Nicks, tous deux situés dans l'État de New York, à 19,0 %. Les taux pour Towards Employment et Madison Strategies, dans l'Ohio et l'Oklahoma respectivement, sont de 15 % et 17,6 %. Tous ces chiffres incluent les impôts sur le revenu fédéraux et d'État. Nous utilisons la moyenne des quatre taux, soit 17,2 %.

Taxes sur les ventes. Conformément à l'hypothèse de Schaberg et Greenberg (2020), nous supposons que l'intégralité du revenu net est consacrée à l'achat de biens et services soumis à la taxe sur les ventes. Après pondération par la population de chaque État, le taux moyen national de la taxe sur les ventes au 1^{er} janvier 2026 était de 5,69 ⁷³ %.

- Recours aux aides sociales. La diminution du recours aux aides sociales, telles que les bons alimentaires et l'assurance chômage, représente une perte pour les stagiaires et un gain pour l'État.

L'estimation présentée ici est approximative. Dans le cadre du projet WorkAdvance, l'enquête de la deuxième année interrogeait les participants sur les prestations sociales qu'ils percevaient. Cela a permis aux chercheurs d'estimer l'impact de la formation sur la probabilité de recours à chaque prestation sociale. Ils ont ensuite intégré des données distinctes sur les dépenses typiques par bénéficiaire afin d'estimer l'impact sur le recours aux prestations sociales en valeur monétaire. Sur les quatre sites WorkAdvance, le ratio moyen entre la réduction des prestations et l'augmentation des revenus est de 14,8 % (Schaberg et Greenberg 2020, tableaux 3.1, 3.3, 3.4 et 3.5). Nous utilisons ce ratio pour estimer la réduction moyenne dans nos échantillons d'études.

Les dépenses liées au travail. Travailler davantage implique parfois de dépenser plus en biens et services nécessaires à l'exercice de l'activité professionnelle, mais sans valeur intrinsèque, comme les uniformes et les titres de transport. Nous reprenons ici les estimations de Schaberg et Greenberg (2020, p. 77), qui évaluent le ratio des dépenses marginales liées au travail aux revenus marginaux avant impôt à 8,34 %.

- Temps non rémunéré. Plus de temps consacré à un travail rémunéré signifie moins de temps non rémunéré, ce qui pourrait signifier moins de loisirs, ou moins de temps consacré à des tâches non rémunérées comme la garde d'enfants, sont autant de facteurs qui peuvent impacter les revenus. Citant une étude révélant que certains employés de bureau valorisaient leur temps non rémunéré à 58 % de leur salaire, Schaberg et Greenberg (2020, p. 80) estiment que le coût du temps non rémunéré perdu représente 50 % de l'impact de la formation professionnelle sur les revenus. Toutefois, ce coût dépend en principe de...

73 Taux de taxe combinés étatiques et locaux d'après Abir Mandal, « State and Local Sales Tax Rates, 2026 », Tax Foundation, taxfoundation.org/data/all/state/sales-tax-rates, Consulté le 25 mai 2026. Populations des États selon le Bureau du recensement des États-Unis, « Totaux et évolution de la population des États : 2020-2025 », census.gov/data/tables/time-series/demo/popest/2020s-state-total.html, consulté le 25 mai 2026.

Nous nous intéressons ici aux impacts sur l'emploi, et non sur les revenus. Par conséquent, nous nous distinguons de Schaberg et Greenberg en intégrant l'effet sur l'emploi. Nous modélisons et actualisons les impacts sur l'emploi jusqu'à l'âge de 65 ans de la même manière que nous le faisons pour les impacts sur les revenus, ce qui signifie que le ratio de la VAN des impacts sur les impacts des années 3 à 5 est identique pour l'emploi et les revenus.

Contrairement à Schaberg et Greenberg, nous ne prenons pas en compte la perte sèche liée à l'augmentation des recettes publiques pour financer la formation. Comme le soulignent Hendren et Sprung-Keyser (2020), l'utilisation d'un coefficient multiplicateur de perte sèche donné suppose une théorie sur la manière dont ces recettes sont générées. Nous éludons cette question, préférant séparer l'analyse de la formation professionnelle de la question de son financement.

L'analyse coûts-avantages est présentée dans le tableau G1 pour tous les programmes de formation primaire de notre échantillon, et dans le tableau G2 pour les programmes sectoriels.

Nous résumons les coûts et les avantages à l'aide de cinq statistiques, qui figurent au bas des tableaux :

Nous définissons un ratio avantages-coûts restreint comme l'impact sur la rémunération totale (somme des lignes 3 et 4) divisé par le coût net de la formation (ligne 1). Ce ratio compare l'augmentation totale des rémunérations (en faisant abstraction des effets d'équilibre général) au coût budgétaire initial et ne tient pas compte des effets sur les recettes fiscales ni de la fourniture d'autres prestations sociales, notamment d'autres formations. Cette définition peut s'avérer pertinente pour les législateurs chargés de l'allocation du budget de cette année et qui s'intéressent moins aux impacts fiscaux et sociaux, car ceux-ci se manifesteront à long terme ou relèvent de la compétence d'autres commissions. Elle est également pertinente pour les financeurs privés, puisqu'ils ne perçoivent pas d'impôts ni ne distribuent d'aides alimentaires.

Le ratio avantages-coûts est de 1,44 pour l'ensemble des programmes de formation initiale. Compte tenu des incertitudes inhérentes à l'analyse coûts-avantages, nous considérons ce résultat comme légèrement supérieur au seuil de rentabilité. Le ratio est de 7,18 pour les seuls programmes sectoriels.

Le ratio avantages-coûts élargi est identique au ratio restreint, à ceci près que son dénominateur tient compte de la réduction des dépenses consacrées à d'autres formations. Il est de 1,80 pour les programmes de formation initiale et de 8,07 pour les programmes sectoriels.

Un autre indicateur, la valeur marginale des fonds publics (VMFP ; Hendren et Sprung-Keyser, 2020), adopte le point de vue de l'État et prend en compte tous les effets budgétaires. Il est calculé en divisant le bénéfice net par membre du groupe de traitement par le coût net pour l'État. Ces deux chiffres figurent au bas des tableaux principaux.

Pour les programmes de formation initiale, le ratio MVPF est de 2,52. Pour les programmes sectoriels, ce ratio est négatif (-1,98) car ces programmes s'autofinancent trois fois grâce aux impôts et aux économies réalisées, et produisent donc un bienfait pour un coût négatif. Conformément à Hendren et Sprung-Keyser, nous considérons ce ratio comme infini dans le tableau.

Nous calculons également les ratios entre le coût net et le coût brut pour l'État. Nous estimons que...

Le gouvernement récupère environ les trois quarts de ses dépenses sur la formation primaire moyenne.
programme – grâce à des recettes fiscales plus élevées, des prestations sociales réduites et des dépenses moindres sur d'autres formations, ce qui représente un coût net de 0,24 \$ par dollar dépensé. Pour les programmes sectoriels, Le ratio est de -1,88.

Enfin, puisque les résultats dépendent du choix discutable du taux d'actualisation, nous calculons

le taux de rendement interne (TRI), c'est-à-dire le taux d'actualisation d'équilibre. Le TRI est infini pour les bénéficiaires, puisqu'ils n'ont aucun frais à déboursier. Et c'est peut-être moins pertinent pour

Les gouvernements étant plus nombreux que les deux indicateurs précédents, nous le calculons uniquement à partir de l'indicateur sociétal. point de vue. Le TRI sociétal est de 5,97 % pour les programmes de formation primaire et de 34,03 % pour les programmes de formation continue. programmes sectoriels.

	Traitement	
	membre du groupe Gouvernement	Société
Coût du programme	-13 598	-13 598
Réduction des dépenses consacrées à d'autres formations	2 778	2 778
Revenus avant impôt	14 146	14 146
Avantages sociaux, y compris les cotisations patronales	5 379	5 379
Cotisations sociales (employé et employeur)	-2 420	2 420
Impôts sur le revenu	-2 434	2 434
taxes de vente	-572	572
Utilisation des prestations sociales	-2 096	2 096
Dépenses liées au travail	-1 180	-1 180
Temps non rémunéré	-2 511	-2 511
Total	8 312	-3 299
B/C étroit	1,44	
B/C large	1,80	
MVPF	2,52	
Coût net pour le gouvernement / coût brut	0,24	
TRI sociétal	5,97%	

Tableau G1. Analyse coûts-avantages des programmes de formation primaire

Remarques : Tous les montants sont exprimés en dollars de 2025 par membre du groupe de traitement. B/C représente le rapport bénéfice/coût.

Les « avantages sociaux » désignent la rémunération brute, c'est-à-dire la somme des revenus avant impôt et des avantages sociaux, y compris

Les charges sociales patronales sont à la charge de l'employeur. Les « coûts » correspondent aux coûts bruts de formation, calculés comme la moyenne simple. d'après des études fournissant des estimations. La MVPF est la valeur marginale des fonds publics (Hendren et al.).

Sprung-Keyser 2020), calculé comme le bénéfice net pour le participant divisé par le coût net pour Le TRI sociétal est le taux de rendement interne réel du point de vue de la société.

	Groupe de traitement membre	Gouvernement	Société
Coût du programme		-11 602	-11 602
Réduction des dépenses consacrées à d'autres formations		1 289	1 289
Revenus avant impôt	60 319		60 319
Avantages sociaux, y compris les cotisations patronales	22 938		22 938
Cotisations sociales (employé et employeur)	-10 317	10 317	
Impôts sur le revenu	-10 379	10 379	
taxes de vente	-2 441	2 441	
Utilisation des prestations sociales	-8 937	8 937	
Dépenses liées au travail	-5 031		-5 031
Temps non rémunéré	-3 099		-3 099
Total	43 052	21 762	64 814
B/C étroit	7.18		
B/C large	8.07		
MVPF	∞		
Coût net pour le gouvernement / coût brut	-1,88		
TRI sociétal	34,03%		

Tableau G2. Analyse coûts-avantages des programmes sectoriels

Remarques : Voir les notes du tableau précédent.

Ces calculs coûts-avantages sont soumis à plusieurs limitations. Premièrement, nous supposons que

Les gains de revenus découlant de ces programmes diminuent lentement, conformément aux travaux de recherche sur les chocs de revenus.

(Karahan et Ozkan 2013). Cela importe moins pour les programmes sectoriels, puisqu'ils financent...

eux-mêmes, même en ne considérant que les années où les sujets sont suivis. Mais pour la formation uniquement

Dans ces programmes, l'essentiel des bénéfices se manifeste lors de l'extrapolation au-delà du suivi des études.

fenêtres. Sans cette persistance, leur rapport avantages-coûts serait inférieur à 1. Quand

Les données disponibles montrent des gains persistants : parmi les 28 estimations primaires d'entraînement avec

un gain à moyen terme d'au moins 200 \$ et un suivi plus long (généralement les années 2 et 4), à long terme

Les impacts ont totalisé 110 % des impacts à moyen terme, avec un ratio médian de 0,90.

Deuxièmement, les emplois offrent des avantages autres que le salaire. Ils peuvent, par exemple, réduire la criminalité ou améliorer la santé mentale. Les prendre en compte augmenterait le total des avantages, mais nous avons choisi de ne pas le faire en raison du manque de données scientifiques à ce sujet. Enfin, nous négligeons les effets d'équilibre général et de déplacement. Si les stagiaires prennent la place d'autres travailleurs postulant aux mêmes emplois, les avantages que nous mesurons seront surestimés.