

RAPPORT IPP INTERMEDIAIRE – JUILLET 2019

Evaluation du dispositif CIFRE : Rapport intermédiaire

Christopher Balaam
Arthur Guillouzouic
Clément Malgouyres





L'Institut des politiques publiques (IPP) est développé dans le cadre d'un partenariat scientifique entre PSE-Ecole d'économie de Paris (PSE) et le Centre de Recherche en Economie et Statistique (CREST). L'IPP vise à promouvoir l'analyse et l'évaluation quantitatives des politiques publiques en s'appuyant sur les méthodes les plus récentes de la recherche en économie.

www.ipp.eu



LES AUTEURS DU RAPPORT

Christopher Balaam est étudiant à Sciences Po Paris en Économie et Politiques Publiques (EPP). Il s'intéresse aux questions d'innovation et de d'investissements publics en R&D.

Page personnelle : <https://cjbalaam.wixsite.com/site>

Arthur Guillouzouic est économiste à l'IPP. Ancien élève de l'École Normale Supérieure de Cachan et titulaire d'un doctorat en sciences économiques de Sciences Po, ses recherches portent sur les réseaux d'innovateurs, les choix de localisation des firmes innovantes, ainsi que sur les externalités locales générées par le service public.

Page personnelle : <https://arthurguileco.wixsite.com/aguillouzouiclecorff>

Clément Malgouyres est économiste à l'Institut des politiques publiques (IPP) et chercheur associé à l'École d'économie de Paris (PSE). Il s'intéresse à des questions empiriques en économie publique, économie internationale et économie du travail. Ses travaux récents portent sur la fiscalité des entreprises et l'évaluation quantitative des politiques publiques.

Page personnelle : <https://sites.google.com/site/clementmalgouyres>

SYNTHÈSE DES RÉSULTATS

Ce rapport intermédiaire vise à appréhender les caractéristiques du programme Cifre et celles de ses participants et de présenter la méthodologie qui sera adoptée pour le rapport final afin d'évaluer l'effet du programme sur les entreprises.

À partir de plusieurs appariements, nous effectuons une description du dispositif en trois parties. Nous décrivons en premier lieu l'ampleur du programme, en second lieu nous détaillons les caractéristiques des doctorants qui prennent part à ce programme et enfin, nous décrivons les entreprises qui ont recours au dispositif. Le dispositif Cifre attire un nombre croissant d'entreprises et d'étudiants. Ces étudiants sont des doctorants en moyenne de 24 ans, principalement français et de sexe masculin, et effectuent majoritairement de la recherche dans les sciences physiques, naturelles et de l'ingénieur. Les doctorants en Cifre ont une probabilité plus élevée de soutenir leur thèse, avec une hétérogénéité importante selon les disciplines. Après la thèse, ils bénéficient d'une prime de salaire et de contrats de travail plus stables.

Les employeurs des doctorants Cifre sont, quant à eux, majoritairement des entreprises, concentrées dans des secteurs industriels et scientifiques. Cette concentration est aussi géographique, dans des bassins avec une forte densité d'entreprises de recherche ou bien avec quelques grandes entreprises employant un nombre important de doctorants Cifre.

Une analyse préliminaire montre aussi que le recrutement d'un doctorant en Cifre s'accompagne en moyenne d'une hausse d'effectifs et du nombre d'ingénieurs dans l'entreprise. C'est sur ces résultats que se fonde la suite de notre projet, dont nous présentons en dernière partie la méthodologie. Cette analyse est constituée de deux volets. Le premier présente l'approche par étude d'événements. Le second est une analyse en différence de différences avec plusieurs stratégies d'appariement, exactes ou par score de propension.

SOMMAIRE

Synthèse des résultats	2
Introduction	5
1 Présentation du dispositif Cifre	9
1.1 Origine et objectifs du dispositif Cifre	9
1.2 Modalités et fonctionnement du dispositif Cifre	11
1.2.1 Conditions pour l'attribution d'une Cifre	11
1.2.2 Fonctionnement et rémunération du dispositif Cifre	13
2 Description statistique du dispositif Cifre	15
2.1 Description générale du dispositif	16
2.2 Description des doctorants faisant une Cifre	19
2.2.1 Statistiques de base et appariement ABES	20
2.2.2 Résultats issus de l'enquête Génération	25
2.2.3 Résultats issus de l'appariement DADS	39
2.3 Description des employeurs recourant au Cifre	45
2.3.1 Caractéristiques des employeurs	45
2.3.2 Évolution des caractéristiques des entreprises recourant au dispositif Cifre par rapport à des groupes comparables	52
2.3.3 Évolution des entreprises avant et après l'acquisition d'une Cifre	56
3 Méthodologie	61
3.1 La méthode des études d'évènements	61
3.1.1 Équation d'estimation	63
3.1.2 Discussion des hypothèses d'identification	64
3.2 Différence de différences et appariement	67
3.2.1 Description du processus d'appariement	68
3.2.2 Équation d'estimation	69
References	71

Liste des tableaux	73
Liste des figures	75

INTRODUCTION

Contexte de l'étude

Dans le cadre de la réglementation européenne, les autorités françaises ont notifié à la Commission européenne un plan d'évaluation du régime cadre 40391 d'aides d'État à la recherche, au développement et à l'innovation (RDI), validé par la décision 2015/4445 de la Commission. Le régime inclut des aides à la R&D, des aides à l'innovation et des aides aux pôles d'innovation. L'objectif de l'évaluation est de mesurer les effets du régime tant sur les bénéficiaires que sur l'économie en général. Cette évaluation doit aboutir à un rapport intermédiaire en octobre 2019 et un rapport final en juin 2020. Le secrétariat de cette évaluation est assuré par la Direction générale des entreprises (DGE) du ministère français de l'Économie et des Finances tandis que sa direction est assurée par un comité de pilotage.

Une Convention industrielle de formation par la recherche (Cifre) est un dispositif de financement de thèse qui aide les entreprises pour le recrutement de doctorants. Au sein du régime SA 40391, le dispositif Cifre fait partie des aides à la recherche et au développement dont l'objectif global est d'augmenter l'effort de R&D des entreprises bénéficiaires. En outre, le dispositif Cifre a pour objectif de contribuer à l'emploi de docteurs par les entreprises et de favoriser les collaborations entre entreprises et laboratoires.

Ce dispositif a fait l'objet de plusieurs études descriptives qui exploitaient les résultats d'une enquête menée auprès des entreprises, doctorants et laboratoires

bénéficiaires. Le dispositif Cifre n'a, en revanche, jamais fait l'objet d'une évaluation d'impact économique micro-économétrique telle qu'attendue par la Commission. Le plan d'évaluation des aides du régime SA 40391 fixe le cadre sur la méthodologie et les objectifs du plan d'évaluation. C'est dans celui-ci que s'inscrit cette évaluation économétrique d'impact causal du dispositif Cifre.

Objectifs de l'étude

L'objectif principal des évaluations menées dans le cadre du plan d'évaluation des aides d'État notifié à la Commission européenne est d'apprécier et de quantifier les effets des aides tant sur les bénéficiaires que sur l'économie en général. L'étude porte sur une période plus vaste que celle attendue par la Commission et couvre la période 2003-2018, période pour laquelle des données au format numérique sont disponibles pour les conventions Cifre. Les bénéficiaires directs de la subvention accordée dans le cadre du dispositif Cifre sont les entreprises. Cette étude vise à répondre à trois objectifs :

- Identifier et quantifier l'impact causal du dispositif Cifre sur les entreprises bénéficiaires. En particulier, il s'agit d'estimer l'impact direct du dispositif Cifre en termes de R&D dans l'entreprise : évolution du nombre d'emplois R&D, de la rémunération des emplois R&D ou d'autres formes de dépenses de R&D, et évolution du nombre de brevets déposés. L'étude s'attachera aussi à estimer les impacts indirects sur le développement économique de l'entreprise, notamment la valeur ajoutée, le salaire moyen et la productivité.
- Identifier et quantifier l'impact causal du dispositif Cifre sur les doctorants sous convention Cifre. Les effets à étudier portent sur l'employabilité des doctorants ayant soutenu leur thèse à la fin de leur Cifre ainsi que sur la nature et les caractéristiques de l'emploi obtenu (salaires, nature du poste). Cette analyse est estimée pour un ensemble de sous-groupes, selon les carac-

téristiques des doctorants et des entreprises.

- Identifier et de quantifier l'impact causal du dispositif Cifre sur la structure de recherche (le laboratoire) partenaire de la convention Cifre, dans la mesure de la disponibilité des données.

Plan du rapport

Chapitre 1 : Présentation du dispositif Cifre

Chapitre 2 : Description statistique du dispositif Cifre

Chapitre 3 : Méthodologie

CHAPITRE 1

PRÉSENTATION DU DISPOSITIF CIFRE

Ce premier chapitre vise à donner une vue d'ensemble qualitative du dispositif Cifre. Nous présentons tout d'abord les origines et les objectifs de cette convention, dont le but est promouvoir des collaborations entre laboratoires publics et entreprises privées et stimuler la productivité de la recherche scientifique française (section 2.1). Nous précisons ensuite le fonctionnement d'une thèse Cifre, les conditions d'éligibilité des entreprises, des doctorants et des laboratoires, le déroulement annuel et les mécanismes de subvention, prenant en compte à la fois la subvention de l'ANRT et les retours du Crédit Impôt Recherche (CIR) pour le dispositif Cifre (section 1.2). Nous détaillons enfin les évolutions éventuelles du dispositif.

1.1 Origine et objectifs du dispositif Cifre

C'est dans le cadre d'une politique de rapprochement du système de formation, de recherche et d'innovation avec les milieux socio-économiques français que le ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI) confie à l'ANRT la mise en oeuvre des Conventions industrielles de formation par la recherche (Cifre) en 1981.

Dans le cadre d'une convention Cifre, l'État apporte un soutien financier à toute structure socioéconomique établie sur le territoire français qui recrute un doctorant pour lui confier une mission de recherche qui constituera le sujet de sa thèse, pour une durée maximale de trois ans. Le doctorat est consacré entièrement au projet de recherche, avec un partage des efforts et du temps entre l'employeur et le laboratoire académique. Il y a donc deux bénéficiaires de cette convention : l'entreprise, qui reçoit un soutien de l'État, et le doctorant, qui reçoit une subvention pendant trois ans pour son projet et bénéficie d'une double formation académique et professionnelle. L'État subventionne donc des emplois qualifiés au sein de ces entreprises, qui en contrepartie de cette aide à la recherche s'engagent à former et rémunérer le doctorant.

Ce programme a été lancé dans les années 1980, dans un paysage de la recherche en France qui affichait une séparation franche entre le monde de la recherche universitaire public et celui des entreprises privées. Se fondant sur l'exemple américain, le Secrétaire d'État chargé de la Recherche, Pierre Aigrain, présente alors en 1980 comme condition de l'amélioration de la science et technique française « le renforcement des relations entre les universités, les établissements de recherche et les entreprises. » Le projet du dispositif Cifre voit donc le jour, dans le but de tisser et renforcer les liens entre laboratoires publics et entreprises privées par le biais des doctorants. S'il n'y a pas de limite sur le secteur d'activité de l'entreprise, et que le champ de disciplines des thèses des doctorants en Cifre s'est étendu ces dernières années, la recherche effectuée par ces doctorants est à majorité dans le domaine scientifique et technique, c'est-à-dire dans les sciences physiques, les sciences de l'ingénieur et les sciences naturelles.

Au-delà des bénéfices pour le monde de la recherche, le dispositif a des conséquences sur l'intégration professionnelle des doctorants. Les doctorants français ont en effet des débouchés limités, à la fois en terme de rémunération et de stabilité de l'emploi, comme le montrent Margolis et Miotti (2015), ce qui a mené

à la mise en place de dispositifs incitatifs comme le dispositif "Jeunes Docteurs". La thèse Cifre représente donc un vecteur d'intégration pour ces doctorants qui profitent de trois ans d'expérience professionnelle au sein d'entreprises qui sont parfois leurs futurs employeurs. En effet, selon l'enquête de l'ANRT en 2016 sur les Cifre, 27 % des doctorants Cifre sont toujours employés au sein de l'entreprise d'accueil un an après leur thèse.

1.2 Modalités et fonctionnement du dispositif Cifre

Cette section précise les modalités des conventions Cifre. Il existe un certain nombre de conditions appliquées aux doctorants, aux entreprises et aux laboratoires qui souhaiteraient s'engager dans une Cifre. Une fois cette Cifre validée, sa continuation et le versement de la subvention, que nous détaillerons, sont aussi soumis à diverses conditions.

1.2.1 Conditions pour l'attribution d'une Cifre

Un certain nombre de conditions s'appliquent aux entreprises, aux candidats et aux laboratoires souhaitant participer à une Cifre, afin d'en limiter la portée à des projets de recherche fondés avec une vraie valeur ajoutée, et au sein d'entreprises capables de s'engager sur l'emploi d'un doctorant Cifre. Nous obtenons ces conditions du document sur les conditions d'octroi d'une Cifre de l'ANRT (2019).

Les doctorants postulant à une Cifre doivent être titulaires d'au moins un diplôme de niveau Master (Bac + 5). Ils s'engagent à préparer une thèse, mais ne doivent pas avoir déjà été en poste pour le même employeur depuis plus de 9 mois - il doit s'agir du début des travaux. Il n'y a pas de restriction d'âge ou de nationalité ; cependant les doctorants doivent disposer d'une autorisation de travail en France. Ils ne peuvent pas avoir déjà travaillé pour l'entreprise avant la convention Cifre, sauf dans le cadre d'emplois ponctuels ou saisonniers sans rapport avec

leur sujet de recherche. Cette restriction est présente afin de favoriser les nouveaux échanges entre doctorants et entreprises et soutenir un projet de thèse original, et non pas subventionner un emploi qui existerait déjà.

En ce qui concerne les employeurs, ils doivent être domiciliés sur le territoire français. Le sujet de recherche du doctorant doit s'inscrire clairement dans l'objet de l'entité d'accueil (secteur d'activité, travaux en cours). Les entreprises éligibles doivent aussi témoigner d'une capacité d'accompagnement et d'encadrement suffisante (notamment selon le profil du tuteur scientifique choisi) : la décision de l'ANRT à ce sujet est prise en partie selon un rapport du délégué régional de recherche et technologie (DRRT). Ces entreprises s'engagent aussi à rémunérer les doctorants avec un salaire minimum déterminé par l'ANRT, qui était de 23 484 euros en 2018, et qui évolue chaque année, dans le cadre d'un CDD ou CDI. Les Cifre peuvent aussi être accueillis dans des organismes publics, à l'exception des services de l'État, des autorités administratives indépendantes (AAI) et des autorités publiques indépendantes (API). Sont donc éligibles les collectivités territoriales, les établissements dits EPIC ou EPA, et les groupements d'intérêts publics (GIP). Les associations peuvent aussi être employeurs.

Les laboratoires de recherche sont ceux rattachés aux établissements doctoraux dans lesquels sont inscrits les doctorants. Un laboratoire étranger peut être associé à un laboratoire français dans le cadre d'une Cifre.

Les dossiers sont présentés à l'ANRT, qui donne un avis favorable ou défavorable. Cet avis dépend de deux critères :

- Un avis sur la qualité scientifique du projet : l'ANRT statue en prenant en compte la qualité du projet du doctorant, son lien avec l'activité et la recherche de l'entreprise et l'adéquation du laboratoire de recherche choisi par rapport aux disciplines engagées. Pour chaque demande de Cifre, l'ANRT mandate une évaluation de la valeur scientifique du projet auprès d'un expert scientifique, qui rend un avis consultatif.

- La pertinence économique de la convention : la solidité financière de l'entreprise et son implication, surtout pour les entreprises n'ayant jamais eu de Cifre, pas de Cifre depuis 5 ans, ou bien en situation financière précaire. C'est le DRRT qui rend un avis à l'ANRT et au comité d'évaluation et de suivi (CES), qui détient le pouvoir final de décision.

Les employeurs employant un nombre élevé de Cifres (plus de six par an en moyenne sur trois ans) et ayant obtenu l'accord de l'ANRT peuvent passer par une procédure simplifiée de recrutement et de convention.

1.2.2 Fonctionnement et rémunération du dispositif Cifre

Lorsqu'une Cifre est validée, l'entreprise reçoit une subvention d'État d'un montant de 14 000 euros par an et par contrat Cifre (montant en 2018). Chaque convention Cifre est valable pour 36 mois dès sa date de validation, et ces subventions sont versées à échéances trimestrielles lors de ces 36 mois et ne sont pas assujetties à la TVA. Chaque année, l'entreprise doit faire parvenir un rapport annuel détaillé présentant l'avancement des travaux de thèse à l'ANRT. La subvention Cifre est conditionnelle à la réception de ce rapport et à sa cohérence par rapport au projet initial de Cifre. Ainsi, si l'ANRT juge que la recherche effectuée ne correspond pas au projet de recherche auquel elle avait donné son accord, elle peut suspendre la Cifre.

Au-delà de la subvention initiale, le dispositif Cifre représente aussi une subvention non négligeable dans le cadre du CIR. Dans le cas d'une Cifre, le CIR est égal à 30 % d'une fois et demi le salaire brut chargé (un coût d'environnement de 50 % du salaire brut chargé représentant les coûts indirects que supporte l'entreprise), dont on soustrait alors la subvention ANRT. Par exemple, une Cifre rémunérée 23 484 euros par an (le salaire plancher) rapporte 10 595 euros par an en CIR en plus des 14 000 euros de subvention directe de l'ANRT. Ainsi, pour un salaire

de 23 484 euros, le salaire brut chargé est de 32 878 euros – mais la subvention Cifre et le CIR font que le coût final du doctorant est de 8 283 euros par an. Il s’agit donc d’une subvention importante du coût du travail pour ces entreprises.

En continuation de la Cifre, le dispositif jeunes docteurs (DJD) est une incitation supplémentaire à l’emploi des doctorants dans leur entreprise d’accueil. Si le doctorant est embauché en CDI par l’entreprise, qu’il s’agit de son premier CDI et que l’effectif de l’employeur n’ait pas diminué, le salaire chargé est compté double dans le calcul du CIR, et l’augmentation d’environnement passe à 100 %. Un salaire brut chargé de 50 000 euros donne donc lieu à 30 % de 200 000 euros par an, soit 60 000 euros par an. Le DJD s’applique à tous les jeunes docteurs, et se présente comme la suite logique de la Cifre, où l’entreprise peut continuer à employer à moindre coût un doctorant qu’elle connaît déjà et qui est intégré.

CHAPITRE 2

DESCRIPTION STATISTIQUE DU DISPOSITIF

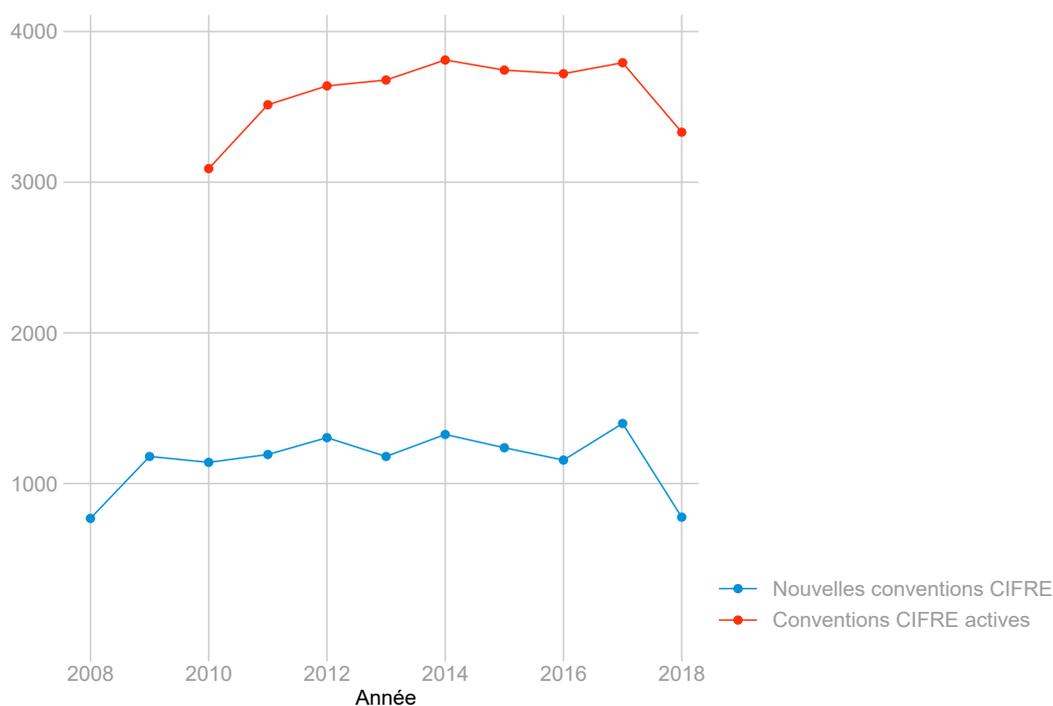
CIFRE

Ce deuxième chapitre est consacré à une description quantitative et statistique du dispositif Cifre, et exploite plusieurs bases de données afin de donner un portrait global du dispositif Cifre, des doctorants et des entreprises qui en font partie. Il ne s'agit pas de mettre en évidence un lien causal, mais simplement de montrer le profil des bénéficiaires du dispositif et les différences entre ceux qui en bénéficient et ceux qui n'en font pas partie.

Nous présentons donc l'évolution du programme Cifre au cours du temps, en termes de volume, de concentration et de répartition des Cifres géographiquement (section 2.1). Ensuite, la description des doctorants se focalisera sur leurs caractéristiques lors de la thèse Cifre (genre, âge, salaire du doctorant), mais aussi sur les salaires et contrats de travail post-Cifre, grâce à des appariements et comparaisons avec les bases de données du CEREQ, de l'ABES et DADS. Enfin, nous présenterons le profil des entreprises qui embauchent des doctorants en Cifre, leurs caractéristiques et leur performance avant et après l'accueil d'un doctorant.

2.1 Description générale du dispositif

FIGURE 2.1 – Évolution du nombre de Cifres validées.

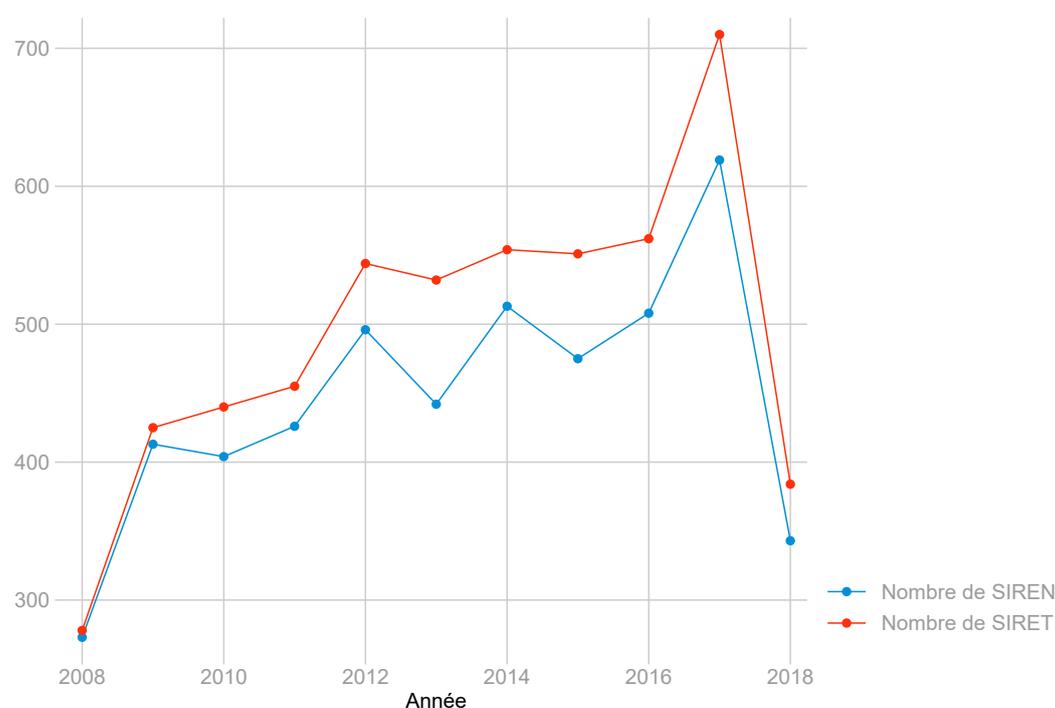


NOTES : L'année 2018 est la dernière année disponible. La baisse observée en 2018 est vraisemblablement liée au fait que l'ensemble des données pour cette année n'avait pas encore été collectée au moment de la création de la base utilisée.

Dans cette première section, nous nous intéressons à une description générale du dispositif Cifre et de son évolution. Depuis sa création, le nombre de conventions Cifre a connu une augmentation régulière. Si depuis son début, le dispositif Cifre s'est étendu rapidement, et que depuis 2003 le nombre de dossiers envoyés a connu une forte augmentation (de 500 dossiers à près de 2000 récemment), le nombre de nouvelles conventions Cifre validées depuis 2008 connaît une croissance plus faible mais constante, comme l'indique la figure 2.1. De 2008 à 2017, le nombre de Cifre par an est passé d'environ 800 à 1 400 conventions. Le nombre de contrats actifs a donc connu une augmentation logique, et en 2017 se situait à près de 3 800 conventions. La baisse observée en 2018 n'est sans doute due qu'au

manque de données pour cette année, car il s'agit de la dernière année de la base. La Cifre est donc un dispositif de plus en plus attractif pour étudiants, entreprises et laboratoires, et dont le coût pour l'État a augmenté. En effet, en se limitant uniquement aux subventions de l'ANRT et excluant les déductions du CIR, 3 800 conventions Cifre actives en 2018 représenteraient donc un budget de 53,2 millions d'euros en subventions.

FIGURE 2.2 – Évolution du nombre d'entreprises ayant recours à une convention Cifre.



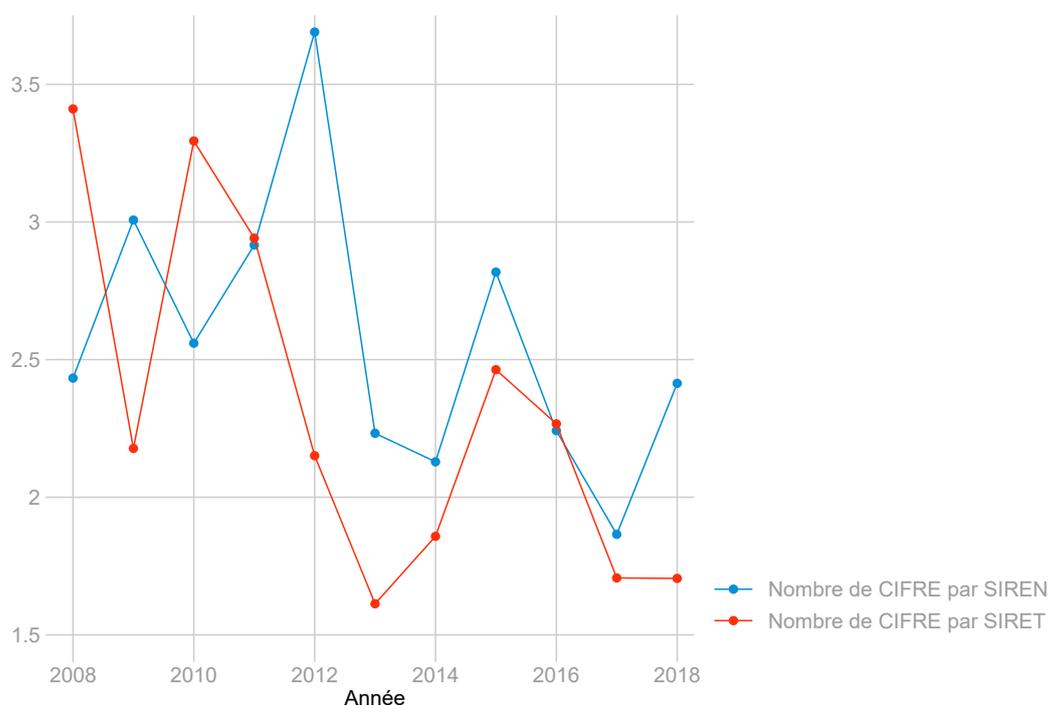
NOTES : L'année 2018 est la dernière année disponible. La baisse observée en 2018 est vraisemblablement liée au fait que l'ensemble des données pour cette année n'était pas encore été collecté au moment de la création de la base utilisée.

Deux dynamiques sous-jacentes peuvent expliquer cette croissance du côté des entreprises. Il peut s'agir d'une augmentation du nombre de conventions Cifre que signe chaque entreprise ; dans ce cas nous devrions voir le nombre de Cifres moyen par entreprise augmenter au cours du temps. Il peut aussi être le résultat d'une augmentation du nombre d'entreprises qui accueillent un doctorant en convention

Cifre.

La figure 2.2 montre l'évolution du nombre d'entreprises ayant accueilli une Cifre par année. On remarque une très nette augmentation du nombre d'entreprises signant une convention, puisque ce chiffre a plus que doublé en dix ans (l'année 2018 n'est pas complète, d'où la baisse). La figure 2.3 montre aussi que cette augmentation n'est pas liée au nombre moyen de Cifres par entité, qui ne montre pas de croissance mais au contraire une diminution sur la dernière décennie. En 2017, plus de 700 entités (niveau SIRET, c'est-à-dire établissement) ont signé une convention Cifre, et chacune accueillait en moyenne 1,7 Cifre. La valeur médiane du nombre de Cifres par SIRET est aujourd'hui d'un seul doctorant : la majorité des établissements n'hébergent qu'une Cifre.

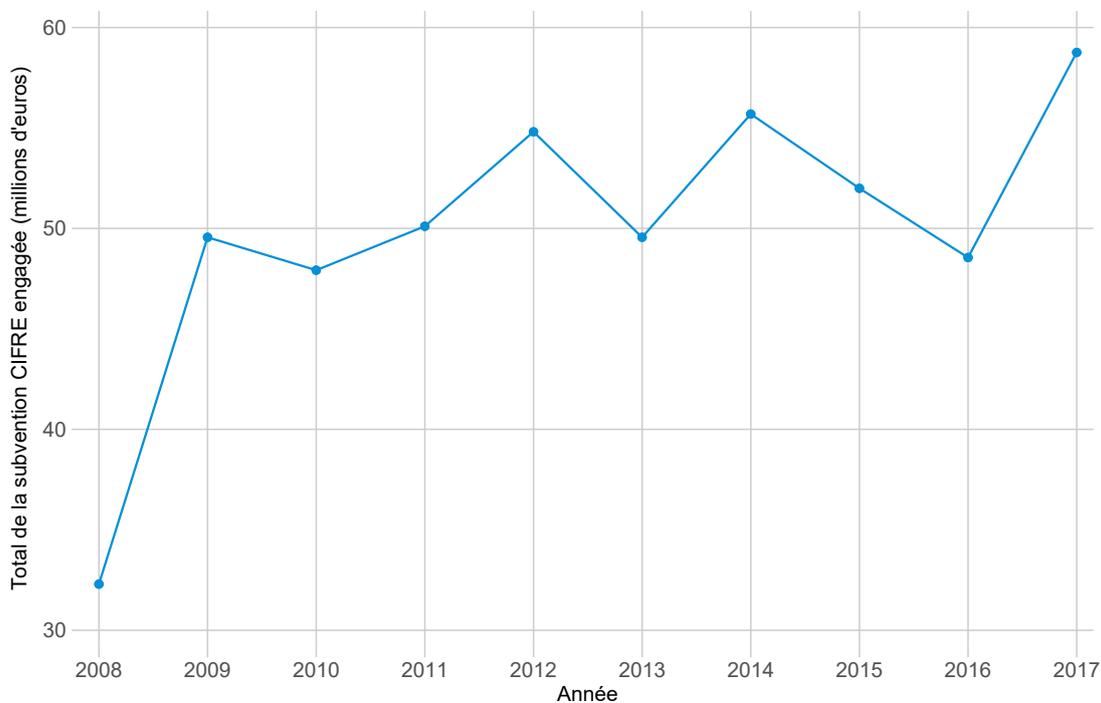
FIGURE 2.3 – Évolution du nombre de Cifres validées par entreprise.



Le graphique 2.4 représente le montant total de subvention engagé par l'ANRT par an sur la période 2008-2017. Pour déterminer le montant total engagé, on sup-

pose que chaque Cifre correspond à un montant de 42 000 euros d’engagement, c’est-à-dire 3 ans de subvention Cifre (le montant versé *in fine* peut différer, par exemple si la thèse est terminée en moins de 3 ans ou si la convention est rompue). Chaque convention est dénombrée l’année de sa date d’effet, disponible dans la base de l’ANRT. La série est relativement stable sur la période 2009-2016, le montant de subvention évoluant autour de 50 millions d’euros.

FIGURE 2.4 – Évolution de la subvention ANRT au cours de la période.



2.2 Description des doctorants faisant une Cifre

Cette section se concentre sur les caractéristiques des doctorants qui entreprennent une Cifre. Les variables qui nous intéressent incluent le salaire, la trajectoire et la stabilité dans l’emploi des Cifre, notamment en comparaison avec les docteurs non-Cifre. Nous présentons ici un certain nombre de statistiques descrip-

tives, en exploitant des données de sources diverses. Dans un premier temps, nous exploitons les données de la base Cifre de l'ANRT afin d'en tirer des statistiques sur le salaire doctoral à l'embauche, le genre, l'âge et la nationalité. Nous effectuons ensuite une opération d'appariement de ces individus avec plusieurs bases de données : celle sur les thèses effectuées de l'ABES (base sous-jacente au site theses.fr), et la base de données d'emploi DADS Postes (INSEE), afin de pouvoir comparer les doctorants Cifre aux non-Cifre, et de suivre le devenir des doctorants à la fin de leur convention. Enfin, nous exploitons aussi l'Enquête « Génération » (CEREQ) afin d'identifier les différences entre Cifre et non-Cifre en termes de salaire et de contrat de travail.

2.2.1 Statistiques de base et appariement ABES

Les statistiques de la base Cifre de l'ANRT permettent de dresser un portrait initial des caractéristiques des doctorants Cifre et de leur évolution. La figure (2.5) montre la distribution de l'âge chez les doctorants Cifre. Ces derniers ont principalement entre 23 et 28 ans à la date d'effet de leur Cifre, avec un pic à 24 - 25 ans. Cependant, on voit quelques étudiants plus âgés, ayant autour de trente ans, bénéficier chaque année de la Cifre.

Les doctorants Cifre sont en majorité français, même si depuis 2008 la proportion d'étrangers effectuant une Cifre augmente et se trouve maintenant au-dessus de 20 %. Comme nous l'avons précisé dans la première partie de ce rapport, il est nécessaire de disposer d'une autorisation de travailler sur le sol français afin d'effectuer une Cifre.

La part d'hommes dans les Cifres est nettement supérieure à celle des femmes. Cependant, la part des femmes montre une légère augmentation depuis 2008 et se situe juste en dessous de la barre des 40 % en 2018.

Cette base nous donne aussi des indications sur les disciplines que suivent les

FIGURE 2.5 – Distribution de l'âge des doctorants (conventions 2011-2018)

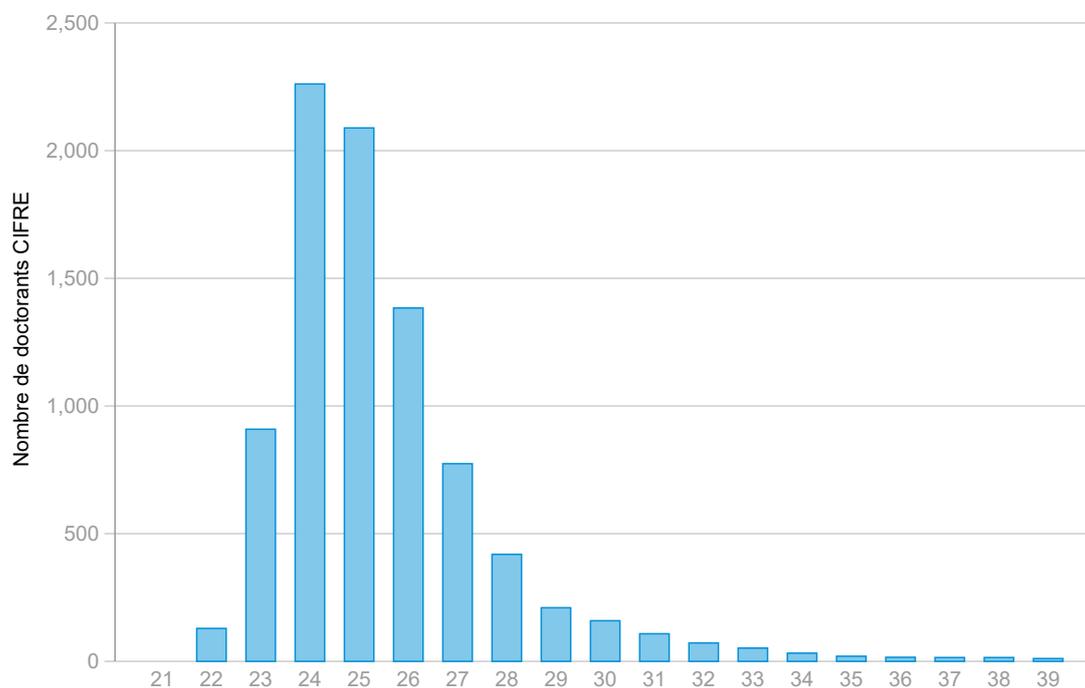


FIGURE 2.6 – Nationalité des doctorants Cifre

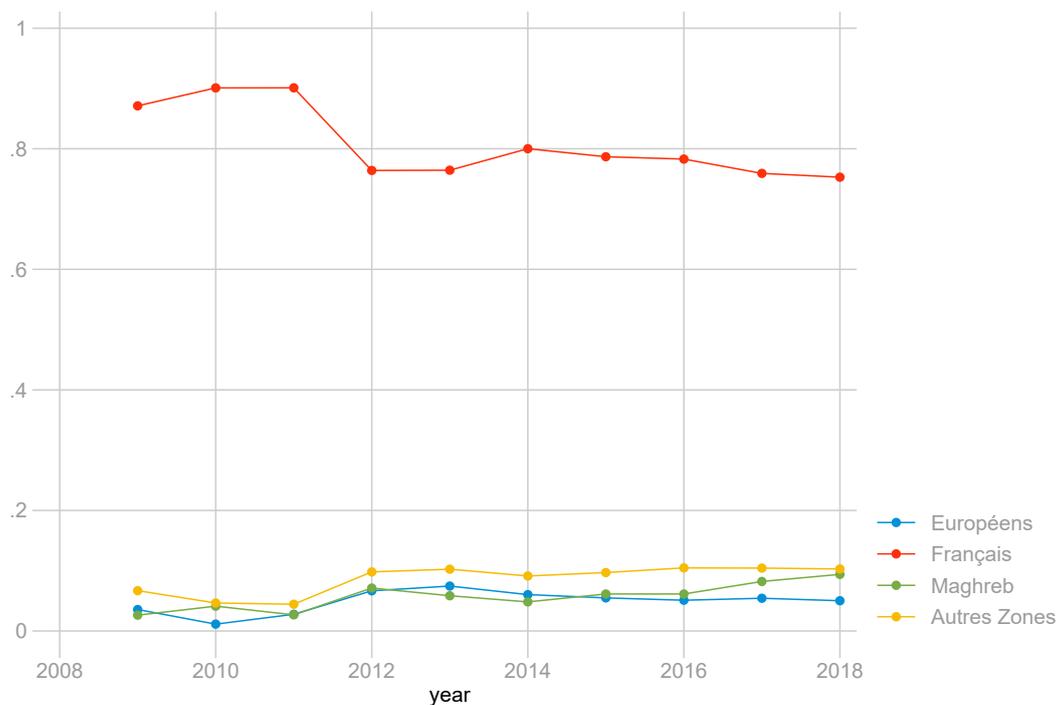
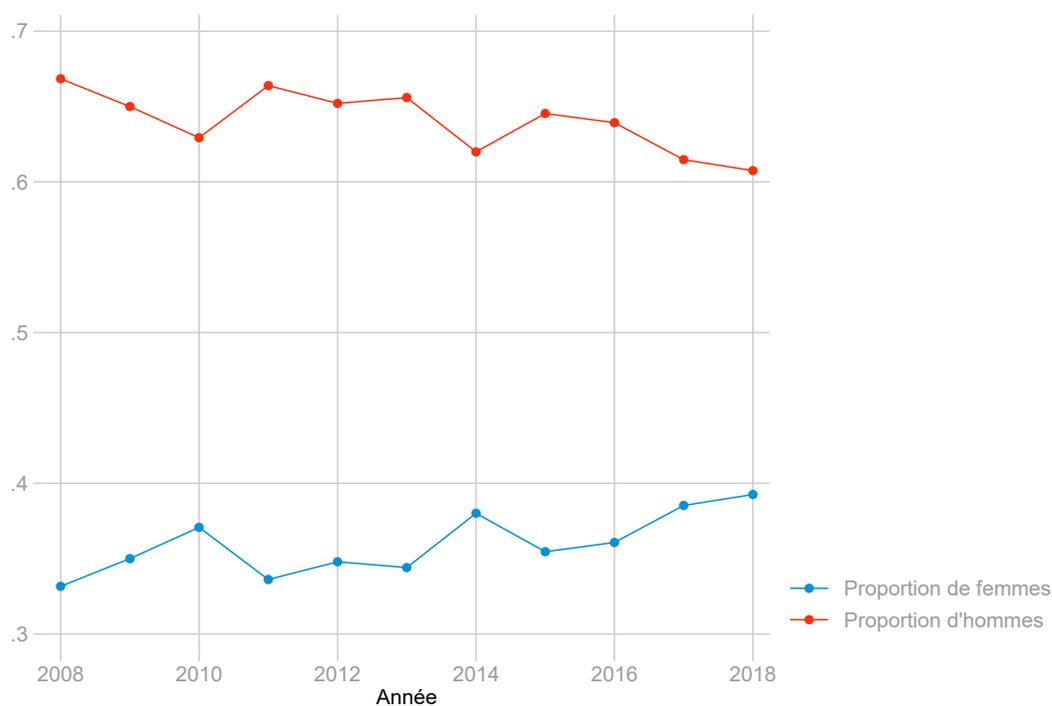


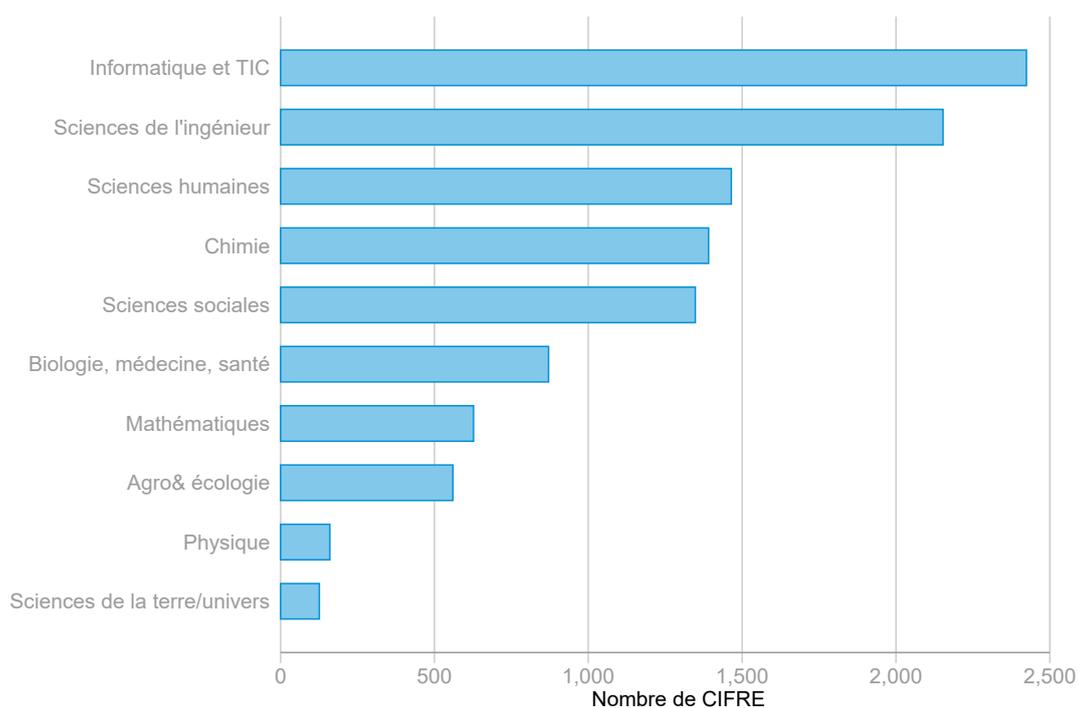
FIGURE 2.7 – Proportion de chaque genre parmi les doctorants



doctorants Cifre (figure 2.8). Cette spécification générale montre une concentration forte autour des sciences dites « STEM », avec les sciences de l'ingénieur et l'informatique en tête. Cependant, il semblerait que le caractère industriel de la recherche présent aux origines du dispositif ne soit plus aussi prégnant, comme en témoigne la proportion non-négligeable de sciences humaines et sociales, qui correspondent à des thèses en droit, sociologie, urbanisme, sciences économiques ou encore démographie.

Nous observons aussi les différents diplômes que détiennent les doctorants qui s'engagent dans une Cifre. Nous observons un grand nombre de masters recherche, suivi d'un nombre élevé de diplômes d'ingénieur. Ces diplômes sont suivis d'un plus petit nombre de masters professionnels et d'un DEA science et technologie. L'orientation vers la recherche et l'aspect « sciences dures » est donc bien présente avec la forte présence de masters recherche et d'ingénieurs dans ces diplômes. Il faut néanmoins noter que de nombreux doctorants ne déclarent pas de

FIGURE 2.8 – Disciplines des doctorants Cifre

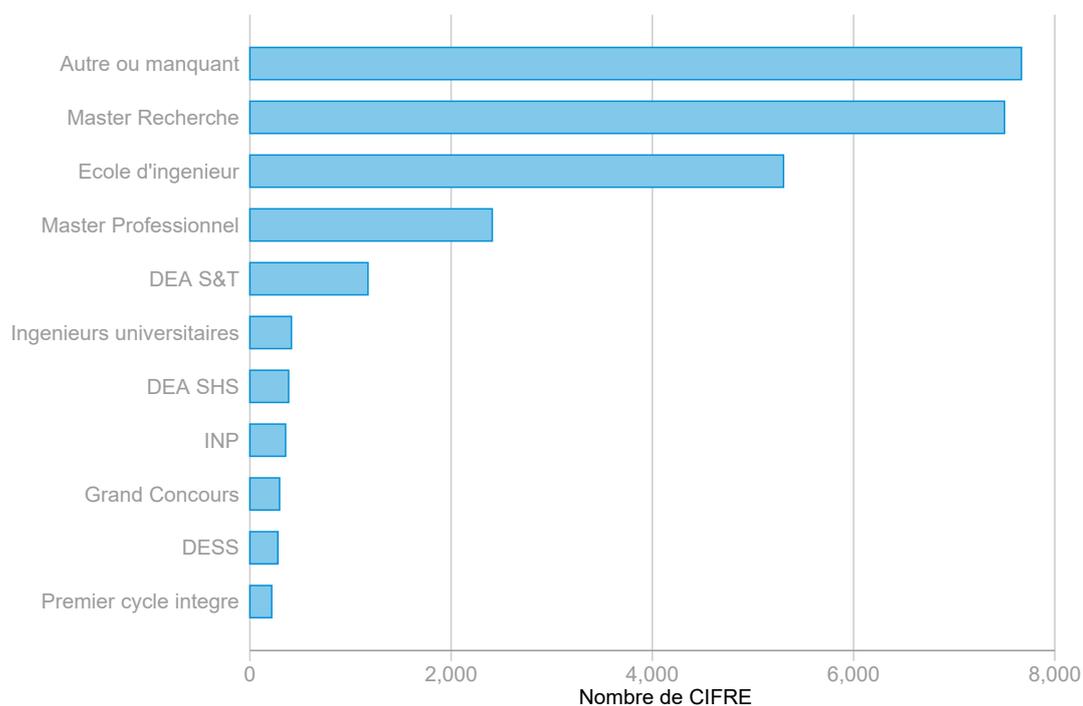


second diplôme, et qu'il est probable que les détenteurs d'un master recherche soient également diplômés, par exemple, d'écoles d'ingénieurs mais ne le mentionnent pas dans l'inscription en thèse. La forte présence des ingénieurs est en effet confirmée par l'analyse que nous menons sur l'enquête Génération, et avait déjà été identifiée dans les travaux de Margolis et Miotti (2015) sur le DJD.

Nous complétons ces statistiques grâce à un appariement avec la base de données de l'ABES, correspondant à des informations librement disponibles sur le site internet theses.fr. Les thèses validées depuis 2008 dans la base Cifre sont appariées avec les thèses de ce site, se basant sur la correspondance entre les noms et prénoms du doctorant, des directeurs de thèse, et le titre de la thèse. Nous appariions environ 9,000 thèses sur les 12,000 de la base Cifre validées sur la période 2008 - 2018. Nous y associons les thèses non-Cifre effectuées pendant la même période dans les mêmes disciplines, composant notre sous-échantillon d'analyse.

Dans la base de données de l'ABES, les thèses peuvent afficher l'un de deux

FIGURE 2.9 – Diplômes des doctorants Cifre



statuts : « soutenues » ou « en cours ». Les thèses abandonnées n’ont pas de statut spécial mais demeurent simplement « en cours » pour une période indéterminée. Nous nous intéressons aux thèses débutées entre 2008 et 2013 afin de limiter le nombre de thèses marquées « en cours » qui n’auraient pas été abandonnées (on s’intéresse donc non pas à la part des thèses finissant par être soutenues, mais à la part des thèses soutenues dans les cinq années suivant l’inscription). Sur ce sous-échantillon, nous régressons le statut de la thèse sur le fait d’avoir bénéficié d’un financement Cifre, en contrôlant pour la discipline (spécification plus large) et l’établissement de la thèse. Le tableau 2.1 montre les résultats de cette régression, et la présence d’un effet positif de la convention Cifre sur le taux de soutenance. Cet effet diminue mais persiste lorsque nous contrôlons pour l’établissement de thèse et de la discipline. Nous ne sommes cependant pas en mesure de contrôler pour l’âge, ni pour l’année de début de la thèse, ni pour le sexe, car ces informations ne sont pas fournies dans la base de données theses.fr.

TABLEAU 2.1 – Régressions linéaire du taux de soutenance sur le statut Cifre

	(1)	(2)	(3)
Convention CIFRE	0.0995*** (21.30)	0.0344*** (7.95)	0.0176*** (4.12)
Constante	0.671*** (449.04)	0.675*** (537.97)	0.677*** (549.98)
R^2	0.003	0.273	0.298
Etablissement		✓	✓
Discipline			✓
Observations	93339	93314	93314

t score entre parenthèses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

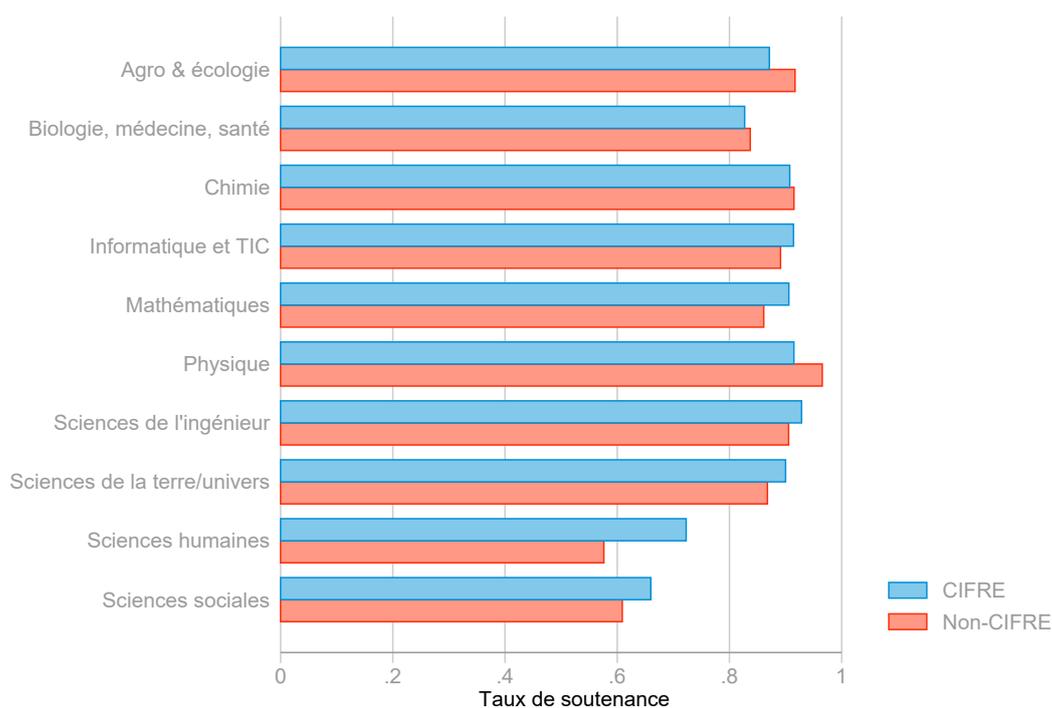
SOURCES : Bases Cifre, theses.fr.

La figure 2.10 montre les taux de soutenance par discipline et par statut Cifre. Il est utile de noter que dans son enquête, l'ANRT trouvait un taux de soutenance qui avoisinait les 98 %. La différence avec notre résultat souligne sans doute une latence dans la mise à jour de la base ; cependant, nous pensons que cet effet n'est pas hétérogène entre les Cifre et non-Cifre, et que les résultats du tableau 2.1 ne sont pas biaisés par cette latence. Nous observons que même si la régression indique l'existence d'un effet positif pour les doctorants Cifre, il existe beaucoup d'hétérogénéité entre disciplines dans cet effet. Il est en effet très fort dans les sciences humaines et sociales, l'ingénierie et les TIC, mais ne l'est pas pour autant en agronomie et écologie, physique ou chimie.

2.2.2 Résultats issus de l'enquête Génération

Dans le cadre de notre évaluation du dispositif Cifre, nous utilisons aussi les données fournies par l'enquête Génération, dont le questionnaire recense les doctorants ayant accompli un doctorat Cifre et leurs séquence d'emploi les trois an-

FIGURE 2.10 – Taux de soutenance par discipline theses.fr (2011-2018)



nées qui ont suivi la fin de leur scolarité. Ces données nous permettent de les comparer à différents groupes contrefactuels, comme des doctorants non-Cifre dans les mêmes disciplines ou bien d'autres étudiants ayant suivi une formation similaire, et d'observer leurs débouchés en terme de salaire et de contrats de travail. Nous utilisons les enquêtes pour les années 2010 et 2013 car elles disposent d'un échantillon conséquent sur les Cifre ainsi que sur le type de formation suivi, comme celui de l'école d'ingénieur.

Présentation des données :

L'Enquête Génération est une base de données produite par le Centre d'études et de recherche sur la qualification (CEREQ) depuis 1997. Elle consiste en un suivi très riche des caractéristiques des jeunes sortant du système scolaire pendant une durée de 3 ans suivant leur dernier diplôme, permettant d'observer leur insertion professionnelle. Les Enquêtes Génération disposent d'un module thèse qui renseigne les spécificités des doctorants durant leur thèse, ce qui permet de suivre les

doctorants qui ont bénéficié d'une Cifre. Les doctorants sont aussi sur-échantillonnés par ces enquêtes, ce qui nous permet d'exploiter un nombre significatif de Cifres dans chaque vague; nous corrigeons nos résultats avec les pondérations fournies. L'analyse de ces données s'effectue sur la base de plusieurs échantillons et variables, que nous définissons ci-dessous.

Échantillons :

- Étudiant en STEM : cette classification regroupe des étudiants en Sciences Techniques, Ingénierie et Mathématiques (STEM selon la dénomination anglaise). Ceci regroupe donc les étudiants dont le plus haut diplôme détenu est un doctorat (hors santé) en Sciences, Technologies et Mathématiques; une école d'ingénieur; une formation Bac+5 en Sciences, Technologies et Mathématiques; une formation Bac+4 en Sciences, Technologies et Mathématiques; ou un M2 Sciences, Technologies et Mathématiques (déclinaisons de la variable PHD)
- Ingénieur : Les ingénieurs sont définis comme ceux dont le plus haut diplôme provient d'une école d'ingénieur (PHD = 16 ou PHINSEE = 170), dont la classe de sortie est une école d'ingénieur (NISOR = 16) ou dont le niveau le plus élevé avant la fin des études en 2013 était une école d'ingénieur (Q39H = 24)
- Docteur : ayant effectué comme dernier diplôme un doctorat (sortant de niveau thèse, hors santé, MODTH = 1)
- Cifre : docteurs ayant bénéficié du dispositif Cifre (TH4 =5) et ayant soutenu leur thèse.

Variables :

- CDI : variable binaire, indique si le contrat de travail est un CDI ou un contrat de fonctionnaire. Valeur de 0 pour les individus en non-emploi.

- Salaire : variable correspondant au salaire en fin de séquence d'emploi ; si la fin n'est pas disponible, le salaire de début est utilisé.
- Secteur : secteur professionnel NAF de l'entreprise, 10 possibilités¹.
- CSP – père : classe socioprofessionnelle du père, 9 possibilités
- PHNSF : spécialité du plus haut niveau de formation, 86 possibilités

Nombre d'observations :

- Etudiants en STEM : 6 176
- Ingénieurs (y compris docteurs) : 2 503
- Cifres (STEM) : 342
- Docteurs : 2 540
- CifreS – ingénieurs : 158
- Docteurs – ingénieurs (y compris Cifres) : 705
- Cifres – ingénieurs : 158
- Ingénieurs non docteurs : 1 798

Il est important de noter que si la variable CDI concerne aussi à l'origine les individus en non-emploi, les contrôles comme le secteur d'activité restreignent l'échantillon à ceux en situation d'emploi. Les contrôles de base sont le genre, une fonction quadratique de l'âge, des effets fixes de génération.

Statistiques descriptives :

Nous nous concentrons dans cette section sur quatre comparaisons qui permettent d'établir les différences entre les Cifres et non-Cifres au sein de plusieurs sous échantillons. Tout d'abord, nous comparons les Cifres en STEM avec les docteurs en STEM, puis les Cifres en STEM avec les non-docteurs en STEM. Nous comparons ensuite les Cifres ingénieurs avec les ingénieurs non-docteurs. Enfin, nous comparons les Cifres ingénieurs et les docteurs ingénieurs.

1. Agriculture, Manufacture, Construction, Commerce, Informatique et Communication, Finance, Immobilier, Métiers Spécialisés, Public et Social.

TABLEAU 2.2 – Statistiques Descriptives - 2013, 2010

	CIFRE	Doc. non CIFRE	Différence
Secteur Privé	0.70	0.43	0.26***
CDI	0.49	0.31	0.18***
Salaire	2583.11	2303.35	279.77***
Satisfaction	0.79	0.69	0.10***
Longueur max. d'emploi	27.23	23.75	3.48***
Changement de situation	1.61	1.82	-0.21***
Âge	27.92	28.04	-0.12
Genre = M	0.67	0.60	0.07*
Père ouvrier/employé	0.30	0.31	-0.01
Père cadre	0.51	0.55	-0.04
Mention Très Bien	0.11	0.12	-0.01
Nb observations	329.00	2198.00	

Échantillon composé de Docteurs en STEM

Les Cifre en STEM ont donc une probabilité plus élevée d'avoir un emploi dans le secteur privé et d'avoir un salaire plus élevé que l'ensemble des docteurs non-Cifre qui sont en STEM (Tableau 2.2), à plus de 2,500 euros pour les Cifres. Ils ont aussi une différence significative au niveau de la stabilité de l'emploi, ayant un taux de CDI plus élevé, des séquences d'emploi plus longues que les autres docteurs et moins de changements de situation. Ils sont aussi plus satisfaits de leur travail et sont plus souvent des hommes que les non-Cifres.

Les docteurs Cifres travaillent en revanche moins souvent dans le privé que les non-docteurs en STEM (Tableau 2.3). Ces derniers ont un salaire plus faible que les Cifres (une différence de 437 euros par mois), ce qui s'explique aussi en partie par la forte différence d'âge (les Cifres sont en moyenne 3.7 ans plus âgés). Dans ces deux échantillons, on ne relève pas de différences dans la classe socio-professionnelle du père ou dans le taux de mention Très Bien.

Les tableaux 2.4 et 2.5 se concentrent sur l'univers des ingénieurs, en regardant respectivement les différences avec les ingénieurs non-docteurs et docteurs. Dans ces deux catégories, la différence de salaire est significative et favorable aux

TABLEAU 2.3 – Statistiques Descriptives - 2013, 2010

	CIFRE	Non Docteur	Différence
Secteur Privé	0.70	0.77	-0.07**
CDI	0.49	0.44	0.05
Salaire	2583.11	2140.55	442.57***
Satisfaction	0.79	0.77	0.02
Longueur max. d'emploi	27.23	23.98	3.25***
Changement de situation	1.61	1.77	-0.16**
Âge	27.92	24.19	3.74***
Genre = M	0.67	0.61	0.06*
Père ouvrier/employé	0.30	0.32	-0.01
Père cadre	0.51	0.50	0.01
Mention Très Bien	0.11	0.09	0.02
Nb observations	329.00	3632.00	

Échantillon composé de Docteurs et Bac+4 ou plus, en STEM

Cifres. Les Cifres connaissent aussi plus de stabilité avec des périodes d'emploi plus longues. Les ingénieurs non-docteurs travaillent bien plus dans le secteur privé que les Cifres, à hauteur de 84 %. Ces caractéristiques ne s'étendent pas aux docteurs ingénieurs, qui eux travaillent moins dans le privé. La probabilité d'avoir un emploi stable (CDI ou fonctionnaire) est plus élevée pour ces Cifres ingénieurs que pour les docteurs ingénieurs non Cifres, mais à égalité avec les ingénieurs non-docteurs, autour de 55 %.

Le portrait global que l'on peut dresser sur les diplômés Cifres en STEM est celui d'une population à mi-chemin entre les docteurs traditionnels et les STEM non-docteurs dans l'insertion dans le secteur privé, qui bénéficie de taux très élevés d'emplois stables, de séquences d'emploi plus longues, et d'une prime de salaire par rapport à toutes les autres populations, y compris pour la sous-population d'ingénieurs. On ne retrouve pas de différences dans la catégorie socio-professionnelle du père, ou la proportion de mentions "Très Bien" au Baccalauréat.

TABLEAU 2.4 – Statistiques Descriptives - 2013, 2010

	CIFRE Ingé.	Non Docteur Ingé.	Différence
Secteur Privé	0.71	0.84	-0.12***
CDI	0.55	0.55	-0.00
Salaire	2646.40	2373.47	272.93***
Satisfaction	0.85	0.81	0.03
Longeur max. d'emploi	28.56	25.96	2.60***
Changement de situation	1.60	1.57	0.03
Âge	27.65	23.83	3.83***
Genre = M	0.72	0.68	0.05
Père ouvrier/employé	0.26	0.25	0.00
Père cadre	0.54	0.55	-0.01
Mention Très Bien	0.16	0.16	-0.00
Nb observations	156.00	1799.00	

Échantillon composé d'Ingénieurs et Docteurs Ingénieurs

TABLEAU 2.5 – Statistiques Descriptives - 2013, 2010

	CIFRE Ingé.	Doc. non CIFRE Ingé.	Différence
Secteur Privé	0.71	0.56	0.15***
CDI	0.55	0.41	0.14**
Salaire	2646.40	2454.36	192.04**
Satisfaction	0.85	0.78	0.06
Longeur max. d'emploi	28.56	25.30	3.25***
Changement de situation	1.60	1.77	-0.17*
Âge	27.65	27.73	-0.08
Genre = M	0.72	0.66	0.06
Père ouvrier/employé	0.26	0.26	-0.00
Père cadre	0.54	0.62	-0.09
Mention Très Bien	0.16	0.19	-0.03
Nb observations	156.00	542.00	

Échantillon composé de Docteurs Ingénieurs

Régressions :

Nous effectuons ensuite une série de régressions afin d'étudier deux variables dépendantes d'intérêt : le salaire (log du salaire) et le taux d'emplois stables (variable CDI). Nous régressons ces variables sur le statut Cifre, mais aussi sur les statuts d'ingénieur et de docteurs selon l'échantillon. Afin d'affiner nos résultats nous contrôlons progressivement pour plusieurs variables observables : le secteur de l'entreprise, la CSP du père, le genre, l'âge (fonction quadratique), la génération (2010 ou 2013) et la spécialité du plus haut diplôme (spécialité NSF).

Notre équation d'estimation pour l'échantillon le plus large est la suivante (avec pondération) :

$$Y_i = \beta_1 Cifre_i + \beta_2 Docteur_i + \beta_3 Ingenieur_i + \beta_4 Secteur_i + \beta_5 CSP_i + \beta_6 NSF_i + \beta_7 Genre_i + \beta_8 Age_i + \beta_9 Age_i^2 + \beta_{10} Generation_i + \epsilon_i \quad (2.1)$$

Premier échantillon : Études en STEM

Dans cet échantillon, nous régressons le statut de CDI et le salaire sur le statut de docteur, le fait d'avoir fait des études d'ingénieur, et le fait d'avoir bénéficié d'un financement Cifre pour sa thèse. Nous observons qu'il existe un effet très fort des études d'ingénieur, ainsi qu'un effet indépendant et de magnitude supérieure d'avoir effectué une Cifre, sur la probabilité d'avoir un emploi stable : environ onze points de pourcentage de l'ingénieur et seize pour la Cifre, sachant que la moyenne pour la catégorie de référence est de 39.6 %. Il n'y a pas d'effet significatif du doctorat, qui a même un coefficient associé négatif pour la première spécification.

En ce qui concerne le salaire, nous retrouvons un effet fort des études d'ingénieur et du doctorat, avec une augmentation de salaire plus élevée pour les ingénieurs non docteurs que les docteurs non ingénieurs. Il n'y a pas d'impact significatif du Cifre sur le salaire une fois que l'on contrôle pour le secteur de travail, et

la magnitude est très faible.

TABEAU 2.6 – Régression du CDI sur statut Docteur et CIFRE - STEM

	(1)	(2)	(3)	(4)
Non-ingé non-docteur	<i>Catégorie de référence</i>			
Docteur	-0.0124 (0.0279)	0.0103 (0.0331)	0.0275 (0.0370)	0.0298 (0.0379)
CIFRE	0.198*** (0.0497)	0.167*** (0.0478)	0.151*** (0.0484)	0.160*** (0.0488)
Études d'ingénieur	0.165*** (0.0204)	0.155*** (0.0220)	0.108*** (0.0239)	0.111*** (0.0250)
R^2	0.145	0.194	0.222	0.226
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Nombre de CIFRE				329
Observations	6159	6152	5380	5001

Échantillon composé de Docteurs et Bac+4 ou plus, en STEM

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

TABEAU 2.7 – Régression du salaire sur statut CIFRE et Docteur - STEM

	(1)	(2)	(3)	(4)
Non-ingé non-docteur	<i>Catégorie de référence</i>			
Docteur	0.0408** (0.0203)	0.0928*** (0.0258)	0.109*** (0.0265)	0.109*** (0.0277)
CIFRE	0.0826*** (0.0254)	0.0559** (0.0271)	0.0286 (0.0261)	0.0196 (0.0273)
Études d'ingénieur	0.182*** (0.0135)	0.165*** (0.0128)	0.135*** (0.0129)	0.136*** (0.0131)
R^2	0.125	0.199	0.240	0.258
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	5224	5215	5215	4851

Échantillon composé de Docteurs et Bac+4 ou plus, en STEM

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

Deuxième échantillon : Études d'ingénieur

Nous régressons maintenant le statut d'emploi stable et le salaire sur le statut de docteur et le fait d'avoir pris part à une Cifre sur un sous-échantillon qui comprend seulement ceux ayant effectué des études d'ingénieur. Nous observons que la convention Cifre n'est plus significative dans la régression du CDI dans la dernière spécification, ce qui est sans doute l'effet d'une restriction de l'échantillon. Cependant, la magnitude du coefficient est importante et doit être notée. Dans le cas du salaire, il y a un effet de la Cifre et du doctorat sur le salaire des ingénieurs, tout deux de 6 points de pourcentage.

TABEAU 2.8 – Régression du CDI sur statut Docteur et CIFRE - Ingénieurs

	(1)	(2)	(3)	(4)
Ingénieur non-docteur	<i>Catégorie de référence</i>			
Docteur	-0.0291 (0.0624)	0.00594 (0.0711)	0.0756 (0.0734)	0.105 (0.0727)
CIFRE	0.198** (0.0770)	0.192** (0.0779)	0.142* (0.0771)	0.109 (0.0766)
Observations	2497	2489	2292	2164
R^2	0.118	0.163	0.191	0.196
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓

Échantillon composé de Docteurs et Bac+4 ou plus ingénieurs
 Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

TABEAU 2.9 – Régression du salaire sur statut Docteur et CIFRE - Ingénieurs

	(1)	(2)	(3)	(4)
Ingénieur non-docteur	<i>Catégorie de référence</i>			
Docteur	0.0177 (0.0261)	0.0333 (0.0331)	0.0715* (0.0371)	0.0674* (0.0391)
CIFRE	0.100** (0.0403)	0.102*** (0.0397)	0.0825** (0.0390)	0.0677* (0.0394)
R^2	0.028	0.136	0.186	0.205
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	2236	2228	2228	2104

Échantillon composé de Docteurs et Bac+4 ou plus ingénieurs
 Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

Troisième échantillon : Docteurs

Nous nous concentrons ensuite sur les docteurs et régressons le statut de CDI et le salaire sur le fait d'avoir pris part à une Cifre et d'avoir effectué des études d'ingénieur, en interagissant ces termes. Nous observons un effet positif de la Cifre sur la probabilité d'avoir un CDI, ainsi que d'avoir effectué des études d'ingénieur; l'interaction est significative mais inférieure à l'addition des effets indépendants. Les résultats sont similaires pour l'effet sur le salaire des doctorants, avec un effet Cifre et ingénieur indépendamment significatifs et un effet sur le salaire d'effectuer une Cifre et d'avoir été ingénieur inférieur à l'addition des deux primes indépendantes. Globalement, les interactions sont plus petites que la somme des effets dans les deux cas, et l'effet sur le salaire est environ deux fois inférieur à celui sur le contrat de travail.

TABLEAU 2.10 – Régression du CDI sur statut CIFRE et Ingénieur - Docteurs

	(1)	(2)	(3)	(4)
Docteur non-ingénieur	<i>Catégorie de référence</i>			
Non CIFRE × Ingénieur	0.112*** (0.0389)	0.0589 (0.0367)	0.0333 (0.0351)	0.0689** (0.0348)
CIFRE × Non Ingénieur	0.244*** (0.0645)	0.182*** (0.0598)	0.160*** (0.0576)	0.208*** (0.0554)
CIFRE × Ingénieur	0.320*** (0.0672)	0.276*** (0.0633)	0.202*** (0.0693)	0.223*** (0.0671)
R^2	0.130	0.200	0.250	0.268
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	2521	2518	2197	2026

Échantillon composé de Docteurs STEM

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

TABLEAU 2.11 – Régression du salaire sur statut CIFRE et Ingénieur - Docteurs

	(1)	(2)	(3)	(4)
Docteur non-ingénieur		<i>Catégorie de référence</i>		
Non CIFRE × Ingénieur	0.0706*** (0.0248)	0.0482** (0.0208)	0.0321 (0.0209)	0.0479*** (0.0185)
CIFRE × Non Ingénieur	0.128*** (0.0280)	0.0990*** (0.0269)	0.0595** (0.0266)	0.0606** (0.0286)
CIFRE × Ingénieur	0.175*** (0.0356)	0.154*** (0.0337)	0.114*** (0.0343)	0.117*** (0.0336)
R^2	0.048	0.098	0.146	0.171
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	2145	2142	2141	1977

Échantillon composé de Docteurs STEM

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

Quatrième échantillon : Docteurs ingénieurs

Enfin, nous nous intéressons au sous-échantillon le plus petit, celui des docteurs ingénieurs, toujours pour comparer les Cifres et non-Cifres. Nous régressons le statut de CDI et le salaire sur le fait d'avoir pris part à une Cifre pour seulement ceux ayant effectué des études d'ingénieur et ont un doctorat. La variable Cifre a un effet sur le CDI, très proche en magnitude de celui pour l'ensemble des STEM à 17 points de pourcentage, avec une moyenne de 46.2 % pour la référence non-Cifre. Nous trouvons également un effet sur le salaire, qui est cependant plus faible dans sa magnitude que l'effet sur les CDI, autour de 7 points de pourcentage.

Le portrait qui se détache du dispositif Cifre de cette analyse est celui d'un dispositif qui mène dans la majorité des cas à une augmentation de stabilité dans l'emploi, ceci avec une grande dimension des contrôles qui nuisent sans doute à la signification statistique de nos échantillons. Pour les salaires, l'effet est plus faible dans sa magnitude que celui sur le contrat de travail, mais ne laisse pas

TABEAU 2.12 – Régression du CDI sur statut CIFRE - Docteurs Ingénieurs

	(1)	(2)	(3)	(4)
Docteur ingé non CIFRE	<i>Catégorie de référence</i>			
CIFRE	0.208*** (0.0749)	0.244*** (0.0667)	0.194*** (0.0711)	0.172** (0.0685)
R^2	0.170	0.236	0.257	0.292
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	698	695	629	586

Échantillon composé de Docteurs ingénieurs

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

TABEAU 2.13 – Régression du salaire sur statut CIFRE - Docteurs Ingénieurs

	(1)	(2)	(3)	(4)
Docteur ingé non CIFRE	<i>Catégorie de référence</i>			
CIFRE	0.0918** (0.0393)	0.0917** (0.0356)	0.0800** (0.0359)	0.0706** (0.0328)
R^2	0.086	0.191	0.233	0.250
Contrôle de Base	✓	✓	✓	✓
PHNSF		✓	✓	✓
Secteur			✓	✓
CSP pere				✓
Observations	619	615	614	573

Échantillon composé de Docteurs ingénieurs

Contrôle de base : genre, fonction quadratique de l'âge, effet fixe génération

de doute au sujet de sa significativité. Il faut noter que la petite taille de l'échantillon pour les doctorants en Cifre est un facteur qui biaiserait nos résultats vers la non-significativité - ce qui fait donc ressortir encore plus les résultats significatifs et avec un effet non-négligeable que nous obtenons sur le salaire et la probabilité d'avoir un emploi stable.

2.2.3 Résultats issus de l'appariement DADS

Cette section décrit l'appariement entre les données de la base Cifre de l'ANRT et les données administratives d'emploi DADS. La base DADS Postes recense tous les emplois pour chaque année. Notre stratégie d'appariement a donc été de retrouver chaque doctorant Cifre dans la base DADS lors de sa première année complète de Cifre, puis de le retracer dans les années suivantes de la base DADS jusqu'à après la fin de sa thèse, lors de sa première année complète d'emploi. Par exemple, un Cifre qui commence sa thèse en 2009 sera identifié dans le DADS 2010, 2011, 2012 et 2013. Les données DADS étant disponibles jusqu'en 2015, notre appariement concerne l'échantillon d'individus ayant commencé leur thèse Cifre entre 2008 et 2011 (inclus). Cela représente environ 4 000 conventions Cifre signées.

Dans un premier temps nous créons toutes les paires possibles, selon le SIRET de l'établissement et le sexe du doctorant. Pour l'année 2011, nous utilisons aussi l'âge, qui est manquant pour les années 2008 – 2010. Nous éliminons ensuite les paires « impossibles » pour lesquelles l'âge est trop élevé (au dessus de 31 ans dans la base DADS, afin d'être conservateurs) ou bien pour lesquelles le salaire du salarié est inférieur au salaire du doctorant – en effet, le salaire brut DADS chiffre aussi des avantages non comptés dans le salaire initial du doctorant. Enfin, nous éliminons les doublons, afin de n'avoir qu'un seul emploi par doctorant et doctorant par emploi.

A ce stade, nous disposons donc de quelques centaines de paires pour chaque

année, pour un millier de conventions Cifre validées par année environ. Nous retrouvons ensuite dans les DADS des années suivantes ces individus pendant trois ans. Lors de cette étape, il y a un appauvrissement supplémentaire de notre échantillon, car certains individus ne sont plus retrouvés d'une année à l'autre. La taille de l'échantillon final est de 583 individus; il s'agit donc d'un échantillon assez faible avec une portée statistique limitée, qui n'apporte pas d'améliorations considérables relativement aux analyses effectuées sur l'enquête Génération. Le nombre d'identifications serait encore plus faible si, quand nous retrouvons les doctorants d'une année DADS à une autre, nous n'autorisons pas un écart d'un an pour l'âge.

L'attrition dont a souffert notre échantillon peut être expliquée par plusieurs facteurs. Tout d'abord, le manque d'information concernant l'âge des doctorants nous empêche d'utiliser cette variable, ce qui rend l'appariement plus complexe. Notre stratégie souffre aussi du manque d'information sur le contrat de travail du doctorant, qui est rarement précisé dans la base de Cifre (variable statut). A cause de cela, il a été nécessaire d'exclure dans un premier temps les appariements avec des CDI, car autoriser des contrats de Cifre en CDI menait à une proportion irréaliste des doctorants en CDI durant la Cifre. Cependant, si les exclure de notre analyse préservait notre confiance en la validité de notre appariement, cela biaisait notre échantillon. Nous avons donc effectué l'appariement une première fois sans CDI, puis y avons ajouté le résultat d'un appariement avec CDI seulement pour les doctorants sans paire. Cela nous permet aussi d'obtenir un échantillon de plus grande taille, comme le détaille le tableau ci-dessous.

Notre échantillon souffre aussi de deux autres biais probables, celui des grandes entreprises et des doctorants qui restent dans leur entreprise après la Cifre. Ce sont les individus que nous retrouvons le plus facilement, avec dans notre échantillon des taux de rétention dans l'entreprise de Cifre et de conventions avec des grandes entreprises plus élevés que les statistiques communiquées par l'ANRT. Nous sur-échantillonons donc probablement les grandes entreprises. Il paraît donc évident

que nos résultats ne sont pas d'une grande fiabilité statistique, étant donné les nombreux biais et le faible échantillon dont nous disposons. Les résultats de l'enquête Génération y seront complémentaires et plus fiables.

TABLEAU 2.14 – Tableau décrivant l'évolution de la taille de l'échantillon lors de l'appariement

Étape de l'appariement	2008	2009	2010	2011	Total
Échantillon de base	852	1182	1152	1236	4422
Échantillon valable	685	1055	1060	1161	3961
Post-appariement DADS	184	266	178	150	778
Individus retrouvés trois ans après	89	120	74	85	368
Total individus après ajout des CDI	204	264	181	131	780

SOURCES : Bases Cifre et DADS Postes

Nous effectuons une brève analyse de l'échantillon résultant de cet appariement. Comme le montre le tableau 2.15, le salaire médian que nous retrouvons est de 39 530 euros, avec comme équivalent de salaire moyen 41 010 euros. C'est une valeur proche de ce que retrouve l'enquête de l'ANRT de 2016, qui trouve 38 782 euros comme salaire moyen. Nous observons aussi une proportion de CDI très élevée, comme dans l'enquête Génération (50 % des individus, mais incluant ceux qui ne sont pas en emploi), confirmant notre analyse que les Cifre s'intègrent très bien dans le marché du travail. Nous observons aussi que les âges observés sont concordants avec notre analyse de l'enquête Génération, qui ont presque 28 ans lorsqu'ils achèvent leurs études; alors que nous trouvons 29 ans lors de la première année de travail.

Dans notre échantillon, nous observons deux groupes : les étudiants qui restent dans la boîte d'origine et ceux qui partent. Les trajectoires sont sans doute différentes entre ces deux groupes, avec des profils différents; il est possible que la décision de partir ou de rester soit déterminante sur le contrat de travail ou le

TABEAU 2.15 – Statistiques descriptives des doctorants post-Cifre

Equivalent du salaire moyen	41010.75
Equivalent du salaire médian	39530
CDI	0.78
Même entreprise	0.65
Proportion d'ingénieurs	0.57
Proportion de techniciens	0.29
Education - recherche	0.74
Âge moyen	29.5
Genre = M	0.68
Observations	780

SOURCES : Bases Cifre, DADS Postes

salaire. Nous régressons donc l'équivalent annuel du salaire gagné et le contrat de travail sur l'année, le salaire au début de la Cifre, le statut d'ingénieur (PCS 38), la tranche d'effectif (treffen) et le fait d'être resté dans la même entreprise (contrôlant aussi pour l'âge (fonction quadratique) et le genre. Ces régressions confirment qu'il y a bien un effet sur le contrat de travail et sur le salaire qui est positif de rester dans la même entreprise. Il reste présent même en contrôlant pour le secteur NAF de l'entreprise et le fait d'être ingénieur pour le CDI, mais pas pour le salaire, pour lequel l'effet disparaît avec l'inclusion du secteur NAF (18 catégories). Cet effet positif est potentiellement l'expression d'une ancienneté dans la carrière dans l'entreprise, ou du fait que les doctorants qui acceptent de rester ne restent que pour un CDI et un salaire élevé. Il peut aussi s'agir d'un filtre, faisant que les doctorants les plus performants sont ceux qui sont embauchés ensuite ; et cette performance est récompensée par un salaire plus élevé et un contrat de travail permanent.

TABLEAU 2.16 – Régressions linéaire du CDI sur le le fait de rester dans la même entreprise

	(1)	(2)	(3)	(4)	(5)
Même entreprise	0.273*** (8.20)	0.286*** (8.73)	0.285*** (8.82)	0.269*** (9.07)	0.112*** (3.89)
Âge	0.349*** (3.41)	0.335** (3.09)	0.336** (3.27)	0.329** (3.22)	0.314* (2.39)
Genre = M	0.0493 (1.58)	0.0420 (1.36)	0.0385 (1.27)	0.0307 (1.06)	0.0259 (1.01)
Salaire du doctorant			0.0000138*** (3.75)	0.0000101** (2.88)	0.00000317 (1.08)
Ingénieur				0.205*** (7.71)	0.0340 (1.33)
Tranche de l'effectif		✓	✓	✓	✓
Secteur NAF					✓
R^2	0.123	0.157	0.172	0.2727	0.466
Observations	780	780	780	780	780

t score entre parenthèses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

SOURCES : Bases Cifre, DADS Postes

TABLEAU 2.17 – Régressions linéaire du salaire sur le fait de rester dans la même entreprise

	(1)	(2)	(3)	(4)	(5)
Même entreprise	0.0933*** (4.27)	0.0839*** (3.73)	0.0826*** (3.82)	0.0637** (3.28)	0.0209 (1.06)
Âge	-0.0130 (-0.16)	-0.00904 (-0.10)	-0.00701 (-0.08)	-0.0153 (-0.17)	-0.00729 (-0.09)
Genre = M	0.0594* (2.56)	0.0658** (2.79)	0.0609** (2.70)	0.0517* (2.50)	0.0412 (1.96)
Salaire du doctorant			0.0000196*** (7.83)	0.0000153*** (6.42)	0.0000152*** (5.98)
Ingénieur				0.241*** (13.56)	0.210*** (10.78)
Tranche de l'effectif		✓	✓	✓	✓
Secteur NAF					✓
R^2	0.037	0.059	0.121	0.278	0.305
Observations	780	780	780	780	780

t score entre parenthèses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

SOURCES : Bases Cifre, DADS Postes

2.3 Description des employeurs recourant au Cifre

Cette section se concentre sur une description des employeurs recourant au Cifre. Il s'agit dans un premier temps de présenter le profil de ces entreprises, à la fois juridique, d'activité mais aussi géographique. Nous décrivons aussi l'évolution de certains résultats d'intérêt, comme le chiffre d'affaires ou l'emploi, suivant le recrutement de doctorants en Cifre. Nous présentons ensuite une comparaison initiale entre les entreprises traitées et notre groupe de comparaison.

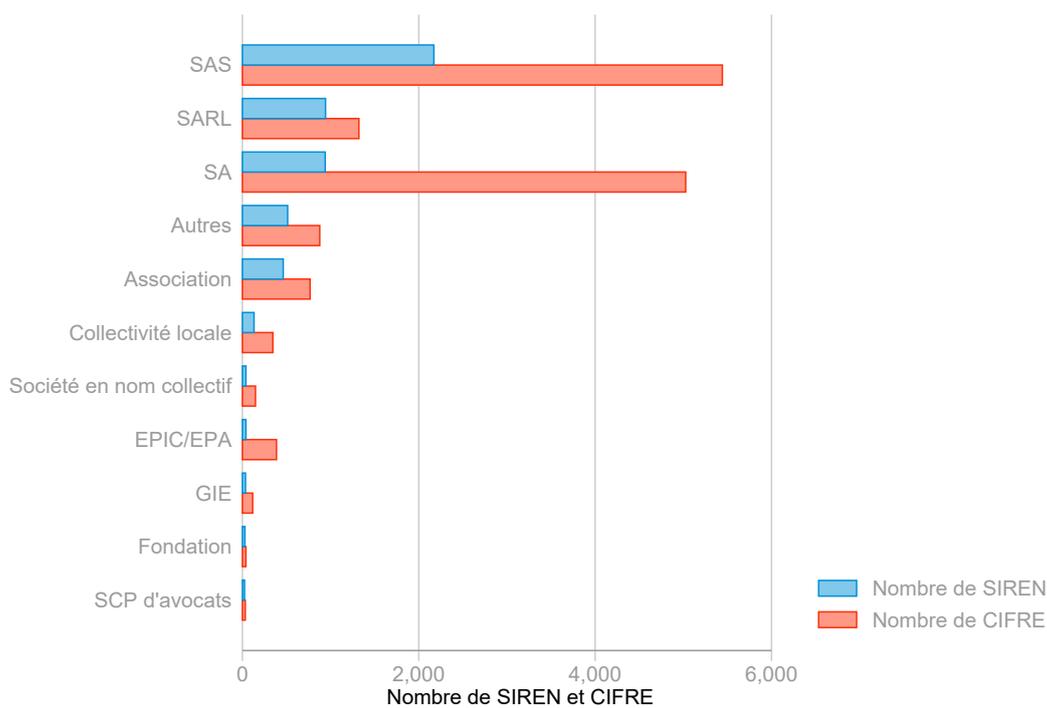
2.3.1 Caractéristiques des employeurs

Nous présentons ici nos résultats concernant les caractéristiques des entreprises accueillant des Cifre. La figure 2.11 montre la distribution des employeurs accueillant des Cifres par catégorie juridique. La vaste majorité des entités d'accueil sont des entreprises (SAS, SA, SARL) : elles représentent un total d'environ 4 000 unités légales, soit près de 80 % du total des employeurs de doctorants Cifre. Les associations ainsi que les collectivités locales (communes, départements) sont également assez nombreuses à recourir au dispositif.

Il est également intéressant de comparer ces chiffres à ceux de la distribution des doctorants dans ces différentes structures. On note notamment que les SA, bien que moitié moins nombreuses que les SAS, prennent presque autant de doctorants Cifre, alors que le nombre de doctorants est très bas pour les SARL qui prennent moins de deux doctorants par entité. Il y a une véritable différence de structure entre ces entreprises, qui fait qu'à nombre égal elles ont des demandes de doctorants divergentes. On peut aussi noter le nombre élevé de doctorants dans les EPIC et EPA, malgré leur nombre très faible.

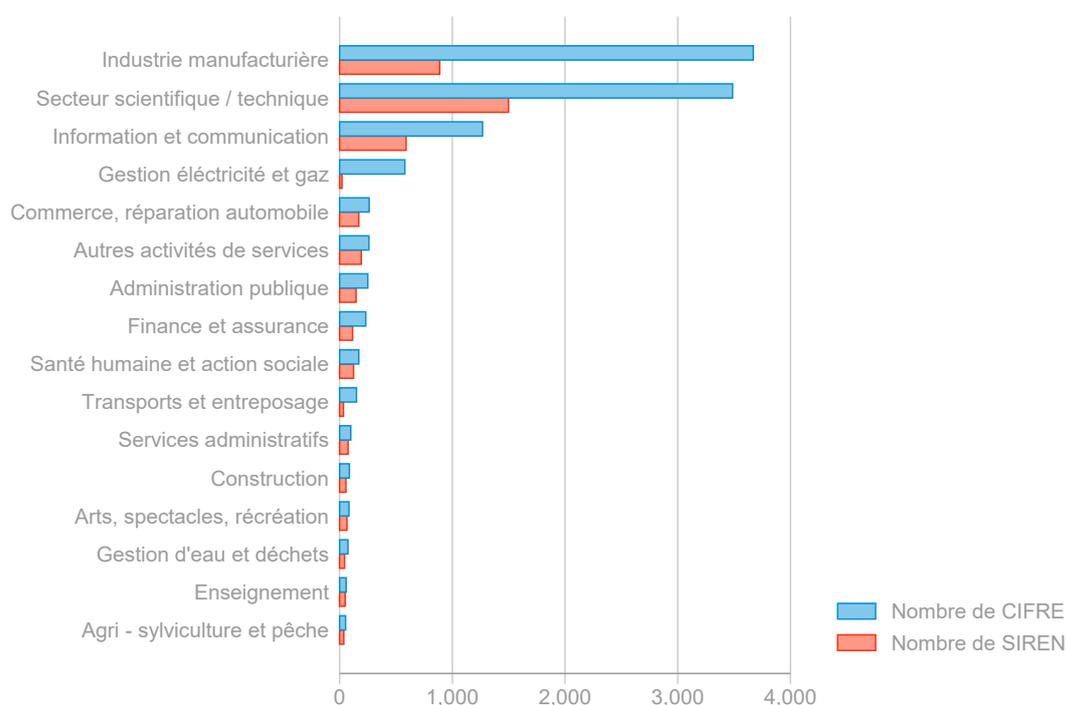
La figure 2.12 détaille les secteurs d'activité des entreprises employant des doctorants Cifre. Nous voyons que si les SIREN sont assez bien répartis, une majorité écrasante des Cifre se trouve dans les industries manufacturières et les activités

FIGURE 2.11 – Catégorie juridique des employeurs



spécialisées et scientifiques. On observe aussi que la production et distribution d'électricité, secteur avec très peu d'entreprises, emploie un grand nombre de doctorants par rapport au nombre de SIREN.

FIGURE 2.12 – Secteur d'activité des employeurs.



Nous nous intéressons aussi à la distribution géographique des entreprises accueillant une Cifre. Tout d'abord, nous observons la distribution géographique des doctorants Cifre et des entreprises accueillant des Cifre par zone d'emploi. Ces cartes montrent notamment les "pôles" de Cifres : l'Île-de-France, notamment autour de Paris, les alentours de Grenoble et de Lannion, ainsi que certaines zones d'emploi dans le Sud de la France. Ces pôles correspondent souvent à des pôles d'innovation spécialisés.

Afin de contextualiser ces chiffres, nous représentons aussi le ratio de Cifre divisé par le nombre d'entreprises employant des ingénieurs (PCS commençant par 38) ou employant de la main d'œuvre hautement qualifiée (PCS commençant par 3). Nous définissons celles-ci comme des entreprises dont la moitié ou plus des emplois sont ingénieurs ou hautement qualifiés. Nous avons donc le nombre de Cifres par le nombre d'entreprises susceptibles d'effectuer de la recherche par zone d'emploi. Cette analyse montre des concentrations élevées de Cifre par entreprise

dans le Sud, en région parisienne et Île-de-France, autour de Grenoble et en Bretagne. Cet indicateur favorise les pôles qui sont dans des zones moins denses, qui ont donc moins d'entreprises. On ne peut pas parler d'une répartition homogène : les Cifre ont tendance à se trouver dans des clusters denses, allant jusqu'à (trouver le maximum pour une ZE), alors que 59 ZE n'ont pas de Cifre effectué dans les données dont nous disposons.

FIGURE 2.13 – Nombre de doctorants Cifre par ZE.

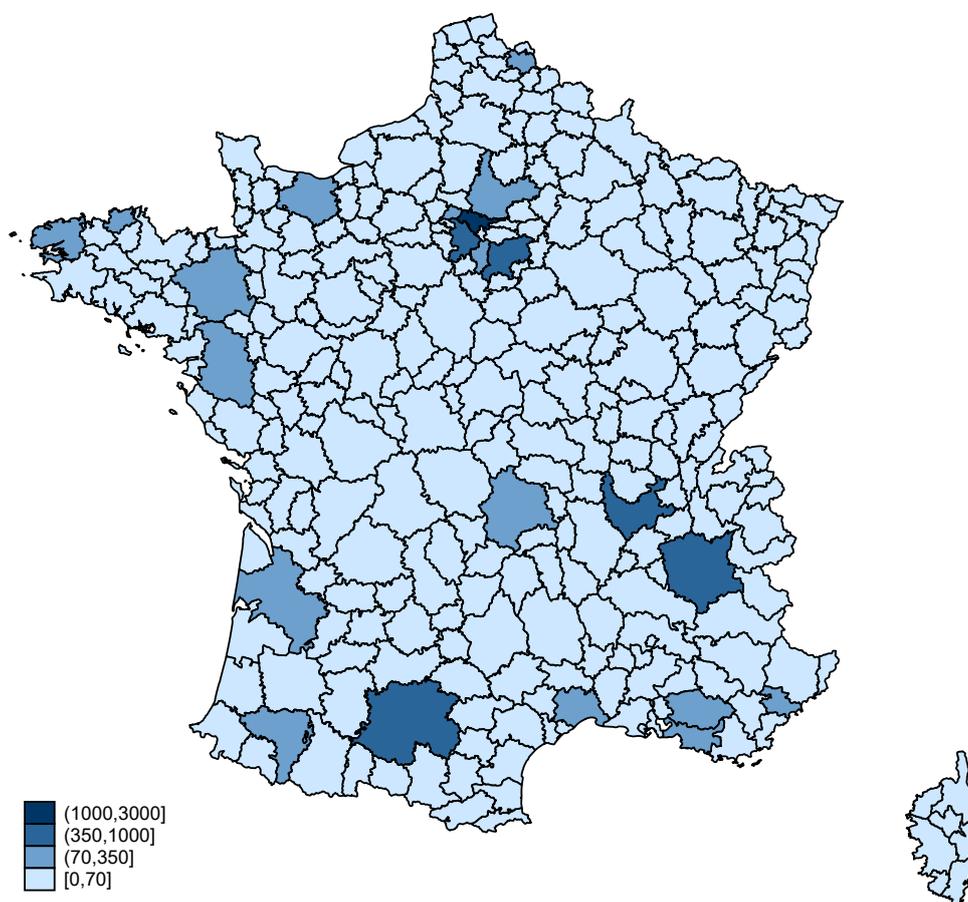


FIGURE 2.14 – Nombre d'établissements ayant accueilli une Cifre par ZE.

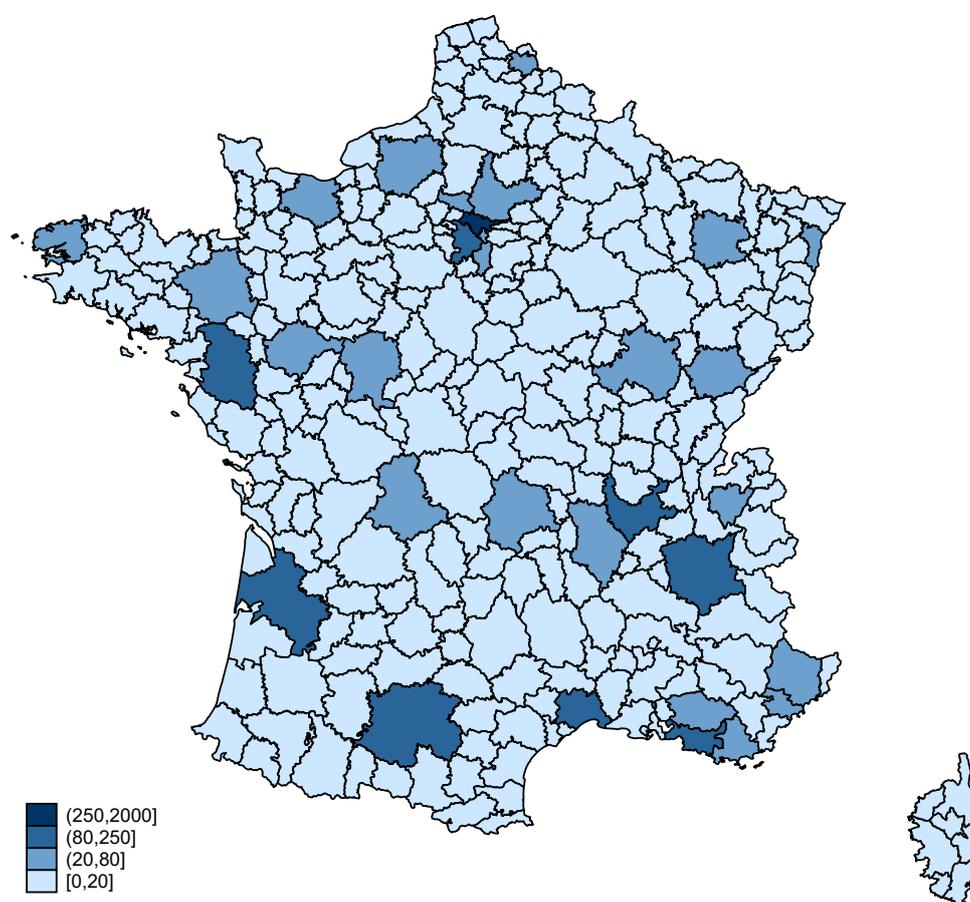


FIGURE 2.15 – Ratio du nombre de doctorants Cifre / entreprises avec emplois hautement qualifiés .

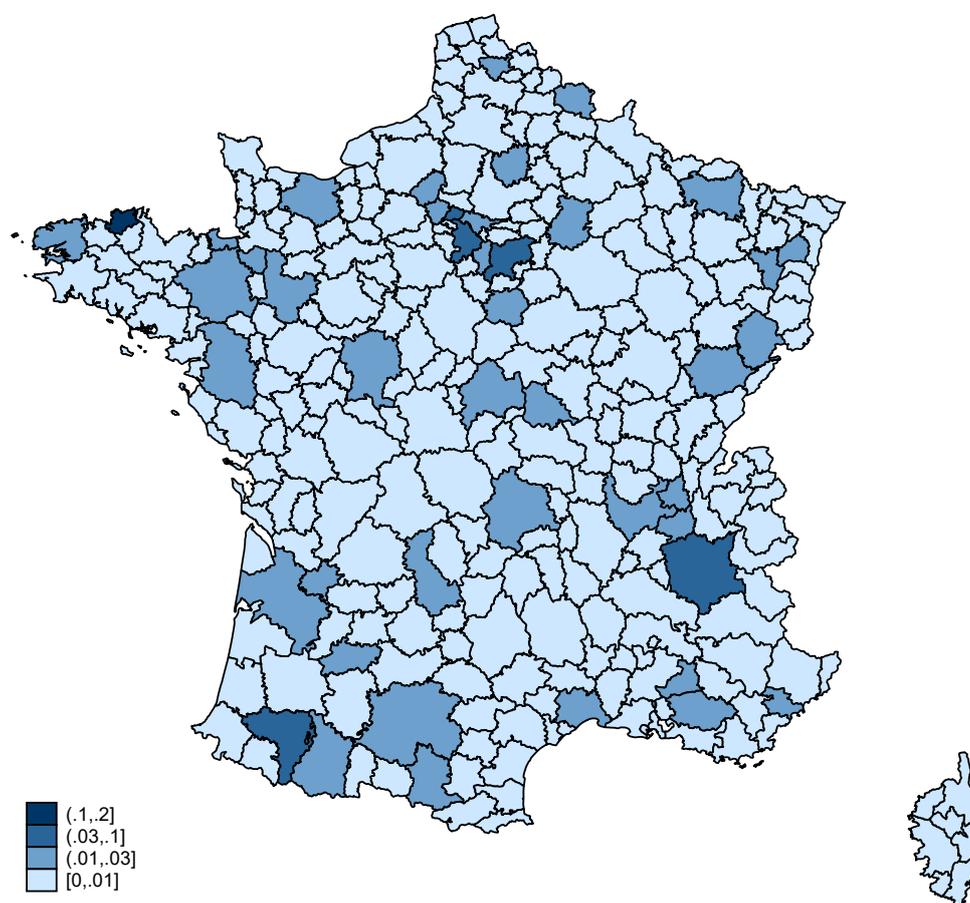
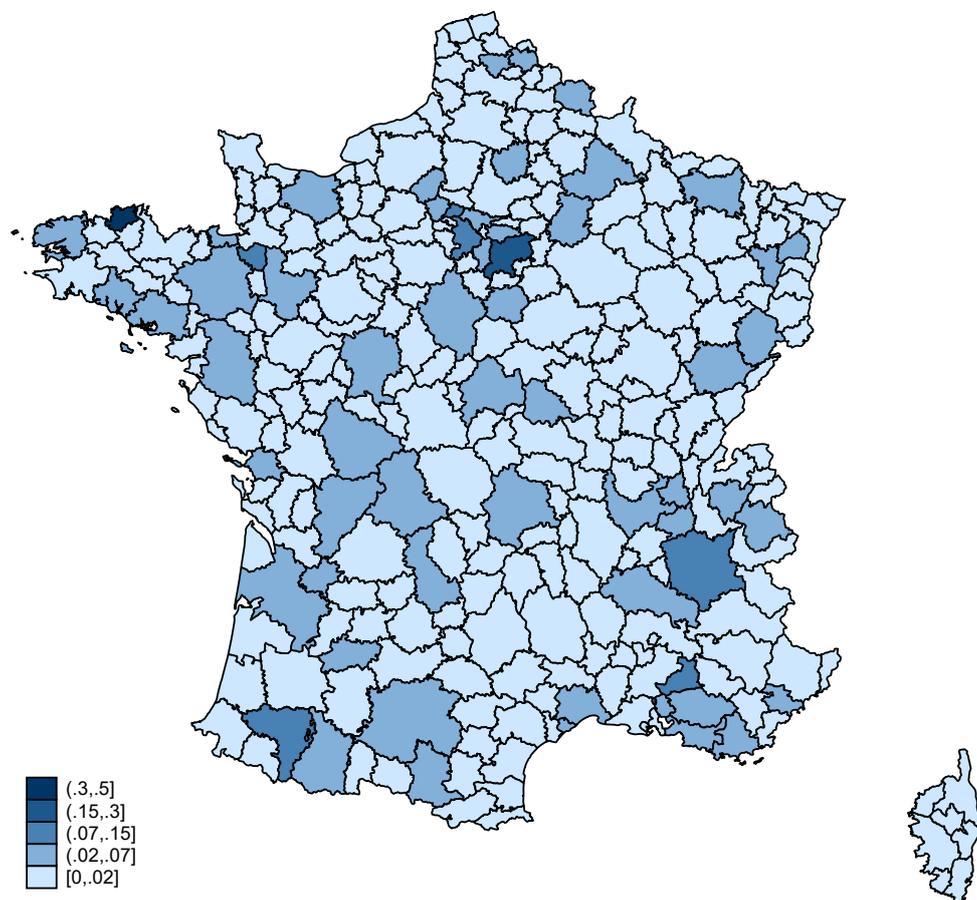


FIGURE 2.16 – Ratio du nombre de doctorants Cifre / entreprises employant des ingénieurs.



2.3.2 Évolution des caractéristiques des entreprises recourant au dispositif Cifre par rapport à des groupes comparables

Cette section propose une comparaison des évolutions sur la période étudiée (2008-2016) des caractéristiques des entreprises ayant recours au dispositif Cifre, par rapport à deux groupes d'entreprises susceptibles d'avoir des traits communs avec le groupe d'entreprises traitées. Le premier groupe de comparaison est naturellement celui des entreprises ayant recours au Crédit Impôt Recherche au cours de la période : ce crédit d'impôt porte sur l'ensemble des dépenses de recherche-développement, et est très complémentaire du dispositif Cifre puisque les salaires versés à des doctorants Cifre sont éligibles au versement du Crédit Impôt Recherche. Le second groupe d'entreprises avec lequel on peut tenter d'établir une comparaison est l'ensemble des entreprises employant des ingénieurs : il s'agit d'un groupe beaucoup plus vaste que les seules entreprises innovantes, mais il permet d'approcher le plus possible le groupe d'entreprises susceptibles d'avoir recours au dispositif Cifre.

Comme le montre le tableau 2.18, les entreprises ayant recours au dispositif Cifre sont en moyenne beaucoup plus importantes en termes de nombre d'employés que les entreprises ayant recours au CIR mais pas au Cifre : le ratio des effectifs moyens était d'environ 15 en 2008, et la différence s'est accrue au cours de la période avec un ratio supérieur à vingt en 2015 sur l'emploi salarié total entre firmes Cifre et firmes recourant au CIR seulement. La différence de taille s'explique très simplement par le fait que la grande majorité des géants technologiques français utilisent à la fois les dispositifs Cifre et CIR, mais sont exclus du groupe CIR pour s'assurer de comparer des entreprises différentes. Cette différence ne se traduit en revanche pas par une composition très différente des effectifs salariés, puisque les parts d'employés ingénieurs et techniciens sont très compa-

TABLEAU 2.18 – Tableau comparant les caractéristiques des entreprises Cifre à celles ayant seulement recours au CIR, et à celles employant seulement des ingénieurs, en 2008 et 2015

En 2008	CIFRE	CIR seulement	Ingé. seulement
Emploi	3641.03	245.18	37.10
Part emploi R&D	0.10	0.06	0.01
Part emploi R&D : ingénieur	0.20	0.18	0.10
Part emploi R&D : techniciens	0.10	0.10	0.06
Nombre de demandes de brevets	49.92	1.13	0.03
Observations	296	9472	130811
En 2015			
Emploi	4202.65	183.84	38.51
Part emploi R&D	0.13	0.13	0.03
Part emploi R&D : ingénieur	0.14	0.13	0.11
Part emploi R&D : techniciens	0.09	0.08	0.07
Nombre de demandes de brevets	0.00	0.00	0.00
Observations	445	11905	116723

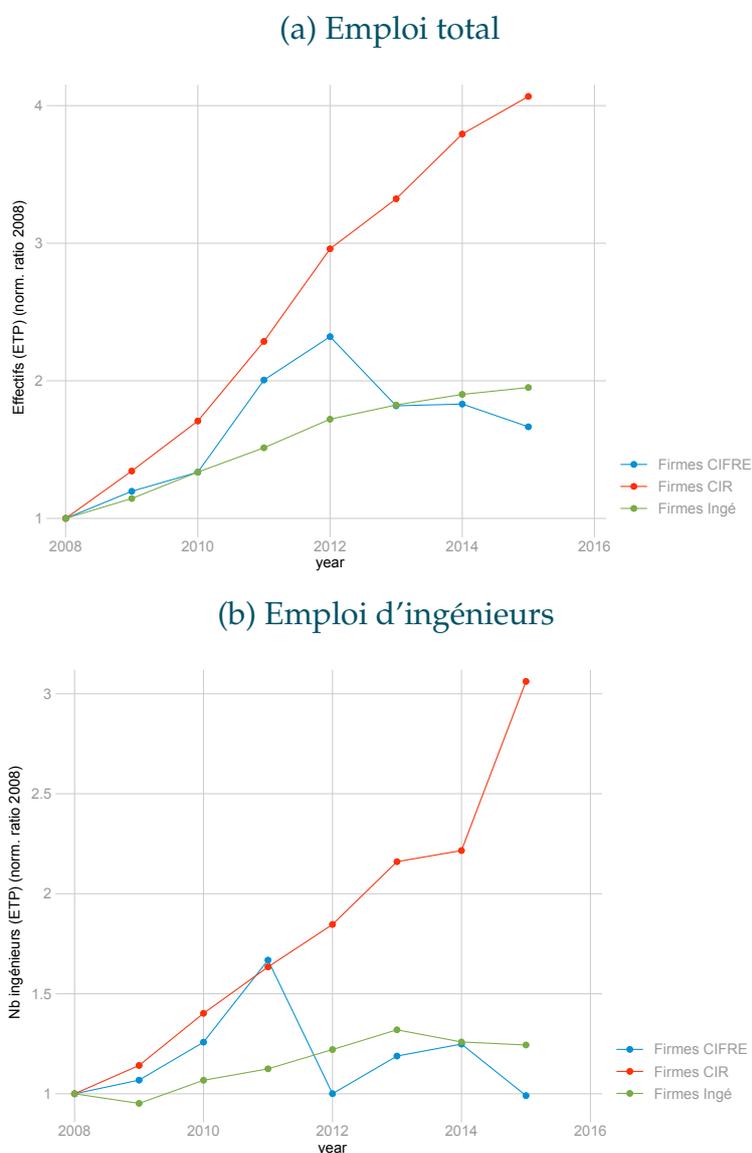
NOTES : La première colonne contient le groupe des entreprises ayant recours au dispositif Cifre pendant l'année concernée, la seconde colonne le groupe des entreprises ayant recours au CIR mais pas au dispositif Cifre pendant cette même année, et la troisième colonne le groupe des entreprises employant des ingénieurs mais ne recourant à aucun des deux dispositifs précités cette même année.

SOURCES : Bases Cifre, GECIR, DADS Postes.

rables entre les deux groupes. L'ensemble des firmes employant des ingénieurs et ne recourant ni au CIR ni au Cifre, présenté dans la troisième colonne, est en revanche très différent des deux autres sur l'ensemble de la période : les entreprises qui le composent sont en moyenne cent fois plus petites en termes d'effectifs, et ont également des parts de l'emploi consacré à la R&D beaucoup plus faibles.

Le graphique 2.17 représente l'évolution sur la période 2008-2015 de l'emploi total (panneau a) et de l'emploi d'ingénieurs (panneau b), en base 100 au début de la période. Il montre que les trois groupes d'entreprises (Cifre, CIR seulement, employeurs d'ingénieurs seulement) sont sur des trajectoires ascendantes d'emploi sur la période. Ces groupes sont définis comme l'ensemble des entreprises respectant le critère les caractérisant au moins une année dans la période. La dynamique d'emploi des firmes ayant recours au dispositif Cifre est positive mais plus modérée que celle des entreprises ayant recours au CIR seulement, qui contient vraisem-

FIGURE 2.17 – Évolution moyenne de l’emploi total et de l’emploi ingénieur des firmes Cifre, CIR seulement et employeuses d’ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.

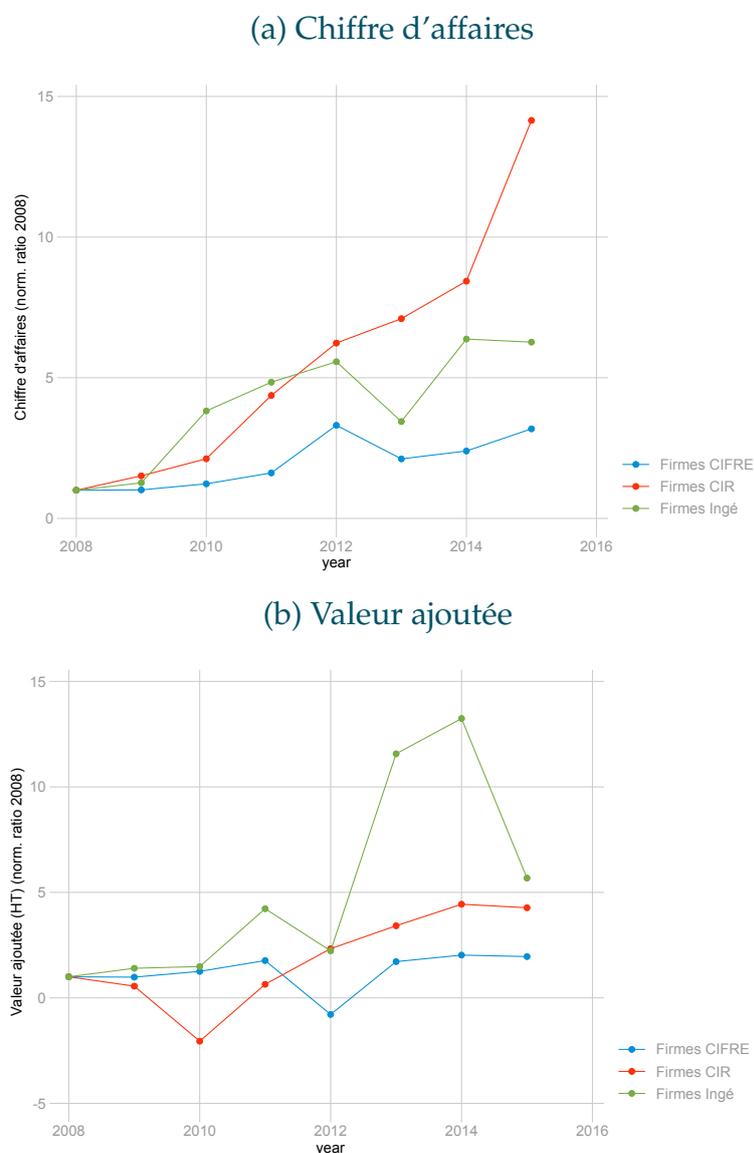


NOTE : La moyenne est non-pondérée.

SOURCE : Bases Cifre, GECIR, DADS Postes.

blement le plus d’entreprises jeunes à très forte croissance. Le graphique 2.18 confirme cette image d’une évolution plutôt positive des firmes ayant recours au Cifre au cours de la période, mais la nuance légèrement. En effet, les firmes Cifre sont en moyenne en croissance en termes de chiffre d’affaires et de valeur ajoutée,

FIGURE 2.18 – Évolution moyenne du chiffre d'affaires et de la valeur ajoutée (HT) des firmes Cifre, CIR seulement et employeuses d'ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.



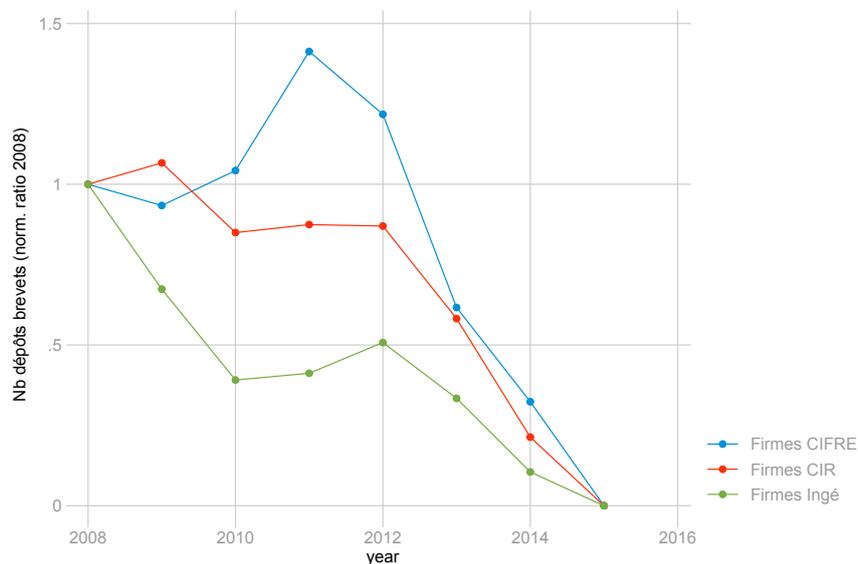
NOTE : La moyenne est non-pondérée.

SOURCE : Bases Cifre, GECIR, DADS Postes, FARE.

mais ces tendances sont moins marquées que pour les firmes CIR seulement ou employeuses d'ingénieurs seulement.

Le très faible nombre de demandes de brevets observable dans le tableau 2.18 pour l'ensemble des groupes dans l'année 2015 s'explique par le retard avec lequel

FIGURE 2.19 – Nombre de demandes de brevets par les entreprises Cifre, CIR et ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.



SOURCE : Bases Cifre, GECIR, DADS Postes, PATSTAT.

les données de brevets sont accessibles. En effet, la base PATSTAT à partir de laquelle sont construites ces statistiques est progressivement enrichie à mesure que les offices de brevets communiquent sur les demandes. Le graphique 2.19 illustre ce problème : les demandes de brevets faites par des firmes Cifre semblent augmenter légèrement mais être comparables à celles des firmes demandant le CIR jusqu'en 2012 (malgré un bruit statistique important lié à la relative rareté des demandes de brevets), mais chutent ensuite conjointement avec celles des autres groupes.

2.3.3 Évolution des entreprises avant et après l'acquisition d'une Cifre

Cette section s'intéresse à la description de l'évolution des entreprises ayant eu recours au dispositif Cifre au cours de la période, pour les années précédant et suivant leur premier recours au dispositif. Afin de s'assurer que la thèse Cifre

considérée est bien la première de l'entreprise, nous nous concentrons sur la période 2008-2016, et utilisons les informations dont nous disposons sur la période 2003-2007 pour exclure les entreprises déjà présentes dans la base Cifre à cette période. L'ensemble d'entreprises considéré est donc l'ensemble des groupes (toutes les informations sont consolidées au niveau groupe) ayant eu recours au dispositif Cifre pendant la période 2008-2016 mais pas pendant la période 2003-2007. L'année de référence est l'année d'effet du premier contrat Cifre alloué.

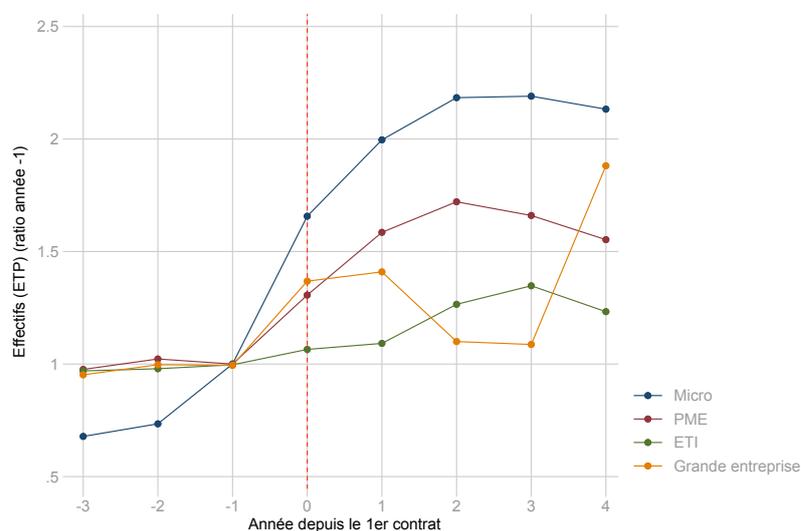
Le graphique 2.20 montre l'évolution moyenne des effectifs salariés des entreprises ayant recours au Cifre, année par année autour du recrutement de leur premier doctorant Cifre. Il montre que le recrutement d'une thèse Cifre s'accompagne en moyenne d'une forte hausse des effectifs, en particulier sur les petites entreprises, qui sont également des entreprises qui recrutent peu de Cifre sur la période.

Le graphique 2.21 présente les variations du nombre d'ingénieurs autour du recrutement d'une thèse Cifre. Afin de pouvoir présenter des variations définies pour les firmes n'ayant aucun ingénieur l'année précédant le recrutement, nous utilisons la différence de la fonction asinh (fonction inverse hyperbolique du sinus) qui approche le logarithme naturel mais est définie en zéro. Ce graphique montre que le recrutement d'une thèse Cifre correspond en moyenne à une augmentation du nombre d'ingénieurs dans le groupe : il correspond à une augmentation d'environ 25 % de l'effectif d'ingénieurs dans les micro-entreprises, mais à une augmentation bien plus faible en proportion dans les entreprises de taille plus importante.

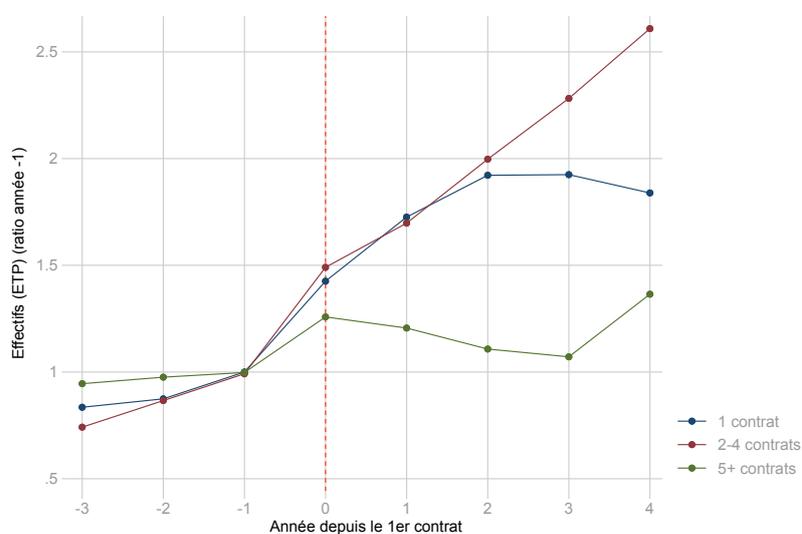
Le graphique 2.22 illustre le même phénomène que celui évoqué dans le graphique 2.19 : toutes les catégories d'entreprises semblent déposer moins de brevets après le recrutement d'une Cifre qu'avant. Ceci est simplement dû au fait que les données sur le nombre de brevets se raréfient lorsqu'on analyse des années plus tardives, mais sera corrigé par des effets fixes année dans l'analyse causale par

FIGURE 2.20 – Évolution du nombre de salariés, avant et après le début d’une thèse Cifre - normalisé à 1 l’année précédant le recrutement

(a) Par nombre de Cifre recrutés



(b) Par taille d’entreprise

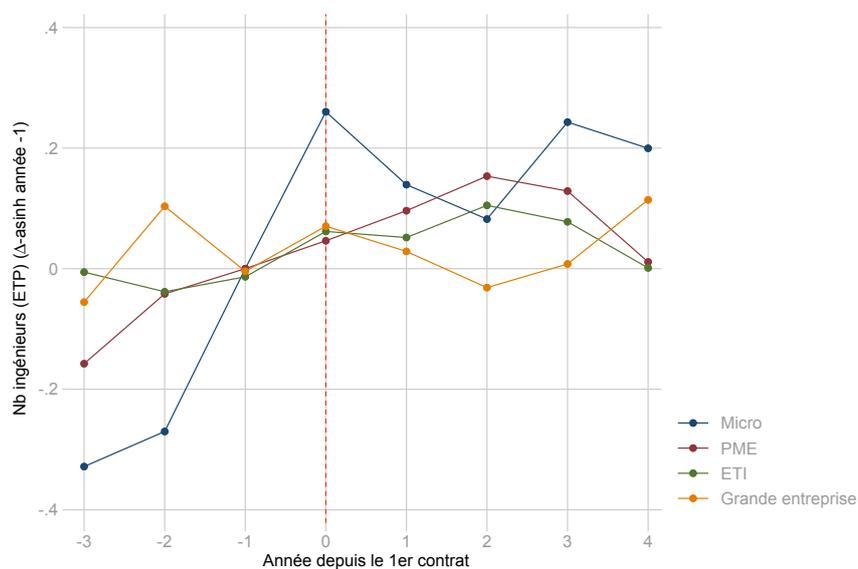


NOTES : Les catégories de taille d’entreprise sont définies au niveau groupe selon les critères de la loi LME de 2008. Les catégories de nombre de contrats sont définies à partir du nombre de contrats alloués sur la période 2008-2016. Les données sont consolidées à niveau groupe, les groupes étant constitués à partir de LIFI et des périmètres des groupes fiscaux.

SOURCE : Bases Cifre, FARE, LIFI, Périmètres fiscaux.

étude d’événement.

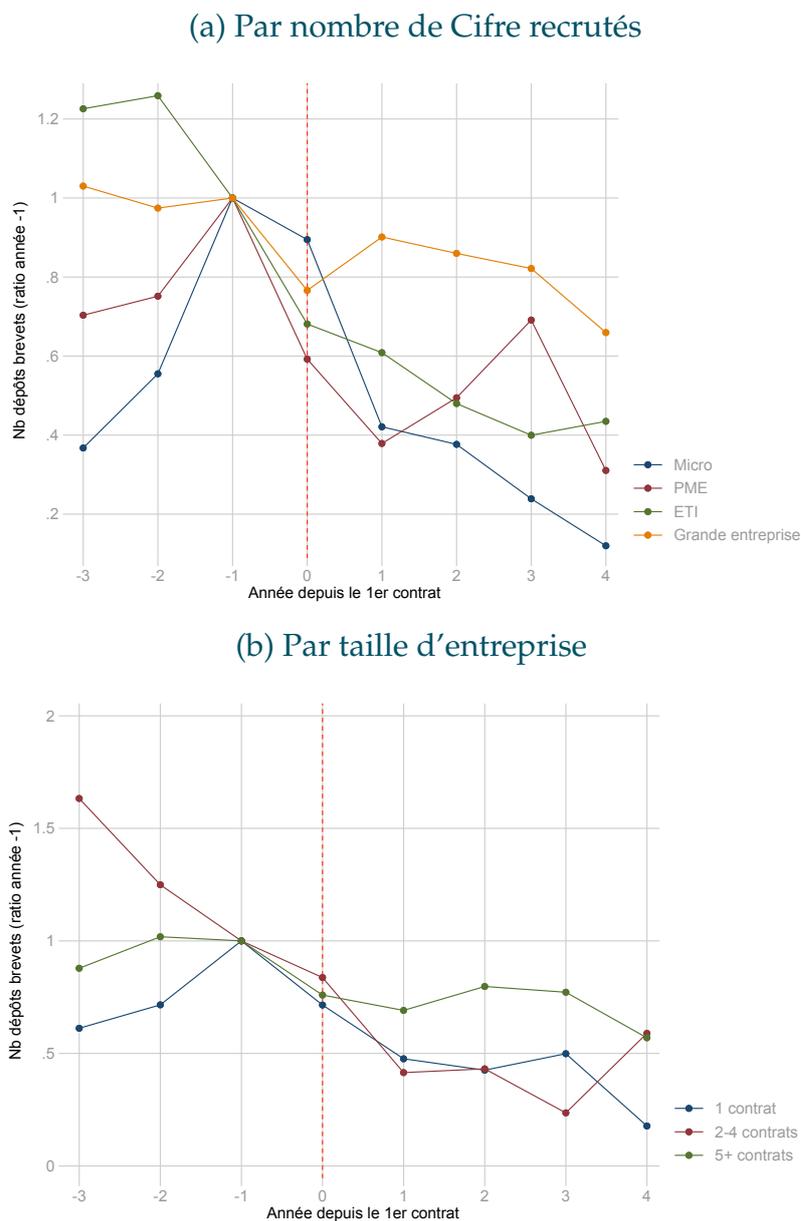
FIGURE 2.21 – Variations du nombre d’ingénieurs dans le groupe avec le recrutement d’une thèse Cifre - fonction asinh normalisée à 0 l’année précédant le recrutement



NOTES : Les catégories de taille d’entreprise sont définies au niveau groupe selon les critères de la loi LME de 2008. Les données sont consolidées à niveau groupe, les groupes étant constitués à partir de LIFI et des périmètres des groupes fiscaux.

SOURCE : Bases Cifre, GECIR, DADS Postes, PATSTAT.

FIGURE 2.22 – Évolution du nombre de brevets, avant et après le début d'une thèse Cifre - normalisé à 1 l'année précédant le recrutement



NOTES : Les catégories de taille d'entreprise sont définies au niveau groupe selon les critères de la loi LME de 2008. Les données sont consolidées à niveau groupe, les groupes étant constitués à partir de LIFI et des périmètres des groupes fiscaux.

SOURCE : Bases Cifre, LIFI, Périmètres fiscaux, PATSTAT.

CHAPITRE 3

MÉTHODOLOGIE

Comme indiqué dans l'introduction du rapport intermédiaire, l'objectif initial de l'étude est d'évaluer les effets du dispositif Cifre sur trois types d'acteurs qui sont partie prenante d'un contrat Cifre : les entreprises, les doctorants et les structures de recherche.

Le mauvais appariement au niveau individuel entre les bases de suivi du dispositif et les données administratives rendent l'évaluation des effets du Cifre sur les doctorants difficile. De plus, en l'absence de données de panel fiables sur les performances des laboratoires, l'évaluation d'impact sur ces unités apparaît délicate. Dès lors, la méthodologie présentée dans cette section va se concentrer sur l'évaluation des effets du dispositif Cifre sur les entreprises.

3.1 La méthode des études d'évènements

Comme suggéré dans les clauses techniques de l'appel d'offre et dans la réponse à l'appel d'offre, une première stratégie d'identification sera menée, en suivant la méthodologie des études d'évènements. Celle-ci restreindra l'échantillon aux entreprises entrant dans le dispositif au cours de la période, et comparera l'évolution de différentes variables d'intérêt de ces entreprises en fonction du mo-

ment de leur entrée dans le programme (recrutement d'un doctorant ou d'une doctorante en Cifre). Il s'agit donc d'une étude d'événement comparant uniquement les traités sur une fenêtre temporelle, en exploitant la variation entre entreprises de la date du traitement.

Par définition, les entreprises traitées – c'est-à-dire ayant recours au dispositif Cifre – dès le début de la période de l'échantillon seront exclues. Cette restriction mènera probablement à exclure l'ensemble des grands groupes français. S'ils occupent une place importante et ont a priori tous recruté en Cifre depuis la création de la mesure, le grand nombre de PME recrutant des doctorants Cifre (environ 60 % des entreprises pour 40 % des financements) devrait permettre de construire l'étude d'événement sur un nombre important d'entreprises traitées une seule fois au cours de la période.

Une autre restriction potentielle concerne les entreprises traitées de multiples fois. Si l'on considère un échantillon allant de 2008 à 2016, la question se pose de savoir s'il est préférable de conserver ou bien d'exclure de l'échantillon les entreprises ayant recours aux Cifres plusieurs fois, par exemple en 2010 et 2012. Inclure des entreprises correspondant à ce cas de figure peut être problématique car l'intensité du traitement sera alors très hétérogène entre unités. Néanmoins les exclure revient implicitement à conditionner sur la trajectoire future des entreprises présentes dans l'échantillon, créant potentiellement une sélection négative. En effet, conditionnellement à avoir recours au Cifre au moins une fois, le fait de n'y avoir recours qu'une seule fois est potentiellement associé à une croissance et une dynamique moindre de l'entreprise, voire à une expérience infructueuse la première fois. En l'absence de solution parfaite à ce problème, nous procéderons à une estimation avec et sans la restriction aux entreprises n'ayant recours au dispositif que pour un seul contrat.

Enfin, une condition importante pour isoler l'impact de Cifre en particulier sera de contrôler pour le recours à d'autres dispositifs de soutien à la R&D qui sont sus-

ceptibles d’être corrélés avec le recours à Cifre. Si la grande diversité de ces aides rend très difficile de contrôler pour l’ensemble, il sera fait usage de la base GECIR afin de contrôler pour le montant de CIR reçu par les entreprises étudiées ainsi que des données de la liasse fiscale afin de contrôler pour le recours au programme JEI (jeune entreprise innovante) – tout du moins dans sa composante fiscale.

3.1.1 Équation d’estimation

Nous estimerons une spécification dynamique qui autorise l’effet du traitement à varier en fonction du temps écoulé depuis le traitement.

L’année de traitement d’une entreprise i est notée t_{i0} . On peut alors indexer le temps écoulé par rapport au traitement avec $d = t - t_{i0}$ où t est une année calendaire. L’indice d est négatif avant le traitement et positif après. Notre échantillon d’estimation couvrira a minima les années 2008 à 2016. La fenêtre d’estimation sera définie par le biais d’une borne inférieure m_0 et supérieure m_1 . Les observations incluses dans l’échantillon d’estimation seront donc telles que : $d \in \{-m_0, -m_0 + 1, \dots, m_1 - 1, +m_1\}$.

L’équation d’estimation de base s’écrit alors :

$$Y_{it} = \sum_{\substack{d=-m_0+1 \\ d \neq -1}}^{d=m_1} \beta_d \times \mathbf{1}\{t = d + t_{i0}\} + \mathbf{x}'_{it} \boldsymbol{\delta} + \alpha_i + \psi_t + \varepsilon_{it}, \quad (3.1)$$

où $\mathbf{1}\{\}$ est la fonction indicatrice ; α_i et ψ_t sont des effets fixes année et entreprise ; \mathbf{x}'_{it} est vecteur de variables de contrôle qui varient au cours du temps.

On propose ici d’omettre deux variables indicatrices pour $d = -m_0$ et $d = -1$. Cette restriction est nécessaire pour éviter la multicollinéarité et pour identifier le processus de génération de données dynamiques sous-jacent si l’ensemble des unités incluses dans l’échantillon d’estimation est traité Borusyak et Jaravel (2017); Gross et al. (2019). Pour nous assurer que cette restriction n’a pas d’influence trop

forte sur nos résultats, nous présenterons les résultats avec des normalisations alternatives. Nous pouvons évaluer par exemple nos résultats en combinant les coefficients associés à m_0 et $m_0 - 1$, c'est-à-dire en les contraignant à être égaux comme proposé par Schmidheiny et Siegloch (2019).

3.1.2 Discussion des hypothèses d'identification

Bien que la méthode des études d'évènements soit très fréquemment employée, les conditions sous lesquelles elle permet d'aboutir à une estimation non-biaisée des effets qu'elle vise à mesurer ne sont pas toujours explicitées. Cette sous-section a pour objet de clarifier les hypothèses d'identification sous-tendant les études d'évènement. Pour étudier sous quelles hypothèses l'estimation de l'équation (3.1) estime l'effet de traitement moyen sans biais, il est utile de suivre Abraham et Sun (2019) et d'introduire une notation en résultat potentiel (*potential outcome*).

Dans les études d'évènements, le vecteur des états de traitement d'une unité peut être caractérisé de façon unique par la période du traitement initial, la date de l'évènement. Nous désignons cette variable aléatoire discrète par $t_{i0} = \min\{t : D_{i,t} = 1\}$. En effet, le traitement est un état absorbant (une fois traitée, une unité est considérée comme toujours traitée quelque soit la valeur de d) et le vecteur des valeur de la variable de traitement $D_{i,t}$ est dès lors une séquence non-décroissante (en t , les années calendaires, mais aussi en d le temps relatif au traitement) de zéros puis de uns, c'est-à-dire $D_{i,t} < D_{i,s}$ pour $s < t$.

Nous notons $Y_{i,t}^c$ pour indiquer le résultat potentiel dans la période t sous une trajectoire de traitement hypothétique (date de l'évènement) c . Soit $Y_{i,t}^B$ pour indiquer le résultat si l'unité i ne reçoit aucun traitement. Il s'agit de la trajectoire de base, c'est-à-dire celle qu'aurait suivi l'unité i en l'absence de traitement.

Nous définissons l'effet du traitement pour un temps depuis la date de l'évènement c comme la différence entre la trajectoire de base $Y_{i,t}^B$ et la trajectoire lorsque

le traitement a eu lieu en c :

$$\alpha_{i,t}^c \equiv Y_{i,t}^c - Y_{i,t}^B. \quad (3.2)$$

Bien entendu, pour chaque unité, nous observons chaque unité sous un seul chemin de traitement t_{i0} . Dès lors, le résultat observé pour l'unité i est simplement

$$Y_{i,t} = Y_{i,t}^{t_{i0}}.$$

On peut exprimer la variable d'intérêt observée comme la somme de la trajectoire de base et des effets spécifiques à chaque date de traitement :

$$Y_{i,t} = Y_{i,t}^B + \sum_c (Y_{i,t}^c - Y_{i,t}^B) \times \mathbf{1}\{c = t_{i0}\} \quad (3.3)$$

Abraham et Sun (2019) montrent que sous trois ensemble d'hypothèses, l'estimation de l'équation (3.1) par moindres carrés ordinaires aboutit à une estimation causalement interprétable de l'effet moyen.

1. La première hypothèse¹ concerne l'homogénéité des effets moyens entre "cohortes" où une cohorte fait référence à l'ensemble des unités qui ont reçu le traitement à une date donnée (par exemple la cohorte c est définie comme : $\{i \in N : t_{i0} = c\}$). Le traitement moyen pour la cohorte c pour une distance au traitement donnée l est défini comme $CATT_l^c = E(\alpha_{i,c+l}^c | t_{i0} = c)$. L'hypothèse s'écrit donc :

$$CATT_l^c = CATT_l, \quad \forall c$$

2. La deuxième hypothèse² est une généralisation de l'hypothèse de tendance commune entre cohortes. Formellement :

$$E(Y_{i,t}^B - Y_{i,s}^B | t_{i0} = c) = E(Y_{i,t}^B - Y_{i,s}^B), \quad \forall c \in \text{supp}(t_{i0}), \forall t, s$$

1. Il s'agit de *Assumption 3* dans Abraham et Sun (2019).

2. Il s'agit de *Assumption 1* dans Abraham et Sun (2019).

où supp est le support de la variable aléatoire.

3. La troisième hypothèse³ concerne l'absence d'anticipation. Elle revient à supposer que le traitement futur n'a pas d'effet dans le passé. Formellement :

$$Y_{i,c+d}^e = Y_{i,c+d}^B, \quad \forall d < 0$$

Elle implique qu'avant la date de traitement ($d < 0$), la variable observée correspond à la variable en l'absence de traitement : $Y_{i,c+d}^B$. Dès lors, sous cette hypothèse, le comportement pour $d < 0$ de la variable dépendante observée $Y_{i,t}$ avec $t < t_{i0}$ est informatif quant à la validité de l'hypothèse de tendance commune énoncée ci-dessus.

Dans le cadre de notre application, l'hypothèse d'homogénéité de l'effet du traitement entre cohortes ne semble pas particulièrement problématique.

Les entreprises ayant recours au dispositif Cifre pour la première fois au cours de la période 2008-2016 vont tendre à être des PME et auront certainement des caractéristiques assez proches, au moins entre cohortes. De plus, l'hétérogénéité liée à des évolutions structurelles de l'économie au cours de la période est vraisemblablement limitée du fait de la courte durée de la période couverte par l'échantillon.

L'hypothèse vraisemblablement la plus problématique est l'absence d'effet d'anticipation. En effet, avoir recours au dispositif Cifre est le fruit d'un choix délibéré de l'entreprise et nécessite vraisemblablement un long travail de préparation (coordination avec un laboratoire, etc.). Dès lors il n'est pas clair que l'absence ou bien la présence de tendance pré-traitement (*pre-trends*) reflète l'anticipation de la politique ou bien une violation de l'hypothèse 2 de tendance commune.

3. Il s'agit de *Assumption 2* dans Abraham et Sun (2019).

3.2 Différence de différences et appariement

La méthode des études d'événements, dans laquelle le groupe témoin est constitué d'entreprises qui sont traitées à un moment différent, ne fonde l'estimation des effets du dispositif que sur la comparaison d'entreprises qui, à un moment ou un autre au cours de la période couverte par l'échantillon, vont effectivement être traitées. Cela peut être perçu comme un avantage si l'on suppose que le choix de rejoindre à un moment ou à un autre le dispositif Cifre est déterminé par une forme d'hétérogénéité inobservable qui est stable au cours du temps et difficilement captée par des variables observables. Néanmoins, il est vraisemblable que cette technique souffre d'une puissance statistique limitée à cause du faible nombre d'entreprises ayant effectivement recours au programme Cifre.

Dès lors que les résultats de l'étude d'évènement apparaîtraient peu concluants car imprécis, il pourra être utile d'avoir recours à une procédure d'appariement conjuguée avec une estimation en double différence. Cette approche permet de considérer un groupe de contrôle plus important en sélectionnant sur la base de critères explicités ci-dessous un sous-ensemble des entreprises n'ayant jamais recours au dispositif Cifre – c'est-à-dire *jamais* traitées.

Nous prévoyons de sélectionner l'ensemble des entreprises témoin au moyen d'une procédure d'appariement avec l'objectif qu'elles ressemblent aux entreprises traitées, bien qu'elles n'aient pas décidé de ne pas participer au programme. L'objectif de la procédure d'appariement est d'identifier les entreprises qui seraient affectées de la même manière en cas de chocs agrégés sur plusieurs séries de résultats d'intérêt, afin de rendre crédible l'hypothèse de tendance parallèle.

Cette stratégie d'identification est complémentaire à l'étude d'évènement. Néanmoins puisque qu'elle s'appuie sur une comparaison avant/après des variables d'intérêt d'une entreprise traitée par rapport à un groupe de contrôle, elle ne permettra pas d'étudier les effets sur dispositifs sur des grandes entreprises qui ont

constamment recours au dispositif.

3.2.1 Description du processus d'appariement

Les traitements ont lieu à des moments différents, il semble approprié de mettre en place une procédure d'appariement séparément pour chaque événement. La date pertinente est donc l'année relative à l'adhésion de l'entreprise au programme Cifre. Comme précédemment nous indiquons t_{i0} les dates auxquelles l'entreprise est traitée et d le nombre de périodes relatives à cet événement. Afin de pouvoir tester l'hypothèse de tendances communes, nous ne procéderons à des appariements que sur la base d'observations collectées jusqu'à 3 ans avant la date de l'événement. Pour être valide, le groupe témoin ne devra pas inclure d'entreprises qui seront traitées au cours de la même période. Pour chaque date d'événement t_{i0} , nous appairerons donc les entreprises traitées durant cette période aux entreprises qui ne sont jamais traitées.

En ce qui concerne les détails de l'appariement, nous proposons deux approches : (i) un appariement exact sur des variables catégorielles ou bien discrétisées – c'est-à-dire initialement continues mais rendues catégorielles via discrétisation (*coarsening*)–, (ii) un appariement sur la base d'un score de propension.

L'appariement exact sur variables catégorielles s'appuiera a minima sur les caractéristiques suivantes :

1. Participation au programme JEI et CIR;
2. Taille de l'entreprise;
3. Productivité apparente du travail (discrétisée);
4. Secteur d'activité;

De cette procédure, nous obtiendrons un grand nombre d'entreprises de contrôle associées à chaque entreprise traitée. Le nombre d'entreprises de contrôle associé

à chaque unité traitée variera. Nous conserverons un grand nombre de firmes témoin appariées, le but de cette approche étant d'améliorer la précision de l'estimation.

L'appariement par score de propension reposera sur des principes similaires mais permettra d'inclure un plus grand nombre de variables continues.

3.2.2 Équation d'estimation

Nous dénotons T l'ensemble des entreprises i telles qu'elles font partie du groupe de traitement sur la période considérée. L'équation d'estimation est très proche de (3.1), elle s'écrit :

$$Y_{it} = \sum_{\substack{d=m_1 \\ d=-m_0 \\ d \neq -1}} \beta_d \times \mathbf{1}\{t = d + t_{i0}\} \times \mathbf{1}\{i \in T\} + \mathbf{x}'_{it} \boldsymbol{\delta} + \alpha_i + \psi_t + \varepsilon_{it}, \quad (3.4)$$

Nous regroupons les erreurs-types au niveau de la firme afin de répondre aux problèmes d'auto-corrélation des erreurs entre périodes (Bertrand et al., 2004) mais aussi de prendre en compte le fait que certaines entreprises témoin peuvent servir de contrôle plusieurs fois.

RÉFÉRENCES

- Abraham, S. et Sun, L. (2019). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Available at SSRN 3158747*.
- ANRT (2019). *Conditions générales d'octroi et d'éligibilité des conventions industrielles de formation par la recherche*.
- Bertrand, M., Duflo, E., et Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1), p. 249–275.
- Borusyak, K. et Jaravel, X. (2017). Revisiting event study designs with an application to the estimation of the marginal propensity to consume. *mimeo*, p. 33.
- Gross, T., Notowidigdo, M. J., et Wang, J. (2019). The marginal propensity to consume over the business cycle. *American Economic Journal : Macroeconomics*.
- Margolis, D. et Miotti, L. (2015). Évaluation de l'impact du dispositif jeunes docteurs du crédit d'impôt recherche. *Rapport, Ministère de l'Éducation nationale, de l'enseignement supérieur et de la recherche*.
- Schmidheiny, K. et Siegloch, S. (2019). On event study designs and distributed-lag models : Equivalence, generalization and practical implications. *CEPR Discussion Paper No. DP13477*.

LISTE DES TABLEAUX

2.1	Régressions linéaire du taux de soutenance sur le statut Cifre	25
2.2	Statistiques Descriptives - 2013, 2010	29
2.3	Statistiques Descriptives - 2013, 2010	30
2.4	Statistiques Descriptives - 2013, 2010	31
2.5	Statistiques Descriptives - 2013, 2010	31
2.6	Régression du CDI sur statut Docteur et CIFRE - STEM	33
2.7	Régression du salaire sur statut CIFRE et Docteur - STEM	34
2.8	Régression du CDI sur statut Docteur et CIFRE - Ingénieurs	35
2.9	Régression du salaire sur statut Docteur et CIFRE - Ingénieurs	35
2.10	Régression du CDI sur statut CIFRE et Ingénieur - Docteurs	36
2.11	Régression du salaire sur statut CIFRE et Ingénieur - Docteurs	37
2.12	Régression du CDI sur statut CIFRE - Docteurs Ingénieurs	38
2.13	Régression du salaire sur statut CIFRE - Docteurs Ingénieurs	38
2.14	Tableau décrivant l'évolution de la taille de l'échantillon lors de l'appariement	41
2.15	Statistiques descriptives des doctorants post-Cifre	42
2.16	Régressions linéaire du CDI sur le le fait de rester dans la même entreprise	43
2.17	Régressions linéaire du salaire sur le fait de rester dans la même entreprise	44

2.18	Tableau comparant les caractéristiques des entreprises Cifre à celles ayant seulement recours au CIR, et à celles employant seulement des ingénieurs, en 2008 et 2015	53
------	---	----

LISTE DES FIGURES

2.1	Évolution du nombre de Cifres validées.	16
2.2	Évolution du nombre d'entreprises ayant recours à une convention Cifre.	17
2.3	Évolution du nombre de Cifres validées par entreprise.	18
2.4	Évolution de la subvention ANRT au cours de la période.	19
2.5	Distribution de l'âge des doctorants (conventions 2011-2018)	21
2.6	Nationalité des doctorants Cifre	21
2.7	Proportion de chaque genre parmi les doctorants	22
2.8	Disciplines des doctorants Cifre	23
2.9	Diplômes des doctorants Cifre	24
2.10	Taux de soutenance par discipline theses.fr (2011-2018)	26
2.11	Catégorie juridique des employeurs	46
2.12	Secteur d'activité des employeurs.	47
2.13	Nombre de doctorants Cifre par ZE.	48
2.14	Nombre d'établissements ayant accueilli une Cifre par ZE.	49
2.15	Ratio du nombre de doctorants Cifre / entreprises avec emplois hautement qualifiées	50
2.16	Ratio du nombre de doctorants Cifre / entreprises employant des ingénieurs.	51

2.17	Évolution moyenne de l'emploi total et de l'emploi ingénieur des firmes Cifre, CIR seulement et employeuses d'ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.	54
2.18	Évolution moyenne du chiffre d'affaires et de la valeur ajoutée (HT) des firmes Cifre, CIR seulement et employeuses d'ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.	55
2.19	Nombre de demandes de brevets par les entreprises Cifre, CIR et ingénieurs seulement sur la période 2008-2015, normalisé à 1 en 2008.	56
2.20	Évolution du nombre de salariés, avant et après le début d'une thèse Cifre - normalisé à 1 l'année précédant le recrutement	58
2.21	Variations du nombre d'ingénieurs dans le groupe avec le recrutement d'une thèse Cifre - fonction asinh normalisée à 0 l'année précédant le recrutement	59
2.22	Évolution du nombre de brevets, avant et après le début d'une thèse Cifre - normalisé à 1 l'année précédant le recrutement	60



L'Institut des politiques publiques (IPP) est développé dans le cadre d'un partenariat scientifique entre PSE-Ecole d'économie de Paris (PSE) et le Centre de Recherche en Economie et Statistique (CREST). L'IPP vise à promouvoir l'analyse et l'évaluation quantitatives des politiques publiques en s'appuyant sur les méthodes les plus récentes de la recherche en économie.

PSE a pour ambition de développer, au plus haut niveau international, la recherche en économie et la diffusion de ses résultats. Elle rassemble une communauté de près de 140 chercheurs et 200 doctorants, et offre des enseignements en Master, École d'été et Executive education à la pointe de la discipline économique. Fondée par le CNRS, l'EHESS, l'ENS, l'École des Ponts-ParisTech, l'INRA, et l'Université Paris 1 Panthéon Sorbonne, PSE associe à son projet des partenaires privés et institutionnels. Désormais solidement installée dans le paysage académique mondial, la fondation décloisonne ce qui doit l'être pour accomplir son ambition d'excellence : elle associe l'université et les grandes écoles, nourrit les échanges entre l'analyse économique et les autres sciences sociales, inscrit la recherche académique dans la société, et appuie les travaux de ses équipes sur de multiples partenariats. www.parisschoolofeconomics.eu



Le CREST est un centre de recherche regroupant des chercheurs de l'ENSAE, de l'ENSAI et du département d'économie de l'École Polytechnique. Centre interdisciplinaire spécialisé en méthodes quantitatives appliquées aux sciences sociales, le CREST est organisé en 4 thématiques: Economie, Statistiques, Finance-Assurance et Sociologie. La culture commune des équipes est celle d'un attachement fort aux méthodes quantitatives, aux données, à la modélisation mathématiques, et d'allers-retours continus entre les modèles théoriques et les preuves empiriques permettant d'analyser des problématiques sociétales et économiques concrètes. <http://crest.science>



Avis sur le rapport intermédiaire de l'évaluation des Conventions industrielles de formation par la recherche (Cifre)

Egidio Luis Miotti, OST

Le rapport intermédiaire présente les caractéristiques du programme CIFRE et celles de ses participants, aussi bien les entreprises que les doctorants.

La description des caractéristiques du programme est claire et ne présente pas de problèmes autres que la faiblesse de la base administrative pour la période 2004-2011.

L'analyse des doctorants est la partie la plus intéressante du rapport intermédiaire et ses conclusions semblent robustes :

1. Les doctorants ont en moyenne 24 ans, majoritairement français et masculins.
2. Ils effectuent de la recherche dans les sciences physiques, naturelles et de l'ingénieur.
3. Ils ont une probabilité plus élevée de soutenir leur thèse.
4. Après la thèse, ils bénéficient d'une prime de salaire et de contrats de travail plus stables.

La conclusion 4 mérite un traitement plus approfondi, particulièrement en ce qui concerne la prime de salaire.

L'analyse des employeurs est dans un état moins avancé que celui des doctorants. Certaines conclusions préliminaires semblent robustes et cohérentes avec les objectifs du dispositif CIFRE :

1. Les employeurs des doctorants CIFRE sont majoritairement des entreprises, concentrées dans des secteurs industriels et scientifiques.
2. Cette concentration est aussi géographique, dans des bassins avec une forte densité d'entreprises de recherche.
3. Le recrutement d'un doctorant en CIFRE s'accompagne en moyenne d'une hausse d'effectifs et du nombre d'ingénieurs dans l'entreprise.

L'analyse en termes de performances technologiques mérite un approfondissement. L'utilisation de nombre de brevets ne paraît pas convaincante, davantage pour le manque de clarté méthodologique (quels brevets ? quels offices (INPI, USPTO ? OMPI, EPO ?)) que pour les résultats présentés.

Avis : le rapport est de bonne facture et respecte l'état de l'art en matière des techniques d'évaluation. Il est rigoureux et jusqu'ici convaincant. Les résultats obtenus sont ceux attendus. Un point doit être approfondi : il faut, dans la mesure du possible, trancher sur la question concernant la prime de salaires aux doctorants bénéficiaires d'un CIFRE : s'agit-elle d'une prime à la recherche ou d'une captation de rente résultat de l'aide ?