

---

**RAPPORTS**

Andrea ALBANESE  
Joël MACHADO  
Michel TENIKUE  
Bertrand VERHEYDEN

# Evaluation du CAI : caractéristiques des participants et faisabilité d'une analyse d'impact

Evaluation du CAI:  
caractéristiques des participants et  
faisabilité d'une analyse d'impact<sup>1</sup>

## Rapport Final

---

<sup>1</sup> **Remerciements.** Ce rapport résulte d'une convention spécifique entre L'Etat du Grand-Duché de Luxembourg (Ministère de la Famille, de l'Intégration, et à la Grande Région – MIFA) et le LISER. Les auteurs remercient le MIFA pour son soutien financier et Frédéric Docquier pour ses précieux commentaires.

# Table des matières

1. Introduction

2. Méthodologie

2.1. Les données

2.1.1. Population des signataires

2.1.2. Population des non-signataires.

2.2. Stratégie empirique

3. Analyse de la participation au CAI

3.1. Profil des signataires : Statistiques univariées

3.2. Profil des signataires : Modèles économétriques

4. Profil des signataires ayant obtenu le certificat et de ceux n'ayant pas complété le programme

4.1. Propension à obtenir le certificat : Statistiques univariées

4.2. Propension à obtenir le certificat : Modèles économétriques

5. Faisabilité de l'implémentation d'une étude causale de la participation au CAI sur l'intégration socio-économiques des signataires

5.1. Sources de données et variables d'intérêts

5.2. Evaluation causale du programme CAI sur la trajectoire socio-économique : méthodologie

5.3. Limites de l'approche empirique suggérée et extensions possibles

6. Conclusions et recommandations

Tableaux

Références

Appendice technique

# 1. Introduction

Depuis 2011, le Ministère de la Famille, de l'Intégration et à la Grande Région (MiFa) propose un programme d'intégration volontaire aux résidents de nationalité étrangère, âgés d'au moins 16 ans, appelé Contrat d'Accueil et d'Intégration (CAI). Dans sa forme originale, ce programme est composé de trois parties : (1) une journée d'orientation générale expliquant aux participants les administrations publiques du pays ; (2) un cours civique présentant l'organisation du cadre institutionnel et les us et coutumes de la vie quotidienne au Luxembourg ; et (3) la mise à disposition de bons à tarif réduit permettant l'accès à des cours de langues subventionnés dans l'une des trois langues officielles du pays. Les signataires du CAI ont 12 mois pour entamer l'une de ses trois composantes. Une fois le CAI débuté, le signataire a 24 mois pour clôturer les 3 composantes requises. Sous certaines conditions, et sur base d'une demande dûment justifiée, une extension de 12 mois (officialisée par un avenant au contrat) peut être octroyée au signataire ayant entamé mais non complété le CAI endéans les 2 ans requis.

Dans le contexte de la réforme de la loi sur l'intégration, le MiFa a entrepris de moderniser le CAI afin de le rendre plus attractif auprès de la population cible. Une phase de ce processus vise à analyser les déterminants de la participation au CAI et à évaluer l'impact de cette participation sur l'intégration (sur le marché du travail, le marché du logement, l'éducation, la vie sociale, etc.) des participants.

Une première étape de cette évaluation consiste à exploiter les données recueillies par le MiFa entre 2017 et 2022, pour dresser le profil des participants. Précisément, le premier objectif du présent rapport est de décrire les groupes de participants potentiels qui ont signé le contrat, honoré ses prestations et les groupes de populations éligibles qui ont été plus difficiles à atteindre. Afin de comparer les participants au CAI à la population éligible, la base du MiFa sera reliée à la base de données administratives de l'IGSS, ce qui permettra d'obtenir les caractéristiques socio-professionnelles de l'ensemble de la population d'intérêt. Les participants et les non-participants seront comparés sur la base de multiples caractéristiques observables (âge, sexe, nationalité, localisation géographique, composition familiale, statut professionnel, etc.). Cette comparaison permettra d'identifier les caractéristiques associées à une participation élevée ou faible au programme volontaire, de même que les sous-groupes de la population éligible qui ne sont pas (ou faiblement) représentés. De la même manière, on identifiera les caractéristiques qui permettent de mieux distinguer les candidats signataires des participants ayant honoré les prestations du CAI.

Dans un deuxième temps, ce projet explore les possibilités de mettre en œuvre une évaluation de l'effet causal de la participation au CAI sur différentes variables de résultat disponibles soit dans les données administratives (emploi, revenu, logement, mobilité géographique, etc.), soit dans des données d'enquêtes existantes (conditions de vie, force de travail, valeurs, etc.) ou encore dans des données à collecter de façon spécifique (bien-être, acculturation, sentiments d'intégration, d'appartenance, relations interculturelles, etc.). La méthodologie proposée exige de disposer de données longitudinales qui recouvrent les périodes avant et après la participation au CAI.

## 2. Méthodologie

### 2.1. Les données

L'analyse centrale de ce rapport repose sur la fusion de deux bases de données administratives. La première est mise à disposition par le Ministère de la Famille et de la Grande Région (MiFa) et contient l'information sur les signataires du CAI. Pour la période 2017-2022, cette base renseigne également quelques variables socio-démographiques auto-déclarées (tel que le niveau d'éducation ou le type de titre de séjour) des signataires et le statut de chaque contrat signé tel qu'enregistré en date du 30 juin 2022 (en cours, conclu ou arrêté). Pour chaque signataire, il est possible de connaître le nombre de contrats signés, leurs dates de signatures ainsi que le dernier statut du participant connu en date du 30 juin 2022 : intégré si la personne a honoré les trois éléments constituant le CAI (journée d'orientation, cours d'instruction civique, cours de langue) ; non-intégré/non-intéressé si la personne n'a pas conclu le CAI ; ou suivi si le dossier du signataire est toujours ouvert. Pour la période 2012-2016, et dans le respect des réglementations d'anonymisation des données, le MiFa a fourni une base de données comprenant l'identifiant pseudonymisé des signataires ainsi que leur statut final (intégré ou non) uniquement. Étant donné le manque d'informations additionnelles sur ces signataires (et notamment la date de signature de leur CAI), cette information additionnelle est uniquement utilisée pour préciser le groupe des non-signataires, en supprimant de ce groupe de population les signataires de la période 2012-2016.

La deuxième source de données est la base de données mise à disposition par la sécurité sociale luxembourgeoise et comprend le statut socio-professionnel mensuel de chaque immigrant éligible au CAI. Cette base de données fournit les caractéristiques socio-démographiques qui permettent d'éclairer la sélection dans la participation au CAI, notamment l'âge, le sexe, la nationalité, la durée de résidence au Luxembourg

depuis 2010 ainsi que plusieurs variables décrivant le statut socio-professionnel de l'individu.<sup>2</sup> Les variables utilisées dans l'analyse sont définies dans l'appendice technique. Les deux bases de données sont appariées et permettent de disposer des informations à la fois sur les signataires et non-signataires du CAI parmi la quasi-totalité de la population étrangère éligible au CAI sur la période 2017-2022.<sup>3</sup>

L'unité d'observation de notre analyse est l'individu. La population cible est composée de tout non-luxembourgeois, âgé d'au moins 16 ans, vivant (au moins à un moment donné) au Luxembourg sur la période 2017-2022. Nous détaillons d'abord la définition de la population des signataires dans la Section 2.1.1 et de la population des non-signataires dans la Section 2.1.2. Ensuite, nous décrivons notre stratégie empirique dans la Section 2.2.

### 2.1.1. Population de signataires

Les signataires du CAI sont identifiés (par un identifiant pseudonymisé) à travers la base du MiFa. Entre 2017 et juin 2022, 6.392 individus ont signé le CAI une seule fois et 297 individus, n'ayant pas accompli avec succès le CAI dans le temps imparti après leur première signature, ont signé un deuxième contrat CAI sur cette période. Un seul contrat (et donc une seule date de signature) est retenu par signataire et nous privilégions la première signature.<sup>4</sup> Cette date de signature est utilisée pour déterminer les caractéristiques de chaque participant au CAI. Plus précisément, pour éviter les fluctuations possibles endéans le mois de la signature du CAI, le profil des signataires est établi au mois qui précède leur signature.

Parmi ces 6.689 individus de la base du MiFa, 9 n'ont pas pu être identifiés dans la base IGSS. Nous disposons donc au total d'un échantillon de 6.680 individus signataires. Néanmoins, nous excluons de l'analyse les individus âgés de moins de 20 ans, bien que l'âge d'éligibilité au CAI soit de 16 ans. Ce choix se justifie par le fait que les données dont nous disposons ne permettent pas de distinguer les individus âgés de moins de 16 ans de ceux âgés de 16 à 20 ans. L'inclusion de ces individus impliquerait donc la présence dans le groupe des non-signataires de nombreux

---

<sup>2</sup> Nous n'avons accès qu'aux données individuelles de l'IGSS ultérieures à 2010 et ne pouvons donc exclure que la présence dans la base de l'IGSS antérieure à cette date ne soit uniquement liée à un emploi et non à la résidence (c'est-à-dire se référer à un travailleur frontalier, non éligible au CAI). Nous nous limitons donc à exploiter l'information observable à partir de 2010 afin d'éviter des erreurs de classification (voir l'appendice technique pour plus de détails).

<sup>3</sup> Cette fusion est effectuée grâce à un identifiant pseudonymisé, qui a été généré par la « Luxembourg Microdata Platform on Labour and Social Protection ».

<sup>4</sup> Cependant, pour 83 signataires, le premier CAI a été signé avant le 1 janvier 2017, et la date exacte de cette première signature nous est inconnue. Dès lors, les caractéristiques du signataire au moment de cette signature ne peuvent être observées. Afin de ne pas supprimer ces 83 signataires de l'analyse, nous utilisons la date de signature de leur deuxième CAI ayant eu lieu entre 2017-2022.

individus qui ne sont pas (encore) éligibles parce qu'ils ne remplissent pas la condition d'âge du CAI. Ce critère nous amène à exclure 55 individus qui avaient moins de 20 ans au moment de la signature (soit moins de 1.5% du total des signataires). Notre échantillon définitif de signataires s'élève donc à 6.625 individus.

### 2.1.2. Population de non-signataires

La population éligible mais non-signataire du CAI inclut tout non-luxembourgeois, âgé d'au moins 20 ans, et vivant (au moins à un moment donné) au Luxembourg sur la période 2017-2022. Par construction, les non-signataires n'apparaissent pas dans la base du MiFa mais uniquement dans la base de l'IGSS. Notre analyse vise à comparer leurs caractéristiques à celles des signataires. Pour chaque individu, nous ne retenons qu'une seule observation. Comme décrit ci-avant, les caractéristiques des signataires sont déterminées au mois précédant leur première signature, afin d'éviter les changements de statuts en cours du mois de signature. Les non-signataires n'ont, par définition, pas de date de signature et nous devons donc définir un moment auquel leurs caractéristiques sont déterminées. Afin d'assurer la représentativité de notre échantillon de non-signataires sur la période analysée, la date d'observation de leurs caractéristiques est tirée de façon aléatoire au sein de leur période de présence dans les données de l'IGSS entre 2017 et 2022.<sup>5</sup> Sur cette base, nous disposons de 262.010 individus non-signataires.

## 2.2. Stratégie empirique

L'avantage d'une analyse empirique pour étudier les déterminants de la participation au CAI est donné par la possibilité de prendre en compte simultanément les multiples caractéristiques des individus disponibles dans notre base de données. En effet, ces caractéristiques peuvent être corrélées ; par exemple, 80% des signataires ayant le statut de co-assuré sont originaires d'un pays non-européen. L'analyse empirique permet de mesurer l'effet net de chacune d'entre elles.

---

<sup>5</sup> On pourrait être tenté de prendre un mois particulier comme point de référence (par exemple juin 2022) pour chaque non-signataire. Cependant, une telle approche risquerait de biaiser la composition de la population éligible i) en faisant abstraction d'individus ayant quitté le pays entre-temps et ii) en ne prenant en compte qu'un environnement macroéconomique particulier, potentiellement corrélé à des variables comme la participation au marché du travail. Une approche alternative aurait été de prendre la population éligible au moment de chaque signature observée dans la base des signataires. Cependant, cette approche impliquerait que les individus ne participant jamais au CAI et résidant dans le pays pendant de nombreuses années auraient une pondération plus importante que les individus arrivés ultérieurement ou n'étant présents dans le pays que pendant une courte période. Nous décrivons dans l'appendice technique un dernier ajustement d'ordre technique par rapport aux individus n'ayant aucun lien observable avec la sécurité sociale du Grand-Duché sur la période d'observation.

L'approche empirique adoptée se base sur un modèle de probabilité linéaire (Linear Probability Model). Les modèles estimés se formulent ainsi :

$$Y_i = \sum_j^J \beta^j X_i^j + \sum_k^K \beta^k X_i^k + \sum_l^L \beta^l X_i^l + \sum_m^M \beta^m X_i^m + t + \epsilon_i .$$

Dans la première partie de l'analyse, nous étudions la relation entre **le fait de prendre part au CAI** et les caractéristiques des individus. Dans ce cas, la variable dépendante  $Y_i$  prend la valeur 1 si l'individu a signé le contrat et 0 s'il est non-signataire. Bien que cette variable soit binaire, nous optons pour un modèle de probabilité linéaire (plutôt qu'un modèle non-linéaire basé sur un logit ou probit) qui permet une interprétation simple et directe des coefficients estimés ainsi qu'une comparaison entre nos modèles alternatifs. Wooldridge (2013) indique qu'en pratique, le modèle de probabilité linéaire et les modèles non-linéaires tendent à produire des effets marginaux similaires lorsque les valeurs des variables indépendantes sont proches de leurs valeurs moyennes. En revanche, les estimations des effets partiels aux extrémités de la distribution des variables indépendantes peuvent différer entre les deux approches. Malgré cet inconvénient, étant donné que notre objectif est d'identifier l'effet partiel des variables indépendantes à la moyenne de leur distribution sur la variable dépendante  $Y_i$ , cette approche reste appropriée (Wooldridge, 2013).

Quatre modèles différents sont estimés, sur base de l'inclusion progressive de différents groupes de variables de contrôle qui mesurent les caractéristiques des individus (dénotés respectivement  $X_i^j, X_i^k, X_i^l$  et  $X_i^m$ ). Le premier groupe de variables,  $X_i^j$ , comprend les variables socio-démographiques de l'individu : sexe, groupe d'âge, canton de résidence, état civil, taille du ménage, né au Luxembourg, nationalité (regroupée par groupe géo-linguistique), durée de résidence au Luxembourg depuis 2010. Le deuxième groupe,  $X_i^k$ , comprend des variables mesurant le statut professionnel : employé, au chômage, reçoit le revenu minimum garanti (RMG avant 2019, REVIS depuis le 1er janvier 2019), co-assuré, autre inactif, fonctionnaire, ou indépendant. Pour les individus employés, nous disposons également de certaines caractéristiques de leur emploi : travail manuel (ouvrier), sans CDI, bénéficie d'une mesure d'activation, et le dernier salaire horaire réel connu. Nous incluons également dans ce groupe certaines variables qui contrôlent pour l'expérience professionnelle de l'individu depuis 2010 : sans salaire, % du temps comme « autre inactif », % du temps au chômage. Le troisième groupe de variables,  $X_i^l$ , contrôle pour 21 secteurs d'emploi différents (y compris l'absence de secteur). Enfin, le quatrième groupe,  $X_i^m$ , détaille les 42 nationalités individuelles ainsi que 12 groupes géo-linguistiques agrégeant le reste des nationalités. Dans le modèle 4, les variables détaillant les nationalités remplacent les agrégats géo-linguistiques utilisés dans le groupe de variables de contrôle base,



$X_i^j$ . Tous les modèles incluent également un effet fixe temporel qui contrôle pour l'année de signature ;  $\epsilon_i$  représente le terme d'erreur.

Dans la deuxième partie de l'analyse, nous étudions la relation entre la **réussite du CAI** et les caractéristiques des individus. Cette réussite est déterminée par la validation par le signataire de sa participation aux 3 composantes du programme. Les signataires du CAI ont 12 mois pour entamer l'une de ses trois composantes. Une fois le CAI débuté, le signataire a 24 mois pour clôturer les 3 conditions requises. L'échéancier typique du CAI implique donc que les contrats les plus récents, signés après 2020 susceptibles d'être clôturés dans les délais prévus. Dans cette partie de l'analyse nous nous concentrons dès lors sur les contrats signés entre 2017 et 2020, pour lesquels les participants ont disposé d'au moins deux ans et demi (jusqu'en juin 2022, date à laquelle nous mesurant le statut final du contrat) pour satisfaire aux trois exigences du CAI. Notons que, comme pour la première partie de l'analyse, nous utilisons les caractéristiques des signataires au mois précédant la signature.

La stratégie d'identification des caractéristiques individuelles associées au succès du CAI suit l'approche empirique et les modèles de régression décrits précédemment. La variable dépendante,  $Y_i$ , prend la valeur de 1 si le CAI a été conclu et la valeur de 0 si le CAI n'a pas été complété endéans le temps imparti. Le statut final du contrat (conclu ou non) est défini selon l'information connue en juin 2022, indépendamment du temps mis par le signataire à conclure les trois composantes du contrat. Pour cette partie de l'analyse limitée aux signataires, nous pouvons utiliser deux caractéristiques socio-démographiques additionnelles disponibles uniquement dans la base MiFa (mais pas dans la base IGSS) : le niveau d'éducation et le statut de séjour du signataire. Ces informations, auto-déclarées, ne sont néanmoins disponibles que pour un sous-groupe des signataires. En plus des variables démographiques et socio-professionnelles, contemporaines à la signature et passées, nous contrôlons donc également pour le niveau d'éducation des signataires et leur titre de séjour dans les quatre modèles. La composition des autres groupes de variables de contrôle suit la structure détaillée ci-avant, à l'exception des secteurs d'activités qui sont regroupés en trois catégories : services, non-services et inconnu, dû à la forte concentration des signataires employés dans les secteurs de services.

### 3. Analyse de la participation au CAI

Dans cette section, nous comparons les signataires du CAI aux non-signataires sur base de différentes caractéristiques. Dans la Section 3.1, nous décrivons les statistiques univariées tandis que dans la Section 3.2 nous utilisons quatre modèles

économétriques différents permettant de contrôler simultanément pour les différentes caractéristiques de chaque individu.

### 3.1. Profil des signataires : statistiques univariées

Le Tableau 1 compare les caractéristiques des 6.625 personnes ayant signé le Contrat d'Accueil et d'Intégration (CAI) au Luxembourg entre 2017 et 2022 avec celles des 262.010 non-signataires. L'analyse porte sur plusieurs variables que nous présentons sous différentes catégories: sexe, âge, canton de résidence, état civil, nationalité et statut professionnel. Les résultats sont présentés sous forme de moyennes et d'écart types, et les différences de moyennes entre les groupes sont évaluées à l'aide d'un test de comparaison de moyennes (le test de Student). Le tableau montre la différence entre les moyennes et le niveau de significativité.<sup>6</sup>

Il y a plus de femmes parmi les signataires du CAI (59%) que parmi les non-signataires (47%). La proportion de personnes âgées de 25 à 39 ans est nettement plus élevée parmi les signataires du CAI (70%) que parmi les non-signataires (41%). À l'inverse, la proportion de personnes âgées de plus de 45 ans n'est que de 13% parmi les signataires du CAI, contre 37% parmi les non-signataires. Seule la proportion de personnes âgées de 40 à 44 ans est similaire entre les deux groupes. Près de deux tiers des signataires (65%) sont mariés, contre 53% des non-signataires, bien que la taille des ménages soit légèrement plus faible parmi les signataires (-0.2 personne en moyenne).

Près de deux tiers des signataires (64%) résident dans le canton de Luxembourg (et 48% à Luxembourg-Ville), alors que ce canton héberge 42% de la population des non-signataires. Par contre, la proportion de non-signataires résidant dans le canton d'Esch-sur-Alzette (28%) est plus importante que celle des signataires (19%). Tous les autres cantons, notamment ceux du nord du pays, abritent également une part légèrement moins importante de signataires que de non-signataires.

L'incitation à signer le CAI semble moindre pour les ressortissants de l'Union Européenne. En effet, alors que 8 non-signataires sur 10 (82%) sont originaires de l'UE, cette proportion n'est que de 38% parmi les signataires. Cette différence est particulièrement marquée chez les résidents originaires du Portugal, d'Espagne et d'Italie, qui représentent 42% des non-signataires, mais uniquement 15% des signataires. Inversement, les ressortissants de pays tiers (non UE) sont surreprésentés parmi les signataires (62%). En particulier, les ressortissants d'Afrique représentent 20% des signataires, alors qu'ils ne constituent que 5% des non-signataires. Il en va de même pour les ressortissants d'Asie (17% des signataires

---

<sup>6</sup> La significativité statistique des différences augmente avec les \* : \* pour 5%, \*\* pour 1% et \*\*\* pour 0.1%.

contre 4% des non-signataires) et dans une moindre mesure d'Amérique du Sud (8% des signataires contre 2% des non-signataires) et des anciennes républiques soviétiques (10% des signataires contre 2% des non-signataires).

Le Tableau 2 détaille le nombre de signataires et non-signataires par nationalité dans la population analysée. En nombres absolus, le CAI a été signé principalement par des ressortissants d'Inde (564), du Portugal (416), d'Italie (404), de Russie (343) et de France (334). Ces cinq principaux pays d'origine représentent ainsi plus de 31% des 6.625 signataires sur la période 2017-2022. La colonne (4) du Tableau 2 divise pour chaque nationalité le nombre de signataires par le nombre d'individus éligibles. Les taux de participation ainsi obtenus sont les plus élevés pour les ressortissants de Turquie (21.6%), d'Égypte (21.4%), de Russie (18.1%), du Pakistan (17.3%) et d'Inde (14.9%). A l'inverse, les ressortissants de nombreux pays européens, dont les principaux pays d'origine des immigrants au Luxembourg, ont des taux de participation très faibles. Les cinq nationalités qui ont les taux de participation les plus faibles sont en effet la Belgique (0.4%), les Pays-Bas (0.5%), le Portugal (0.5%), l'Allemagne (0.7%) et la France (0.8%).

La proportion de salariés est identique entre les signataires et les non-signataires du CAI (59%). Cependant, le salaire horaire réel des signataires est plus élevé de 7% en moyenne (2.44€). Seuls 11% des signataires sont ouvriers alors qu'ils représentent 25% des non-signataires. Ces différences s'observent également au niveau des secteurs d'activité, les signataires étant moins actifs dans la construction (2% contre 12% pour les non-signataires) et plus actifs dans le secteur de la finance et des assurances, les activités d'information et de communication et les activités scientifiques et techniques (55% contre 32% pour les non-signataires).

Par ailleurs, les signataires du CAI sont plus souvent co-assurés (21% contre 11% parmi les non-signataires), et moins souvent dans un statut « autre », comprenant les inactifs, les pensionnés ou encore les fonctionnaires européens (13% contre 20% parmi les non-signataires). Les signataires ont généralement une durée de résidence plus courte (environ 3 années contre 6 années pour les non-signataires), ce qui pourrait suggérer que la décision de signer le CAI est généralement prise assez tôt après l'arrivée au Luxembourg. Cette plus courte durée de séjour pourrait en partie expliquer également la plus faible proportion du temps passé en emploi depuis 2010 observée parmi les signataires.

En somme, les signataires du CAI au Luxembourg sont généralement plus jeunes, plus souvent des femmes, plus probable d'être employé ou co-assuré, résident près de la capitale et sont originaires de pays hors UE.

## 3.2. Profil des signataires : modèles économétriques

Dans cette section, nous utilisons les quatre modèles économétriques, enrichis successivement avec des variables de contrôle additionnelles, décrits dans la Section 2. Les coefficients sont estimés sur la population cible totale.

Le Tableau 3 montre les caractéristiques observables corrélées avec la propension à signer le CAI. Le coefficient associé à une variable s'interprète comme la variation dans la probabilité de signer le CAI induite par l'augmentation du niveau de cette variable d'une unité, en maintenant les autres variables constantes (hypothèse 'toutes choses égales par ailleurs'). Les valeurs des caractéristiques de références sont les suivantes : sexe masculin, âge entre 35 et 39 ans, marié, habitant Luxembourg-Ville, n'étant pas né à Luxembourg, ayant une nationalité européenne francophone (France ou Belgique dans les modèles 1 à 3; France dans le modèle 4), salarié, travaillant dans le secteur des activités financières et d'assurance et ayant signé le CAI en 2017. La propension moyenne à signer le CAI pour un individu ayant les caractéristiques de référence, représenté par la constante du Tableau 3, se situe entre 3.1% et 3.5% selon le modèle utilisé. Les quatre modèles donnent des résultats très similaires au niveau de la valeur des paramètres estimés (coefficients et écarts-types). Nous nous référons dès lors principalement au modèle 4 (qui contrôle pour les caractéristiques professionnelles et la nationalité détaillée) dans le descriptif qui suit et concentrons notre analyse sur les variables statistiquement significatives.

Les femmes ont une propension à signer le CAI supérieure de l'ordre de 0.6 point de pourcentage (p.p.) par rapport à celle des hommes. L'âge influence la propension à signer le CAI ; ainsi, les personnes âgées de 20 à 24 ans ont une probabilité de signer le CAI de l'ordre de 1.7 p.p. inférieure à celles âgées entre 35 et 39 ans. Les individus âgés entre 30 à 34 ans ont une propension légèrement supérieure à signer le CAI comparé au groupe de référence. Enfin, les individus âgés de 40 ans ou plus ont une probabilité de signer le CAI plus faible que le groupe de référence. La différence est de l'ordre de 1.2 p.p. pour les individus âgés de plus de 50 ans.

Le canton de résidence est corrélé à la participation au CAI. Les résidents du canton de Luxembourg (capitale et campagne) ont une propension plus grande à signer que celle des résidents d'autres cantons du pays. A la suite du canton de Luxembourg, viennent les cantons de Capellen, Grevenmacher et Remich, tous situés géographiquement autour du canton de Luxembourg, où la propension à signer le CAI est environ 0.5 p.p. inférieure à celle observée à Luxembourg-Ville. La propension à signer diminue avec la distance par rapport à la capitale. Dans le canton d'Esch-sur-Alzette, elle est de -0.8 p.p. tandis que dans les cantons du nord, Clervaux, Wiltz et Vianden, ainsi que Redange, la différence est de l'ordre de 1.3 à 1.5 p.p..

La structure familiale est également corrélée à la propension à signer le CAI. On observe d'abord que les personnes célibataires sont moins susceptibles de signer le CAI que les personnes mariées (-1 p.p.). En parallèle, la taille du ménage est négativement corrélée à la propension à signer le CAI alors que le statut de co-assuré y est positivement associé (+0.7 p.p.).

La propension à signer le CAI diffère selon l'origine, mesurée à travers le pays de naissance et la nationalité. On note en premier que les personnes nées au Luxembourg sont moins susceptibles de signer le CAI comparé aux personnes nées à l'étranger. En effet, très peu de signataires sont nés au Luxembourg, ce qui peut paraître intuitif étant donné que les personnes grandissant dans le pays sont naturellement confrontées aux langues et aux institutions locales. La nationalité des individus est également corrélée à la propension à signer le CAI. Le modèle 3 indique que les individus originaires de pays non-UE ont en générale une propension significativement plus élevée à signer le CAI que les individus originaires des pays européens francophones. Les allemands ont une propension à signer le CAI similaires aux européens francophones tandis que les autres ressortissants européens (anglophones, portugais, espagnols, italiens et autres) ont une propension plus élevée à signer (entre 0.7 p.p. et 2.4 p.p.). Les ressortissants d'Europe de l'Est ont une propension à signer de 3 p.p. plus élevée. Les ressortissants d'Afrique (à l'exception de l'Afrique lusophone) ainsi que d'Amérique du Sud ont une propension à signer le CAI d'environ 8 p.p. plus élevée que le groupe de référence. Cette différence atteint presque 11 p.p. pour les ressortissants d'Asie du Sud. La langue parlée dans le pays d'origine ne semble néanmoins pas être le seul déterminant de la signature. En effet, parmi les ressortissants d'Afrique, par exemple, on ne note qu'un écart modeste (0.7 p.p.) de participation entre les personnes originaires de l'Afrique francophone et celles de l'Afrique anglophone.

Le modèle 4 permet de mettre en avant les pays dont les ressortissants se distinguent par une forte propension de leurs ressortissants à signer le CAI. Les ressortissants d'Égypte, d'Inde, du Pakistan, des Philippines, de la Russie, du Cameroun, de la Turquie et du Nigeria ont ainsi une propension à signer le CAI de plus de 10 p.p. supérieure à celle des français (le groupe de référence du modèle 4).

La propension à signer est également corrélée à l'expérience socio-professionnelle au Grand-Duché. Les nouveaux arrivants sont ainsi plus susceptibles de signer le CAI puisque la durée de séjour depuis 2010 est négativement corrélée à la propension à signer. Les personnes recevant le revenu minimum garanti (-0.6 p.p.) et celles ayant un statut autre (-1.4 p.p.) ont une propension plus faible à signer le CAI que les individus employés. En revanche, les personnes au chômage n'ont pas de propension significativement différente aux employés à signer le CAI une fois les autres

caractéristiques de l'individu prises en compte. Le fait d'être indépendant est associé à une probabilité de signature légèrement plus élevée tandis que l'inverse est observé pour les ouvriers. L'absence de CDI et le salaire horaire ne sont en revanche pas corrélés à la probabilité de signer le CAI. Enfin, le nombre de mois passés sans salaire depuis 2010 est négativement corrélé à la propension de participer au CAI.

Au niveau des secteurs d'activités, on note une propension significativement plus faible à signer le CAI parmi les travailleurs des secteurs de l'agriculture, de la construction, ou encore de l'hébergement et de la restauration. Par rapport au secteur de la finance et des assurances, qui sert de référence, certains secteurs nécessitant une main d'œuvre qualifiée, comme le secteur de l'information et de la communication, sont corrélés à une participation plus élevée (+1.4 p.p.) tandis que d'autres, comme les activités spécialisés, scientifiques et techniques (- 0.5 p.p.) ou encore la santé humaine et les activités sociales (-0.5 p.p.), sont associés à une participation plus faible. Au niveau de l'évolution de la participation dans le temps, on note une reprise de la signature du CAI en 2022, après une baisse en 2020, probablement liée à la crise du COVID-19.

## 4. Profil des signataires ayant obtenu le certificat et de ceux n'ayant pas complété le CAI

Dans cette section, nous comparons les signataires ayant conclu le CAI à ceux n'ayant pas terminé les trois composantes du contrat. L'analyse se concentre donc sur la probabilité de conclure le CAI, conditionnelle au fait d'avoir signé le contrat. Étant donné que les signataires ont, en général, au moins 3 ans pour accomplir toutes les conditions, nous nous concentrons dans cette partie de l'analyse sur les signataires ayant signé le CAI entre début 2017 et fin 2020. Dans la Section 4.1 nous décrivons les statistiques univariées tandis que dans la Section 4.2 nous utilisons quatre modèles économétriques différents permettant de contrôler simultanément pour les différentes caractéristiques individuelles.

### 4.1. Propension à obtenir le certificat : statistiques univariées

Le Tableau 4 compare les caractéristiques des personnes ayant rempli les conditions du CAI avec celles des personnes ayant commencé le contrat sans le conclure. L'analyse, identique à celle décrite à la Section 3.1, porte sur plusieurs variables que nous présentons sous différentes catégories: sexe, âge, canton de résidence, état civil, nationalité et statut professionnel. De plus, deux variables additionnelles, l'éducation

et le type de titre de séjour détenu, sont disponibles pour la population des signataires. Les résultats sont présentés sous forme de moyennes et d'écart types, et les différences de moyennes entre les groupes sont évaluées à l'aide d'un test de comparaison de moyennes (le test de Student). Le tableau montre la différence entre les moyennes et le niveau de significativité.

Sur les 4.306 signataires ayant signé le CAI entre 2017 et 2020, 2.067 ont obtenu le certificat et 2.239 n'ont pas conclu les trois composantes requises (tel que reporté fin juin 2022), soit un taux de réussite de l'ordre de 48%. La proportion de femmes est légèrement plus importante parmi les personnes ayant terminé le CAI (62%) que parmi celles qui ne l'ont pas conclu (60%), mais la différence n'est pas statistiquement significative. Une personne sur 5 (21%) ayant terminé le CAI a moins de 30 ans, alors que 30% des personnes n'ayant pas terminé sont dans cette tranche d'âge. En revanche, 72% des personnes ayant terminé le CAI ont entre 30 et 49 ans, alors que cette tranche d'âge ne représente que 62% parmi les personnes n'ayant pas conclu le CAI. Concernant l'état civil, la proportion de personnes mariées est significativement plus élevée parmi ceux ayant terminé le CAI (70%) que parmi ceux qui ne l'ont pas terminé (62%). Inversement, la proportion de célibataires est significativement plus faible parmi ceux ayant terminé le CAI (26%) que parmi ceux qui ne l'ont pas terminé (32%). La taille des ménages est légèrement plus élevée parmi ceux ayant terminé le CAI (2,49 personnes en moyenne) que parmi ceux qui ne l'ont pas terminé (2,36 personnes en moyenne).

La proportion de résidents dans le canton d'Esch-sur-Alzette est légèrement plus faible parmi ceux ayant terminé le CAI (17%) que parmi ceux qui ne l'ont pas terminé (20%) tandis que l'inverse s'observe à Luxembourg-Ville (49% parmi ceux qui ont terminé et 46% parmi ceux qui ne l'ont pas conclu). Ces différences ne sont toutefois que marginalement significatives d'un point de vue statistique. La durée de résidence est légèrement plus faible pour les signataires ayant terminé le CAI (2.73 années) que pour ceux n'ayant pas terminé le programme (2.9 années).

Les personnes ayant terminé le CAI sont moins souvent originaires de l'UE, en particulier du Portugal (4% contre 11%), de l'Espagne (3% contre 5%) et de l'Italie (5% contre 9%). À l'inverse, les proportions de ressortissants de pays tiers ayant terminé le CAI sont plus élevées que parmi ceux ne l'ayant pas terminé, en particulier pour les personnes originaires d'Afrique (19% contre 14%), d'Asie (17% contre 14%) et des anciennes républiques soviétiques (13% contre 6%). La colonne (5) du Tableau 2 fournit pour chaque nationalité le taux de réussite pour les signataires de la période 2017-2020. Les nationalités ayant le taux de réussite le plus élevé sont : le Liban (83.3%), la Russie (70.5%), la Turquie (70%), la Tunisie (69.6%) et l'Ukraine (63.7%). A l'inverse, les signataires originaires d'Afghanistan (20%), du Nigéria (21.1%), des

pays UE15 : Autres (24.3%), d'Érythrée (25.0%) et du Portugal (27.1%) ont la propension la plus faible à conclure le CAI.

Le statut professionnel n'est que faiblement corrélé à la probabilité de terminer le CAI. Parmi les signataires ayant un emploi, quelques différences ressortent néanmoins. La proportion d'ouvriers est significativement plus faible parmi les personnes ayant terminé le CAI (9%) que parmi celles ne l'ayant pas terminé (18%). Les personnes ayant terminé le CAI ont également un salaire horaire réel supérieur en moyenne (36,03€) que celles ne l'ayant pas terminé (28,90€), et ont davantage tendance à avoir un CDI (89% contre 84%).

En ce qui concerne l'éducation, une proportion significativement plus élevée de personnes ayant terminé le CAI possède un niveau d'éducation tertiaire (80%) par rapport à celles ne l'ayant pas terminé (69%). À l'inverse, la proportion de personnes avec un niveau d'éducation primaire et secondaire est plus faible parmi les personnes ayant terminé le CAI (1% et 11% respectivement) que parmi celles ne l'ayant pas terminé (2% et 19% respectivement).

Enfin, concernant le titre de séjour, on observe des différences significatives dans la répartition des types de visas. Les personnes ayant terminé le CAI sont plus susceptibles de détenir un visa "Carte Bleue" (8%) ou un visa "Salarié" (19%) que celles ne l'ayant pas terminé (4% et 15% respectivement). En revanche, les personnes n'ayant pas terminé le CAI sont plus susceptibles d'avoir un visa "Autorisation" (47%) que celles ayant terminé le programme (38%).

En somme, les personnes ayant terminé avec succès le CAI au Luxembourg sont généralement plus âgées, plus souvent mariées, résident plus souvent à Luxembourg-Ville et ont davantage tendance à être originaire d'un pays tiers. Elles ont généralement un niveau d'éducation plus élevé, ont un statut professionnel plus stable, et sont plus susceptibles de détenir des visas liés à l'emploi.

## 4.2. Propension à obtenir le certificat : modèles économétriques

Le Tableau 5 montre les caractéristiques observables des individus corrélés avec la propension à obtenir le certificat CAI. Les coefficients sont estimés sur l'ensemble des personnes ayant signé le CAI sur la période 2017-2020. L'individu de référence est un homme, âgé entre 35 et 39 ans, marié, habitant Luxembourg-Ville, n'étant pas né à Luxembourg, ayant une nationalité européenne francophone (France ou Belgique ; France dans le modèle 4), salarié, travaillant dans le secteur des services, ayant une éducation universitaire, ayant un visa « Autorisation », et ayant signé le CAI en 2017.



Le taux de réussite du groupe de référence se situe entre 50.1% et 51.1%, selon le modèle. Les quatre modèles donnent des résultats très similaires. Nous nous référons dès lors principalement au modèle 4 dans le descriptif qui suit.

Les femmes sont plus nombreuses à signer le CAI et, une fois que toutes les caractéristiques individuelles sont prises en compte, elles ont aussi une plus grande propension à obtenir le certificat (+4,2 p.p.). Comparés aux personnes âgées de 35 à 39 ans, les 20-29 ans ont une probabilité significativement plus faible de terminer le parcours du CAI (entre 11.4 et 15.8 p.p.). Par contre, les signataires âgés de 40 ans ont une probabilité de terminer le programme similaire au groupe de référence. L'état civil n'est pas corrélé à la probabilité de conclure avec succès le CAI.

A l'exception du canton de Wiltz (-20.9 points de pourcentage), la probabilité de terminer le CAI dans les autres cantons est similaire à celle des résidents de Luxembourg-Ville. L'exception que pose Wiltz est en partie liée au fait qu'il n'y a quasiment pas de signataires du CAI dans ce canton sur la période 2017-2022 (voir Tableau 1) et que leur propension à terminer le contrat a été faible. Si les nouveaux arrivants sont plus susceptibles de signer le CAI, on observe que la durée de séjour n'affecte pas systématiquement la probabilité de réussite. La probabilité de terminer le CAI augmente cependant avec le niveau d'études. Les signataires qui ont le niveau d'études primaires (-15.2 p.p.) et secondaires (-8.8 p.p.) sont moins susceptibles de terminer le CAI comparés à ceux qui ont un niveau d'études universitaires.

En lien avec le niveau d'études, on note que lorsque que l'on contrôle pour l'éducation, détenir une carte bleue européenne n'est pas corrélé à la propension de terminer le CAI. Cependant, il y a quelques différences pour les autres types de titre de séjour. Les détenteurs de visas de regroupement familial (UE et non-UE) sont moins susceptibles de terminer le CAI. Les détenteurs d'un statut de protection ont une propension plus faible à terminer le CAI dans le modèle 1, mais cette différence semble liée à des ressortissants de certains pays en particulier puisqu'elle se dissipe une fois que nous contrôlons pour le statut professionnel et les nationalités détaillées dans le modèle 4.

L'origine, mesurée à travers la nationalité, est corrélée à la propension à terminer le CAI. Les Portugais, les Espagnols et les Italiens sont moins susceptibles de terminer le CAI que les Belges et les Français. On note à l'inverse des propensions de succès significativement plus élevées pour les ressortissants des anciennes républiques soviétiques, de l'Afrique francophone, de l'Amérique du sud hispanophone et de l'Asie. Quelques pays se distinguent par une propension relativement élevée de leurs ressortissants à terminer le CAI. Il s'agit en particulier de l'Albanie, de la Bosnie-

Herzégovine, du Liban, du Maroc, de la Russie, de la Tunisie, de la Turquie et de l'Ukraine.

Au niveau du statut professionnel, le fait d'être au chômage, d'être co-assuré ou de recevoir le revenu minimum garanti est associé à une plus faible propension de terminer le CAI. Le statut d'ouvrier est associé à une baisse de 10 p.p du taux de réussite. En revanche, l'expérience professionnelle dans le pays ainsi que le secteur d'emploi ne sont pas significativement corrélés à la probabilité de conclure le contrat. Enfin, une légère baisse du taux d'obtention du certificat est observée pour les signataires de l'année 2020, ce qui pourrait être dû à la pandémie de COVID-19.

## 5. Analyse d'impact de la participation au CAI sur l'intégration socio-économique des signataires : une étude de faisabilité

Une évaluation rigoureuse des politiques publiques repose idéalement sur l'estimation causale de leurs impacts sur les décisions ou « résultats » des individus visés. Une liste détaillée de variables de résultat potentielles est présentée dans la Section 5.1. Cependant, afin d'illustrer nos propos de manière plus concrète, nous ferons régulièrement référence dans cette section à l'employabilité, qui est une des variables de résultat les plus fréquemment étudiées. Dans le cadre du CAI, il serait en effet particulièrement intéressant d'établir de manière causale si la participation au programme améliore ou non l'employabilité des participants, c'est-à-dire « toutes autres choses égales par ailleurs ».

La littérature économique a développé des méthodes d'évaluation variées permettant d'estimer des relations causales entre un traitement (par exemple, la participation à un programme public comme le CAI) et une variable de résultat (par exemple, l'employabilité d'un individu traité).

L'essai contrôlé randomisé (ECR) est la méthode scientifique de référence pour déterminer les effets causaux de la participation à un traitement. Dans une expérience contrôlée, la participation au traitement est répartie de manière aléatoire entre les participants à l'expérience. Le hasard étant le seul déterminant du traitement, les individus qui ont reçu le traitement "par chance" devraient être en moyenne similaires (ex-ante) aux individus qui n'ont pas reçu le traitement, pour autant que le nombre d'individus dans l'expérience soit suffisamment élevé. Le traitement étant ainsi la seule différence entre un individu traité et un individu non-traité, son effet causal peut être estimé. Toutefois, pour des raisons éthiques, il n'est pas toujours possible de mettre

en œuvre un ECR. Il peut en effet être considéré comme injuste que le traitement ne soit pas accessible à tous les individus, et en particulier aux personnes qui en ont potentiellement le plus besoin. Néanmoins, dans un scénario de ressources financières et humaines limitées, des projets pilotes associés à une évaluation ECR peuvent permettre d'évaluer s'il est souhaitable de généraliser un nouveau programme, potentiellement coûteux, à une population plus large.

Lorsqu'un ECR ne peut être mis en œuvre, il reste possible, dans certains contextes, d'aller au-delà des simples corrélations et d'estimer des relations causales, comme le montrent les travaux des lauréats du prix Nobel d'économie 2021, Card, Angrist et Imbens. Dans les études dites observationnelles, différentes méthodes d'estimation sont à la disposition des chercheurs en fonction du cadre de l'évaluation et du type d'aléa qui peut être observé dans l'allocation du traitement. En effet, les effets causaux ne peuvent être retrouvés en comparant simplement les résultats des individus participant au traitement à ceux des non-participants, car les individus traités seront probablement différents des non-participants au niveau de différentes caractéristiques.

D'une part, les différences peuvent concerner des caractéristiques observables qui pourraient être contrôlées si elles étaient observées (par exemple, l'âge). Ainsi, les individus plus jeunes pourraient être plus susceptibles de s'engager dans le CAI et également être plus employables que leurs aînés. Si l'on ne tient pas compte de la variable âge, on pourrait donc surestimer l'effet du CAI sur l'employabilité. Il est donc nécessaire de contrôler pour un maximum de caractéristiques observables dans une approche économétrique afin de prendre en compte ces effets de composition.

D'autre part, certaines caractéristiques sont par nature inobservables et pourraient affecter les résultats : par exemple, la motivation individuelle intrinsèque est difficilement mesurable. Ainsi, un signataire du CAI pourrait avoir plus de succès sur le marché de l'emploi simplement parce qu'il est plus motivé que d'autres personnes et non parce qu'il a participé au CAI. Dans le cadre des variables inobservables, il n'est par définition pas possible de les prendre en compte explicitement dans l'analyse.

Différentes approches ont été développées au cours du temps pour estimer l'effet causal d'un traitement en présence de ces deux types de différences. Ainsi, les estimateurs d'appariement sont basés sur la construction d'un groupe de contrôle non-traité (témoin) sur la base des caractéristiques observables dans les données, comme le sexe, l'âge ou l'expérience professionnelle antérieure qui est apparié au groupe traité. Ces méthodes garantissent que les personnes traitées et les témoins (faisant partie du groupe de contrôle non-traité) sont en moyenne similaires en ce qui concerne les caractéristiques observables dans les données. Cependant, elles ne peuvent pas traiter l'hétérogénéité des caractéristiques non-observées, telles que la motivation

intrinsèque à s'intégrer au Luxembourg. La méthode que nous proposons dans la Section 5.2 prend en compte cette hétérogénéité additionnelle possible.

Une évaluation causale des effets du CAI se ferait de façon ex-post, sur base de données collectées sur les participants à ce programme volontaire. L'évaluation doit se faire dans la limite des données disponibles, collectées auprès des signataires lors de la signature du contrat, complétées par des sources de données externes permettant de mesurer les variables de résultat.

## 5.1. Sources de données et variables d'intérêts

Pour identifier l'effet causal de la participation au CAI sur l'intégration socio-culturelle des participants, plusieurs types de données pourraient être mobilisées. Un certain nombre de variables de résultat sont disponibles dans les données administratives (emploi, revenu, logement, mobilité géographique, etc.). D'autres pourraient être basées sur des données d'enquêtes existantes (conditions de vie, force de travail, valeurs, etc.). Enfin, certaines dimensions pourraient nécessiter la collecte spécifique de nouvelles données originales, idéalement en dimension longitudinale (bien-être, acculturation, sentiments d'intégration ou d'appartenance, relations interculturelles, etc.).

Nous nous appuierons dans un premier temps sur les microdonnées (données individuelles pseudonymisées) provenant de la liaison de deux bases de données administratives décrites plus en détail à la Section 2. La base de données du MiFa contient l'information sur la date de signature du CAI et le cas échéant, de l'obtention du certificat. La base de données de l'IGSS fournit des données longitudinales sur les caractéristiques sociodémographiques des signataires et non-signataires ainsi que les variables d'intérêt liées à l'intégration socio-culturelle des individus.

Différentes dimensions pourraient être étudiées, à condition que l'information soit disponible et de qualité, comme variables de résultat :

- **Intégration professionnelle** : probabilité d'avoir un emploi (avec une possible distinction par type d'emploi : col blanc / ouvrier), salaire, probabilité d'avoir un contrat permanent, taille de l'entreprise dans laquelle l'individu est employé, secteur d'emploi, mobilité au sein de l'entreprise, changement d'emploi (d'entreprise), perte d'emploi, réception d'indemnité de chômage, réception du revenu d'intégration sociale (REVIS)...
- **Intégration socio-culturelle** : mariage inter-ethnique, naturalisation, mobilité géographique (au niveau cantonal)

La méthodologie que nous élaborons dans la Section 5.2 peut être appliquée pour évaluer l'impact de la signature (ou de la réalisation) du CAI sur ces différentes variables.

## 5.2. Evaluation causale du programme CAI sur la trajectoire socio-économique : méthodologie

Notre stratégie d'estimation est basée sur une méthode qui peut traiter la sélection sur les caractéristiques observables ainsi que sur les caractéristiques inobservables. L'objectif de cette méthode est d'estimer l'effet moyen du traitement sur les individus traités (average treatment on treated - ATT). Dans notre analyse, nous excluons du groupe de contrôle les unités qui n'ont jamais participé au CAI (jamais traitées) et nous comparerons les variables de résultats d'individus traités tôt (que nous appellerons CAI précoces) à ceux traités ultérieurement (que nous appellerons CAI tardifs).

Comme les signataires décident de participer volontairement au CAI, la participation ne peut être considérée comme une allocation exogène. Les signataires sont ainsi vraisemblablement différents des non-signataires, y compris au niveau de caractéristiques non-observables, comme par exemple la motivation intrinsèque à s'intégrer au Luxembourg. En exploitant le moment de signature du CAI, nous comparons des individus qui sont vraisemblablement plus similaires en termes de caractéristiques inobservables puisqu'ils participeront tous au CAI à un moment donné. Pour éliminer les différences non-observées restantes, nous tirerons avantage de la dimension longitudinale des données afin d'éliminer les différences dans les facteurs de confusion qui étaient présents avant la participation au traitement.

Nous utiliserons un estimateur de différence dans les différences (DiD), qui compare l'évolution d'une variable de résultat,  $Y$ , des individus traités ( $D=1$ ) avant et après la participation au traitement (les CAI précoces), aux individus du groupe de contrôle ( $D=0$ ) qui n'ont pas encore commencé le traitement à cette période (les CAI tardifs). En termes mathématiques, l'impact du traitement sur la variable de résultat,  $\hat{\beta}$ , estimé avec une approche DiD est écrit :

$$\hat{\beta} = E(Y_t - Y_{t-1} \mid D=1) - E(Y_t - Y_{t-1} \mid D=0).$$

La stratégie d'identification repose sur trois hypothèses fondamentales : (1) SUTVA, (2) absence d'anticipation, et (3) tendances parallèles. La première hypothèse, connue sous l'acronyme SUTVA qui signifie « valeur unitaire stable du traitement », implique que le traitement ne peut pas contaminer de manière directe ou indirecte le résultat du groupe témoin. La seconde, l'absence d'anticipation, signifie qu'il ne peut y avoir

d'anticipation du traitement par les individus traités, afin d'assurer que leur résultat n'ait pas été déjà affecté dans la période précédant le traitement. Enfin, la troisième hypothèse – l'hypothèse des tendances parallèles - stipule que la variable de résultat des deux groupes d'individus aurait continué à suivre une évolution parallèle si le groupe traité n'avait pas reçu le traitement. C'est grâce à cette hypothèse que les déviations de la tendance commune sont interprétées comme l'effet de la participation au traitement (voir Figure 1 pour une illustration graphique).

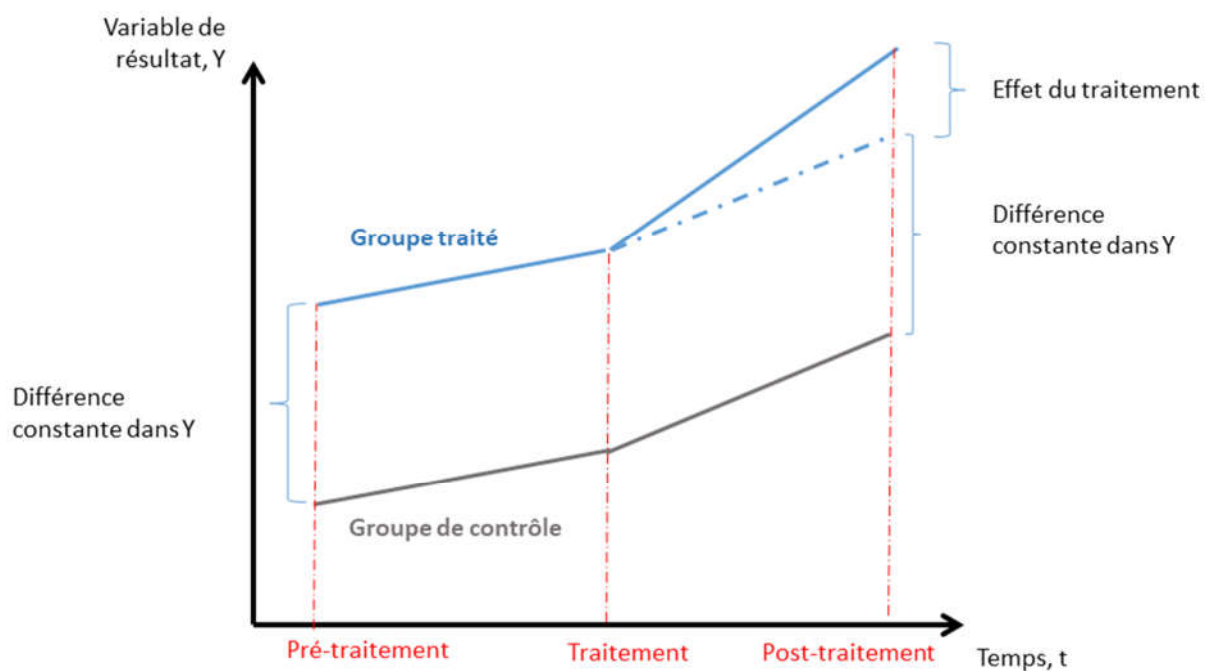


Figure 1 : méthode des différences dans les différences (DiD)

Comme la variable de résultat en l'absence de traitement,  $Y(0)$ , pendant la période de traitement ne peut pas être observée pour les individus traités (représenté par la droite bleue en pointillés dans la Figure 1), cette hypothèse d'identification ne peut pas être testée directement dans les données. Il est courant de s'appuyer sur des tests indirects, dits tests placebo, et de vérifier s'il existe une tendance commune dans les résultats observés pendant la période de prétraitement. Un échec des tests placebo peut émettre de sérieux doutes sur la crédibilité de l'hypothèse des tendances parallèles pendant la période de traitement.

Une façon possible d'assouplir l'hypothèse des tendances parallèles est d'exiger qu'elle ne soit valable que pour les individus partageant des caractéristiques observables similaires (notées  $X$  ci-dessous). Cette approche est basée sur un

estimateur CDiD (conditional-difference-in-differences), qui peut être mis en œuvre de plusieurs façons selon le cadre spécifique du traitement à analyser.<sup>7</sup>

Dans le contexte du CAI, nous sommes dans un cadre d'évaluation échelonnée puisque les individus peuvent signer le CAI à différents moments dans le temps ( $d$ ) et que le traitement est un état d'absorption (c'est-à-dire qu'une fois traité, un individu est traité pour toujours). Dans ce cadre, il est courant de s'appuyer sur des méthodes d'« étude d'événement » telles que l'estimateur à effets fixes à deux niveaux (individus et temps). Cette méthode est directement liée à l'estimateur DiD.

Cependant, la littérature économétrique récente a montré que les estimateurs à effets fixes à deux niveaux sont biaisés si l'effet du traitement varie dans le temps ou entre les groupes traités (voir par exemple Goodman-Bacon, 2021). Cela est dû au fait que les estimateurs à effets fixes à deux niveaux tirent parti de toutes les variations possibles dans les données pour estimer l'ATT. Ces estimateurs se basent sur de nombreuses comparaisons DiD et utilisent également des individus qui ont déjà été traités dans le passé comme l'un des groupes de contrôle. Cette comparaison est à proscrire car elle peut conduire à des estimations biaisées de l'ATT si les individus traités ont une réponse hétérogène au traitement et/ou si le traitement a des effets dynamiques dans le temps.

Pour prendre en compte ces préoccupations, nous utiliserons l'estimateur de différences dans les différences proposé par Callaway et Sant'Anna (2021), qui s'assure que seules des comparaisons DiD « valides » sont effectuées. En pratique, il divise les données en de nombreuses comparaisons de différences dans les différences, deux par deux, et obtient un effet de traitement moyen  $ATT(d,t)$  pour chaque groupe traité (défini comme toutes les unités entrant dans le traitement au moment  $d$ ) et la période calendaire  $t$ .<sup>8</sup> Chaque  $ATT(d,t)$  est estimé en utilisant comme groupe de contrôle les unités qui n'ont pas (encore) participé au traitement au moment  $t$  (les CAI tardifs).

Les estimations  $ATT(d,t)$  sont ensuite agrégées pour chaque période écoulée ( $e$ ) après l'entrée dans le traitement pour obtenir des estimations  $ATT(e)$  de l'étude d'événement. Cela permet d'observer la dynamique de l'effet moyen du traitement et de discerner les effets à court et à long terme.

Par ailleurs, nous isolerons l'éventuel effet d'anticipation du traitement (également connu dans la littérature économique sous le nom de Ashenfelter dip) en utilisant comme base de référence une période de prétraitement suffisamment éloignée de la date de participation au traitement  $d$ . Cet élément justifie de l'importance d'avoir un

---

<sup>7</sup> L'appendice technique détaille le CDiD dit doublement robuste de Sant'Anna et Zhao (2020).

<sup>8</sup> Les périodes  $t > d$  sont des périodes de traitement, tandis que  $t < d$  sont des périodes placebo.

historique suffisamment long des participants au CAI, avant même la signature du contrat. L'hypothèse sous-jacente est que dans  $\delta$  périodes antérieures, la variable de résultat traitée n'a pas encore commencé à réagir au futur traitement. Chaque ATT( $d, t$ ) est alors estimé par l'estimateur doublement robuste DiD de Sant'Anna et Zhao (2020), détaillé dans l'appendice technique.

### 5.3. Limites de l'approche empirique suggérée et extensions possibles

Comme détaillé dans les parties précédentes du rapport, l'évaluation des effets causaux du CAI est compliquée par le fait que les participants ne sont pas alloués au programme de façon exogène mais sont auto-sélectionnés puisqu'ils participent sur une base volontaire. Comme dans toutes les études d'observation, l'estimation n'est causale que si les hypothèses d'identification, qui ne sont pas toutes vérifiables, sont valides.

Dans notre contexte, la principale limitation potentielle de notre stratégie d'identification serait qu'un choc contemporain au traitement impacte simultanément la propension des individus à participer au CAI ainsi que leur variable de résultat. Cela peut se produire, par exemple, en cas de crise économique qui pourrait à la fois provoquer une augmentation de la probabilité de perte d'emploi ainsi que des incitants à entrer dans le CAI. Dans cette situation, les tests placebo ne montreraient pas de différences dans les tendances avant le traitement, mais l'estimation d'ATT serait confondue par le choc simultané sur la participation au traitement et sur la variable de résultat. De façon similaire, des attentes individuelles sur l'évolution future de la variable d'intérêt pourraient être systématiquement corrélées au traitement, sans qu'elles ne puissent être prises en compte par la méthode d'estimation. Ainsi, les individus pourraient signer le CAI avec l'espoir d'améliorer leurs perspectives professionnelles s'ils s'attendent systématiquement à ce que ces perspectives se détériorent. L'évolution négative de la trajectoire professionnelle des signataires du CAI par rapport aux signataires ultérieurs serait dans ce cas plutôt une matérialisation d'une sélection différée dans le traitement, liée aux anticipations personnelles, qu'un effet causal du programme sur la trajectoire professionnelle.

Le fait que plusieurs programmes et offres de cours existent en parallèle au CAI constitue un facteur de risque additionnel. Différents organismes publics et privés proposent des cours de langues et certains employeurs subventionnent même la participation de leurs employés. De plus, la participation au CAI et à d'autres formations, telles que celles proposées par exemple par l'ADEM, pourraient agir de



façon similaire sur les variables de résultat. Une évaluation causale du CAI nécessiterait donc idéalement de connaître l'historique complet des participations des signataires et des non-signataires du CAI à d'autres programmes de formations. La liaison de la base de donnée MiFa-IGSS à des sources additionnelles d'information, contenant l'information sur les cours pour adultes du Ministère de l'Education Nationale ou les formations de l'ADEM suivies, donnerait une image plus complète des programmes suivis par les signataires.

## 6. Conclusions et recommandations

Notre analyse, basée sur des sources de données administratives existantes, met en évidence l'auto-sélection de la population étrangère dans la participation au CAI en analysant les caractéristiques individuelles associées à la signature du contrat ainsi qu'à sa conclusion.

Globalement, seulement 2.5% de la population éligible a signé le CAI sur la période 2017-2022. Les signataires du CAI ont un profil démographique très sélectif. Ainsi, les femmes, les personnes âgées entre 30 et 40 ans, et mariées ont une plus forte propension à signer le CAI. La propension à signer est également plus élevée parmi les individus de nationalité non-européenne que parmi les Européens. En tête des nationalités ayant la propension la plus élevée à participer au CAI se retrouvent les ressortissants de Turquie, d'Égypte, de Russie, du Pakistan et d'Inde. A l'inverse, les ressortissants des pays d'origines traditionnels de l'immigration au Luxembourg, dont la France, l'Allemagne, la Belgique, les Pays-Bas, le Portugal et l'Italie, ont les taux de participation les plus faibles. Au niveau du statut professionnel, les salariés, en particulier ceux qui sont actifs dans des secteurs ayant une main d'œuvre hautement qualifiée, comme le secteur de la finance et des activités d'assurance ou l'information et communication, ont une propension à signer le CAI plus élevée que les personnes inactives, au chômage ou travaillant dans des secteurs comme la construction ou l'hébergement et la restauration. Une sélection s'observe également dans le lieu de résidence. En effet, les résidents de Luxembourg-Ville et de ces communes limitrophes ont une propension significativement plus élevée à participer au CAI. A l'inverse, les résidents des cantons du nord du pays, Wiltz, Clervaux et Vianden ont les propensions les plus faibles à signer le contrat.

Au niveau de la probabilité conditionnelle d'obtenir le certificat, c'est-à-dire de remplir les trois conditions du CAI conditionnel au fait d'avoir signé, certaines caractéristiques se distinguent également. Ainsi, les signataires ayant moins de 30 ans ont une probabilité significativement plus élevée de ne pas arriver au terme du CAI. Les

femmes ont en général une propension plus élevée à conclure le contrat tandis que l'état civil n'est pas associé au taux de réussite. Les individus ne possédant pas une éducation tertiaire ont des propensions à terminer le CAI avec succès significativement inférieure aux individus qui en possèdent une. Les signataires originaires des anciennes républiques soviétiques, de l'Afrique francophone ou encore du Moyen-Orient et Afrique du Nord ont les propensions les plus élevées à réussir le CAI. En particulier, les participants originaires du Liban, de la Tunisie, du Maroc, de Russie et de la Turquie se distinguent de façon particulièrement positive. À l'inverse, les signataires portugais, espagnols et italiens ont des propensions significativement plus faibles à conclure le CAI. Les individus au chômage, travaillant sous le statut d'ouvrier ou, dans une moindre mesure, ayant le statut de co-assuré ont une propension plus faible de réussite que les individus salariés. Au niveau de la résidence, une propension significativement plus faible à terminer le CAI est observé à Wiltz, tandis que le taux de réussite ne diffère guère entre les autres cantons. Les individus ayant un statut de séjour lié au regroupement familial (UE et non-UE) ont également une propension plus faible à conclure le CAI.

Nos résultats fournissent les premiers éléments de réflexions pour mettre en place une analyse causale de l'impact de la participation au CAI sur l'intégration socio-professionnelle des signataires. Notre rapport détaille les contours méthodologiques d'une telle analyse causale, en soulignant les avantages et les inconvénients d'une analyse empirique basée sur des données administratives ainsi que des méthodes alternatives potentiellement mobilisables, tel que l'essai contrôlé randomisé. En attendant l'approfondissement de l'étude des effets du CAI sur l'intégration, l'analyse des déterminants de la participation au contrat ainsi que de sa conclusion nous amène à formuler quatre types de recommandations.

**1. Réévaluer la communication et la sensibilisation au CAI** - Le profil des signataires du Contrat d'Accueil et d'Intégration est très spécifique. L'étude souligne plusieurs critères observables de cette auto-sélection de la population qui signe le CAI. Pour accroître la participation au CAI, il serait utile d'ajuster, de diversifier et de renforcer la communication et la sensibilisation auprès des sous-groupes de la population qui ont une faible propension à signer le CAI, notamment auprès des hommes, des célibataires, des individus âgés de plus de 40 ans, des ressortissants des pays de l'UE, des résidents de plus longue date, et des individus inactifs sur le marché du travail.

**2. Décentraliser l'offre de cours** - Notre analyse met en évidence une concentration géographique très marquée chez les participants. La participation au CAI est particulièrement concentrée autour de la capitale, mais faible à Esch-Sur-Alzette et dans les trois cantons du nord. Ces divergences pourraient être liées aux limites de la

communication et de la connaissance du programme dans ces régions. Elles pourraient également être affectées par une concentration élevée de l'offre de formations et de cours au centre du pays. Afin de faciliter l'accès aux différents modules du CAI, une décentralisation de l'offre de cours et de formations, notamment au sud et au nord du pays, pourrait être envisagée. La digitalisation progressive de certains modules pourrait également atténuer ces considérations d'ordre logistique.

**3. Réévaluer la structure et le contenu du programme** - L'analyse a démontré que seul un faible taux de signataires, en deçà de 50%, concluent le programme avec succès. Les données à notre disposition ne permettent pas de connaître la participation aux différentes composantes du contrat. Une connaissance plus détaillée des raisons des « abandons » permettrait de mieux cerner les leviers qui peuvent être activés pour réduire l'arrêt des participants au programme. Notamment, le faible taux de réussite pourrait être expliqué par une inadéquation du contenu du programme par rapport aux besoins des signataires. Une enquête visant à récolter les avis des signataires (aussi bien ceux qui ont conclu le CAI que ceux qui n'ont pas été à son terme) permettrait d'optimiser à la fois son contenu et renforcer la communication sur le programme.

Au vu des discussions actuelles d'une réforme de la loi sur le vivre-ensemble, avec un élargissement du public cible aux travailleurs frontaliers et aux résidents luxembourgeois, une différenciation de l'offre avec un renforcement de l'individualisation du parcours en fonction des besoins pourrait également augmenter la participation. Par exemple, un immigrant vivant depuis plusieurs années au Luxembourg connaît probablement déjà les institutions du pays, aussi son intérêt pour une journée d'orientation pourrait être très limité, ce qui pourrait le dissuader de participer au CAI. De façon similaire, les travailleurs frontaliers pourraient être intéressés par les cours de langues, mais moins par d'autres parties du programme, telles que l'instruction civique.

**4. Renforcer le suivi et l'évaluation des programmes d'intégration** - Notre analyse met en évidence les difficultés d'implémenter une évaluation ex-post des programmes d'intégration quand cette évaluation n'est pas planifiée en amont de la mise en œuvre des programmes à évaluer. Lorsque la participation à un programme est volontaire, l'auto-sélection des participants rend difficile la mesure de l'impact du programme. Il est alors important de mettre en place, dès le lancement du programme, un dispositif qui permette de tenir compte de la sélection des participants sachant que cette sélection peut être basée sur des facteurs observables ou non-observables. L'évaluation du CAI est sujette à ce type de problème et nous avons montré que la

sélection des participants est basée, en partie tout au moins, sur un ensemble de caractéristiques individuelles observables, qui peuvent être corrélées entre elles. Bien que des informations pertinentes soient disponibles dans les données administratives, elles ne permettent pas de rendre compte de manière exhaustive de certaines spécificités des immigrants, telles que la composition complète du ménage, du réseau familial, social et professionnel, la maîtrise des langues, le niveau préalable d'intégration, les compétences et domaines d'expertise, les traits de personnalités, les croyances, les attentes vis-à-vis du programme, ou les potentielles causes de l'abandon. Certaines de ces limitations pourraient être résolues par une collecte de cette information directement auprès des signataires.

Par ailleurs, certaines dimensions de l'étude mériteraient d'être approfondies. Premièrement, pour les participants n'ayant pas conclu le CAI, il serait pertinent d'analyser quels modules du CAI ils ont suivi avant d'abandonner. Deuxièmement, une analyse dynamique de la participation au CAI serait pertinente, dans la mesure où de nombreuses caractéristiques individuelles évoluent dans le temps (l'âge, la durée de séjour, la situation professionnelle et la structure familiale).

En conclusion, les résultats de cette analyse, soulignant un profil particulier et une auto-sélection marquée dans la signature du CAI, incitent à une réflexion sur le ciblage et la communication du programme, son contenu, son organisation ainsi que la mise en place d'un monitoring plus systématique.

Tableau 1: Statistiques Descriptives - Déterminants de la probabilité de signer le CAI

	Signataires CAI		Non-signataires CAI		Différence
	Moyenne	Ecart-type	Moyenne	Ecart-type	Différence
Femme	0.59	(0.49)	0.47	(0.50)	0.12***
<i>Age</i>					
Age 20-24	0.04	(0.20)	0.10	(0.31)	-0.06***
Age 25-29	0.21	(0.41)	0.14	(0.34)	0.07***
Age 30-34	0.29	(0.46)	0.14	(0.35)	0.15***
Age 35-39	0.20	(0.40)	0.13	(0.34)	0.07***
Age 40-44	0.12	(0.33)	0.12	(0.32)	0.00
Age 45-49	0.07	(0.25)	0.11	(0.31)	-0.04***
Age 50-54	0.03	(0.18)	0.10	(0.29)	-0.06***
Age 55-59	0.02	(0.13)	0.07	(0.26)	-0.05***
Age 60+	0.01	(0.12)	0.09	(0.29)	-0.08***
<i>Canton de résidence</i>					
Capellen	0.05	(0.21)	0.06	(0.23)	-0.01***
Clervaux	0.00	(0.06)	0.02	(0.15)	-0.02***
Diekirch	0.02	(0.13)	0.05	(0.21)	-0.03***
Echternach	0.01	(0.11)	0.02	(0.15)	-0.01***
Esch-sur-Alzette	0.19	(0.39)	0.28	(0.45)	-0.09***
Grevenmacher	0.03	(0.16)	0.03	(0.18)	-0.01**
Luxembourg-Campagne	0.16	(0.37)	0.10	(0.31)	0.06***
Luxembourg-Ville	0.48	(0.50)	0.32	(0.47)	0.16***
Mersch	0.03	(0.16)	0.04	(0.20)	-0.01***
Redange	0.00	(0.07)	0.02	(0.14)	-0.01***
Remich	0.02	(0.13)	0.03	(0.16)	-0.01***
Vianden	0.00	(0.04)	0.01	(0.08)	-0.00***
Wiltz	0.01	(0.09)	0.02	(0.15)	-0.01***
Résidence depuis 2010 (années)	3.06	(2.60)	5.92	(3.88)	-2.86***
<i>État civil</i>					
Célibataire	0.29	(0.46)	0.36	(0.48)	-0.07***
Marié	0.65	(0.48)	0.53	(0.50)	0.13***
Séparé/divorcé/veuf	0.04	(0.20)	0.09	(0.29)	-0.05***
Etat civil inconnu	0.01	(0.11)	0.02	(0.15)	-0.01***
Taille du ménage	2.38	(1.28)	2.57	(1.53)	-0.19***
Né à Luxembourg	0.00	(0.03)	0.05	(0.22)	-0.05***
<i>Statut professionnel</i>					
Revenu minimum garanti	0.02	(0.15)	0.03	(0.17)	-0.01***
Chômage	0.02	(0.14)	0.03	(0.16)	-0.01***
Co-assuré	0.21	(0.41)	0.11	(0.31)	0.10***
Autres (inactifs, fonc. EU, ...)	0.13	(0.34)	0.20	(0.40)	-0.07***
Salarié	0.59	(0.49)	0.59	(0.49)	-0.00
Fonctionnaire	0.00	(0.04)	0.00	(0.06)	-0.00**
Indépendant	0.02	(0.15)	0.04	(0.20)	-0.02***
<i>Caractéristiques de l'emploi et historique d'activité</i>					
Salaire réel horaire (dernier)	34.65	(60.86)	32.22	(83.78)	2.44*
Ouvrier	0.11	(0.32)	0.25	(0.43)	-0.13***
Sans CDI	0.13	(0.33)	0.12	(0.32)	0.01
Mesure d'activation	0.01	(0.10)	0.02	(0.13)	-0.01***
Emploi depuis 2010 (%)	0.52	(0.42)	0.60	(0.39)	-0.08***
Allocation chômage depuis 2010 (%)	0.03	(0.10)	0.05	(0.15)	-0.03***
Autre statut depuis 2010 (%)	0.45	(0.42)	0.35	(0.38)	0.11***

Suite sur la page suivante

Tableau 1 – suite

	Moyenne	Ecart-type	Moyenne	Ecart-type	Différence
<i>Secteur d'activité</i>					
Sect: Agriculture	0.00	(0.03)	0.01	(0.08)	-0.01***
Sect: Indust. extractive	0.00	(0.01)	0.00	(0.02)	-0.00
Sect: Indust. manufact.	0.03	(0.18)	0.04	(0.20)	-0.01**
Sect: Prod./distr. électricité, gaz	0.00	(0.04)	0.00	(0.03)	0.00
Sect: Prod./distr. eau et assain.	0.00	(0.03)	0.00	(0.04)	-0.00*
Sect: Construction	0.02	(0.15)	0.12	(0.32)	-0.10***
Sect: Commerce et rép. automobiles	0.07	(0.25)	0.09	(0.29)	-0.03***
Sect: Transports et entreposage	0.02	(0.14)	0.04	(0.18)	-0.01***
Sect: Hébergement et restauration	0.08	(0.27)	0.10	(0.30)	-0.02***
Sect: Information et communication	0.13	(0.34)	0.05	(0.21)	0.08***
Sect: Act. financ. et d'assurance	0.18	(0.38)	0.13	(0.33)	0.05***
Sect: Activités immobilières	0.01	(0.08)	0.01	(0.10)	-0.00***
Sect: Act. spéc., scienti. et techn.	0.24	(0.42)	0.14	(0.34)	0.10***
Sect: Act. serv. administratifs	0.09	(0.29)	0.11	(0.31)	-0.02***
Sect: Admin. publique	0.01	(0.12)	0.03	(0.17)	-0.01***
Sect: Enseignement	0.04	(0.21)	0.02	(0.14)	0.02***
Sect: Santé humaine, act. sociale	0.03	(0.17)	0.06	(0.24)	-0.03***
Sect: Arts et act. récréatives	0.00	(0.07)	0.01	(0.09)	-0.00***
Sect: Autres services	0.01	(0.12)	0.02	(0.13)	-0.00
Sect: Act. ménage en tant qu'employeur	0.02	(0.15)	0.03	(0.18)	-0.01***
Sect: Act. extra-territoriales	0.00	(0.05)	0.00	(0.06)	-0.00
<i>Nationalité</i>					
UE: Français	0.06	(0.24)	0.23	(0.42)	-0.17***
UE: Allemand	0.01	(0.11)	0.04	(0.20)	-0.03***
UE: Anglais	0.03	(0.16)	0.02	(0.14)	0.01***
UE: Portugais	0.06	(0.24)	0.32	(0.46)	-0.25***
UE: Espagnol	0.03	(0.18)	0.02	(0.15)	0.01***
UE: Italien	0.06	(0.24)	0.08	(0.27)	-0.02***
UE: Pays de l'Est	0.10	(0.30)	0.06	(0.23)	0.04***
UE: Autres	0.03	(0.18)	0.05	(0.23)	-0.02***
Anciennes républiques soviétiques	0.10	(0.29)	0.02	(0.12)	0.08***
Afrique: Français	0.09	(0.28)	0.02	(0.15)	0.06***
Afrique: Anglais	0.02	(0.14)	0.00	(0.07)	0.01***
Afrique: Portugais	0.01	(0.09)	0.01	(0.11)	-0.00***
Afrique: Autres	0.08	(0.27)	0.02	(0.14)	0.06***
USA-Océanie	0.02	(0.15)	0.01	(0.10)	0.01***
Amerique du Sud: Portugais	0.04	(0.20)	0.01	(0.10)	0.03***
Amerique du Sud: Espagnol	0.04	(0.20)	0.01	(0.10)	0.03***
Asie: Sud	0.10	(0.30)	0.02	(0.13)	0.09***
Asie: Reste	0.07	(0.25)	0.02	(0.15)	0.04***
Pays Balkan	0.05	(0.21)	0.04	(0.19)	0.01***
Autres (inconnu)	0.00	(0.04)	0.00	(0.02)	0.00
<i>Année</i>					
Année 2017	0.14	(0.34)	0.19	(0.39)	-0.06***
Année 2018	0.21	(0.41)	0.16	(0.37)	0.05***
Année 2019	0.25	(0.43)	0.16	(0.36)	0.09***
Année 2020	0.05	(0.22)	0.16	(0.37)	-0.11***
Année 2021	0.17	(0.37)	0.19	(0.39)	-0.03***
Année 2022	0.18	(0.39)	0.14	(0.34)	0.05***
Observations	6625		262010		268635

Tableau 2: Participation et réussite du CAI par nationalité

Nationalité	Signataires CAI	Non-signataires	Taux de participation	Taux de réussite
Inde	564	3215	14.9	52.9
Portugal	416	82541	0.5	27.1
Italie	404	20466	1.9	34.1
Russie	343	1557	18.1	70.5
France	334	43198	0.8	48.5
Amérique	288	2393	10.7	54.7
Brésil	265	2576	9.3	36.4
Turquie	249	905	21.6	70.0
Chine	243	3510	6.5	50.6
Roumanie	221	5241	4.0	55.5
Espagne	216	6308	3.3	33.5
Pologne	176	4326	3.9	44.2
Ukraine	170	1616	9.5	63.7
Royaume-Uni	169	4894	3.3	56.2
Maroc	166	1434	10.4	56.3
Afrique-Francophone	140	1705	7.6	52.7
Etats-Unis	133	2235	5.6	49.4
Anciennes républiques soviétiques	124	758	14.1	52.3
UE10 Pays de l'Est	116	2870	3.9	41.4
Grèce	98	2811	3.4	43.3
Asie du Sud/ Sud-Est	95	1429	6.2	54.9
Cameroun	88	672	11.6	50.0
Tunisie	87	1060	7.6	69.6
Afrique-Anglophone	86	899	8.7	55.0
Philippines	85	588	12.6	42.9
Serbie	80	1933	4.0	48.2
Allemagne	79	11225	0.7	54.0
Syrie	76	1613	4.5	39.7
Belgique	71	17299	0.4	38.0
Bulgarie	70	1251	5.3	54.5
Albanie	67	775	8.0	61.1
Egypte	66	243	21.4	61.3
Hongrie	59	1364	4.1	51.3
Pakistan	59	282	17.3	57.9
Bosnie-Herzégovine	57	1903	2.9	55.3
Monténégro	56	3128	1.8	42.6
UE15 Autres	55	6057	0.9	24.3
Maurice	54	567	8.7	60.0
Iran	50	539	8.5	55.2
Moyen-Orient Afrique du Nord	47	297	13.7	63.2
Croatie	46	928	4.7	50.0
Algérie	43	485	8.1	50.0
Cap-Vert	42	2660	1.6	30.3
Nigéria	41	270	13.2	21.1
Iraq	37	645	5.4	51.6
Liban	33	318	9.4	83.3
Asie de l'Est	31	752	4.0	47.1
Kosovo	25	1257	2.0	31.6
Macédoine	23	390	5.6	60.0
Pays-Bas	17	3195	0.5	36.4
Océanie	13	271	4.6	57.1
Europe Autres	12	1095	1.1	100.0
Erythrée	12	1008	1.2	25.0
Afghanistan	11	444	2.4	20.0
Autres	9	161	5.3	42.9
Afrique-Lusophone	8	414	1.9	50.0
Yugoslavie (inconnu)	0	34	0	.

Notes: Les colonnes (2) et (3) donnent le nombre de signataires et de non-signataires par nationalité détaillée. La colonne (4) calcule le taux de participation par nationalité comme le ratio des signataires sur l'ensemble de la population éligible (signataires + non-signataires). La colonne (5) calcule le taux de réussite par nationalité comme le nombre de signataires ayant obtenu le certificat sur le nombre de signataires.

Tableau 3: Déterminants de la probabilité de signer le CAI

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Femme	0.008*** [0.01,0.01]	0.006*** [0.00,0.01]	0.006*** [0.00,0.01]	0.006*** [0.00,0.01]
<i>Age; référence: Age 35-39</i>				
Age 20-24	-0.015*** [-0.02,-0.01]	-0.020*** [-0.02,-0.02]	-0.019*** [-0.02,-0.02]	-0.017*** [-0.02,-0.01]
Age 25-29	-0.003 [-0.01,0.00]	-0.004** [-0.01,-0.00]	-0.004** [-0.01,-0.00]	-0.003 [-0.01,0.00]
Age 30-34	0.007*** [0.00,0.01]	0.006*** [0.00,0.01]	0.007*** [0.00,0.01]	0.007*** [0.00,0.01]
Age 40-44	-0.005*** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]	-0.004*** [-0.01,-0.00]
Age 45-49	-0.011*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]
Age 50-54	-0.014*** [-0.02,-0.01]	-0.013*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]
Age 55-59	-0.014*** [-0.02,-0.01]	-0.012*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]	-0.011*** [-0.01,-0.01]
Age 60+	-0.016*** [-0.02,-0.01]	-0.012*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]	-0.013*** [-0.02,-0.01]
<i>Canton de résidence; référence: Lux.-Ville</i>				
Capellen	-0.004** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]
Clervaux	-0.017*** [-0.02,-0.01]	-0.016*** [-0.02,-0.01]	-0.016*** [-0.02,-0.01]	-0.015*** [-0.02,-0.01]
Diekirch	-0.013*** [-0.02,-0.01]	-0.013*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]
Echternach	-0.011*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]	-0.008*** [-0.01,-0.00]
Esch-sur-Alzette	-0.010*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]	-0.008*** [-0.01,-0.01]
Grevenmacher	-0.006*** [-0.01,-0.00]	-0.006*** [-0.01,-0.00]	-0.006*** [-0.01,-0.00]	-0.004** [-0.01,-0.00]
Luxembourg-Campagne	0.005*** [0.00,0.01]	0.004** [0.00,0.01]	0.004** [0.00,0.01]	0.004*** [0.00,0.01]
Mersch	-0.009*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]	-0.009*** [-0.01,-0.01]	-0.007*** [-0.01,-0.00]
Redange	-0.015*** [-0.02,-0.01]	-0.014*** [-0.02,-0.01]	-0.014*** [-0.02,-0.01]	-0.013*** [-0.02,-0.01]
Remich	-0.008*** [-0.01,-0.00]	-0.008*** [-0.01,-0.00]	-0.007*** [-0.01,-0.00]	-0.006*** [-0.01,-0.00]
Vianden	-0.017*** [-0.02,-0.01]	-0.016*** [-0.02,-0.01]	-0.015*** [-0.02,-0.01]	-0.013*** [-0.02,-0.01]
Wiltz	-0.016*** [-0.02,-0.01]	-0.015*** [-0.02,-0.01]	-0.015*** [-0.02,-0.01]	-0.013*** [-0.02,-0.01]
Résidence depuis 2010 (années)	-0.002*** [-0.00,-0.00]	-0.002*** [-0.00,-0.00]	-0.002*** [-0.00,-0.00]	-0.002*** [-0.00,-0.00]
<i>État civil; référence: Marié</i>				
Célibataire	-0.013*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]
Séparé/divorcé/veuf	-0.006*** [-0.01,-0.00]	-0.002* [-0.00,-0.00]	-0.001 [-0.00,0.00]	-0.002* [-0.00,-0.00]
Etat civil inconnu	-0.026***	-0.020***	-0.019***	-0.020***

Suite sur la page suivante



Tableau 3 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Taille du ménage	[-0.03,-0.02] -0.002***	[-0.02,-0.02] -0.001***	[-0.02,-0.02] -0.001***	[-0.02,-0.02] -0.001***
Né à Luxembourg	[-0.00,-0.00] -0.004*** [-0.00,-0.00]	[-0.00,-0.00] -0.007*** [-0.01,-0.01]	[-0.00,-0.00] -0.007*** [-0.01,-0.01]	[-0.00,-0.00] -0.007*** [-0.01,-0.01]
<i>Nationalité regroupée; référence: UE: Français</i>				
UE: Allemand	0.002* [0.00,0.00]	0.002* [0.00,0.00]	0.002* [0.00,0.00]	
UE: Anglais	0.024*** [0.02,0.03]	0.023*** [0.02,0.03]	0.021*** [0.02,0.03]	
UE: Portugais	0.007*** [0.01,0.01]	0.011*** [0.01,0.01]	0.012*** [0.01,0.01]	
UE: Espagnol	0.022*** [0.02,0.03]	0.024*** [0.02,0.03]	0.024*** [0.02,0.03]	
UE: Italien	0.013*** [0.01,0.01]	0.014*** [0.01,0.02]	0.015*** [0.01,0.02]	
UE: Pays de l'Est	0.027*** [0.02,0.03]	0.029*** [0.03,0.03]	0.030*** [0.03,0.03]	
UE: Autres	0.007*** [0.00,0.01]	0.007*** [0.00,0.01]	0.007*** [0.00,0.01]	
Anciennes républiques soviétiques	0.119*** [0.11,0.13]	0.118*** [0.11,0.13]	0.118*** [0.11,0.13]	
Afrique: Français	0.075*** [0.07,0.08]	0.077*** [0.07,0.08]	0.077*** [0.07,0.08]	
Afrique: Anglais	0.082*** [0.07,0.10]	0.084*** [0.07,0.10]	0.084*** [0.07,0.10]	
Afrique: Portugais	0.015*** [0.01,0.02]	0.020*** [0.01,0.02]	0.020*** [0.02,0.02]	
Afrique: Autres	0.078*** [0.07,0.09]	0.086*** [0.08,0.09]	0.085*** [0.08,0.09]	
USA-Océanie	0.038*** [0.03,0.05]	0.037*** [0.03,0.05]	0.037*** [0.03,0.05]	
Amerique du Sud: Portugais	0.078*** [0.07,0.09]	0.079*** [0.07,0.09]	0.079*** [0.07,0.09]	
Amerique du Sud: Espagnol	0.091*** [0.08,0.10]	0.090*** [0.08,0.10]	0.090*** [0.08,0.10]	
Asie: Sud	0.110*** [0.10,0.12]	0.110*** [0.10,0.12]	0.108*** [0.10,0.12]	
Asie: Reste	0.052*** [0.05,0.06]	0.051*** [0.04,0.06]	0.051*** [0.05,0.06]	
Pays Balkan	0.028*** [0.02,0.03]	0.031*** [0.03,0.03]	0.031*** [0.03,0.04]	
Autres (inconnu)	0.047** [0.01,0.08]	0.056*** [0.02,0.09]	0.056*** [0.02,0.09]	
<i>Statut professionnel; référence: Salarié</i>				
Revenu minimum garanti		-0.017*** [-0.02,-0.01]	-0.017*** [-0.02,-0.01]	-0.006** [-0.01,-0.00]
Chômage		-0.003 [-0.01,0.00]	-0.002 [-0.01,0.00]	-0.001 [-0.00,0.00]
Co-assuré		0.009*** [0.01,0.01]	0.009*** [0.01,0.01]	0.007*** [0.00,0.01]
Autres (inactifs, fonc. EU, ...)		-0.016*** [-0.02,-0.01]	-0.015*** [-0.02,-0.01]	-0.014*** [-0.02,-0.01]
Fonctionnaire		-0.001	0.000	0.000

Suite sur la page suivante

Tableau 3 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Indépendant		[-0.01,0.01] 0.003* [0.00,0.01]	[-0.01,0.01] 0.004** [0.00,0.01]	[-0.01,0.01] 0.005*** [0.00,0.01]
<i>Caractéristiques de l'emploi et historique d'activité</i>				
Salaire réel horaire (dernier)		-0.000 [-0.00,0.00]	-0.000 [-0.00,0.00]	-0.000 [-0.00,0.00]
Ouvrier		-0.010*** [-0.01,-0.01]	-0.007*** [-0.01,-0.01]	-0.006*** [-0.01,-0.00]
Sans CDI		0.001 [-0.00,0.00]	0.001 [-0.00,0.00]	0.002 [-0.00,0.00]
Mesure d'activation		-0.007** [-0.01,-0.00]	-0.007* [-0.01,-0.00]	-0.005 [-0.01,0.00]
Sans salaire depuis 2010		-0.011*** [-0.01,-0.01]	-0.012*** [-0.01,-0.01]	-0.013*** [-0.01,-0.01]
Autre statut depuis 2010 (%)		0.015*** [0.01,0.02]	0.014*** [0.01,0.02]	0.016*** [0.01,0.02]
Allocation chômage depuis 2010 (%)		-0.004* [-0.01,-0.00]	-0.004* [-0.01,-0.00]	-0.002 [-0.01,0.00]
<i>Secteur d'activité</i>				
<i>référence: Sect: Act. financ. et d'assurance</i>				
Sect: Agriculture			-0.011*** [-0.01,-0.01]	-0.011*** [-0.02,-0.01]
Sect: Indust. extractive			0.008 [-0.01,0.02]	0.006 [-0.01,0.02]
Sect: Indust. manufact.			0.002 [-0.00,0.01]	0.001 [-0.00,0.00]
Sect: Prod./distr. électricité, gaz			0.017 [-0.00,0.04]	0.015 [-0.01,0.04]
Sect: Prod./distr. eau et assain.			0.001 [-0.01,0.01]	0.000 [-0.01,0.01]
Sect: Construction			-0.005*** [-0.01,-0.00]	-0.005*** [-0.01,-0.00]
Sect: Commerce et rép. automobiles			-0.003* [-0.01,-0.00]	-0.004** [-0.01,-0.00]
Sect: Transports et entreposage			-0.002 [-0.01,0.00]	-0.003 [-0.01,0.00]
Sect: Hébergement et restauration			-0.006*** [-0.01,-0.00]	-0.007*** [-0.01,-0.00]
Sect: Information et communication			0.016*** [0.01,0.02]	0.014*** [0.01,0.02]
Sect: Activités immobilières			-0.005 [-0.01,0.00]	-0.006* [-0.01,-0.00]
Sect: Act. spéc., scienti. et techn.			-0.002 [-0.01,0.00]	-0.005** [-0.01,-0.00]
Sect: Act. serv. administratifs			-0.002 [-0.00,0.00]	-0.003 [-0.01,0.00]
Sect: Admin. publique			-0.002 [-0.01,0.00]	-0.002 [-0.01,0.00]
Sect: Enseignement			0.003 [-0.00,0.01]	0.001 [-0.00,0.01]
Sect: Santé humaine, act. sociale			-0.004** [-0.01,-0.00]	-0.005** [-0.01,-0.00]
Sect: Arts et act. récréatives			-0.012*** [-0.02,-0.01]	-0.013*** [-0.02,-0.01]

Suite sur la page suivante

Tableau 3 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Sect: Autres services			-0.004	-0.005
			[-0.01,0.00]	[-0.01,0.00]
Sect: Act. ménage en tant qu'employeur			-0.005**	-0.007***
			[-0.01,-0.00]	[-0.01,-0.00]
Sect: Act. extra-territoriales			-0.008	-0.008*
			[-0.02,0.00]	[-0.02,-0.00]
Sans secteur d'emploi			0.001	0.003
			[-0.00,0.00]	[-0.00,0.01]
<i>Nationalité détaillée; référence: France</i>				
Afghanistan				0.020**
				[0.01,0.03]
Albanie				0.069***
				[0.05,0.09]
Algérie				0.071***
				[0.05,0.09]
Allemagne				0.002*
				[0.00,0.00]
Belgique				0.001
				[-0.00,0.00]
Bosnie-Herzégovine				0.028***
				[0.02,0.04]
Brésil				0.079***
				[0.07,0.09]
Bulgarie				0.040***
				[0.03,0.05]
Cap-Vert				0.018***
				[0.01,0.02]
Cameroun				0.103***
				[0.08,0.13]
Chine				0.050***
				[0.04,0.06]
Croatie				0.037***
				[0.02,0.05]
Egypte				0.189***
				[0.14,0.23]
Erythrée				-0.001
				[-0.01,0.01]
Espagne				0.024***
				[0.02,0.03]
Etats-Unis				0.040***
				[0.03,0.05]
Grèce				0.018***
				[0.01,0.02]
Hongrie				0.027***
				[0.02,0.04]
Inde				0.122***
				[0.11,0.13]
Iran				0.072***
				[0.05,0.09]
Iraq				0.050***
				[0.03,0.07]
Italie				0.015***
				[0.01,0.02]
Kosovo				0.017***
				[0.01,0.02]

Suite sur la page suivante

Tableau 3 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Liban				0.076*** [0.05,0.11]
Macédoine				0.050*** [0.03,0.07]
Maroc				0.089*** [0.07,0.10]
Maurice				0.068*** [0.05,0.09]
Monténégro				0.018*** [0.01,0.02]
Nigéria				0.118*** [0.08,0.16]
Pakistan				0.149*** [0.11,0.19]
Pays-Bas				0.005** [0.00,0.01]
Philippines				0.111*** [0.09,0.14]
Pologne				0.031*** [0.02,0.04]
Portugal				0.011*** [0.01,0.01]
Roumanie				0.029*** [0.02,0.03]
Royaume-Uni				0.022*** [0.02,0.03]
Russie				0.158*** [0.14,0.17]
Serbie				0.035*** [0.03,0.04]
Syrie				0.038*** [0.03,0.05]
Tunisie				0.058*** [0.04,0.07]
Turquie				0.195*** [0.17,0.22]
Ukraine				0.075*** [0.06,0.09]
Yugoslavie (inconnu)				0.001 [-0.00,0.01]
Afrique-Anglophone				0.072*** [0.05,0.09]
Afrique-Francophone				0.068*** [0.06,0.08]
Afrique-Lusophone				0.019** [0.01,0.03]
Amérique				0.091*** [0.08,0.10]
Asie de l'Est				0.018* [0.00,0.03]
Asie du Sud/ Sud-Est				0.047*** [0.03,0.06]
UE10 Pays de l'Est				0.027*** [0.02,0.03]
Europe Autres				0.002 [-0.00,0.01]

Suite sur la page suivante

Tableau 3 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
UE15 Autres				-0.000 [-0.00,0.00]
Moyen-Orient Afrique Nord				0.116*** [0.08,0.15]
Océanie				0.025* [0.00,0.05]
Anciennes républiques soviétiques				0.120*** [0.10,0.14]
Autres				0.050** [0.02,0.08]
<i>Année; référence: Année 2017</i>				
Année 2018	0.015*** [0.01,0.02]	0.015*** [0.01,0.02]	0.015*** [0.01,0.02]	0.015*** [0.01,0.02]
Année 2019	0.020*** [0.02,0.02]	0.020*** [0.02,0.02]	0.020*** [0.02,0.02]	0.020*** [0.02,0.02]
Année 2020	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]	-0.010*** [-0.01,-0.01]
Année 2021	-0.000 [-0.00,0.00]	-0.001 [-0.00,0.00]	-0.001 [-0.00,0.00]	-0.001 [-0.00,0.00]
Année 2022	0.001 [-0.00,0.00]	0.002 [-0.00,0.00]	0.002 [-0.00,0.00]	0.002* [0.00,0.00]
Constante	0.031*** [0.03,0.03]	0.035*** [0.03,0.04]	0.034*** [0.03,0.04]	0.032*** [0.03,0.04]
Observations	268635	268635	268635	268635
Notes: 95% intervalles de confiance entre parenthèses; * $p < 0.05$ , ** $p < 0.01$ , *** $p < 0.001$				

Tableau 4: Statistiques Descriptives - Déterminants de la probabilité d'obtenir le certificat

	Certificat CAI		Pas de certificat CAI		Différence
	Moyenne	Ecart-type	Moyenne	Ecart-type	Différence
Femme	0.62	(0.48)	0.60	(0.49)	0.03
<i>Age</i>					
Age 20-24	0.03	(0.17)	0.07	(0.25)	-0.04***
Age 25-29	0.18	(0.39)	0.23	(0.42)	-0.05***
Age 30-34	0.30	(0.46)	0.27	(0.44)	0.03
Age 35-39	0.21	(0.41)	0.18	(0.39)	0.03*
Age 40-44	0.14	(0.35)	0.11	(0.31)	0.03**
Age 45-49	0.07	(0.26)	0.06	(0.24)	0.01
Age 50-54	0.04	(0.19)	0.03	(0.18)	0.01
Age 55-59	0.02	(0.13)	0.02	(0.15)	-0.01
Age 60+	0.01	(0.12)	0.02	(0.14)	-0.01
<i>Canton de résidence</i>					
Capellen	0.05	(0.22)	0.04	(0.21)	0.01
Clervaux	0.00	(0.06)	0.00	(0.04)	0.00
Diekirch	0.02	(0.14)	0.02	(0.15)	-0.00
Echternach	0.02	(0.13)	0.01	(0.12)	0.00
Esch-sur-Alzette	0.17	(0.38)	0.20	(0.40)	-0.03*
Grevenmacher	0.03	(0.18)	0.03	(0.16)	0.01
Luxembourg-Campagne	0.16	(0.37)	0.17	(0.37)	-0.01
Luxembourg-Ville	0.49	(0.50)	0.46	(0.50)	0.03
Mersch	0.02	(0.16)	0.02	(0.15)	0.00
Redange	0.01	(0.08)	0.00	(0.06)	0.00
Remich	0.01	(0.12)	0.02	(0.13)	-0.00
Vianden	0.00	(0.04)	0.00	(0.05)	-0.00
Wiltz	0.00	(0.06)	0.01	(0.12)	-0.01***
Résidence depuis 2010 (années)	2.73	(2.40)	2.90	(2.77)	-0.18*
<i>État civil</i>					
Célibataire	0.26	(0.44)	0.32	(0.47)	-0.06***
Marié	0.70	(0.46)	0.62	(0.48)	0.08***
Séparé/divorcé/veuf	0.04	(0.19)	0.05	(0.21)	-0.01
Etat civil inconnu	0.01	(0.09)	0.01	(0.12)	-0.01
Taille du ménage	2.49	(1.31)	2.36	(1.28)	0.14***
Né à Luxembourg	0.00	(0.04)	0.00	(0.03)	0.00
<i>Statut professionnel</i>					
Revenu minimum garanti	0.02	(0.15)	0.04	(0.19)	-0.01*
Chômage	0.02	(0.13)	0.03	(0.16)	-0.01*
Co-assuré	0.25	(0.43)	0.22	(0.41)	0.03*
Autres (inactifs, fonc. EU, ...)	0.15	(0.36)	0.14	(0.34)	0.01
Salarié	0.53	(0.50)	0.56	(0.50)	-0.02
Fonctionnaire	0.00	(0.03)	0.00	(0.05)	-0.00
Indépendant	0.02	(0.16)	0.02	(0.15)	0.00
<i>Caractéristiques de l'emploi et historique d'activité</i>					
Salaire réel horaire (dernier)	36.03	(79.38)	28.90	(56.23)	7.13*
Ouvrier	0.09	(0.29)	0.18	(0.38)	-0.09***
Sans CDI	0.11	(0.32)	0.16	(0.36)	-0.04**
Mesure d'activation	0.01	(0.11)	0.01	(0.12)	-0.00
Emploi depuis 2010 (%)	0.47	(0.43)	0.48	(0.41)	-0.02
Allocation chômage depuis 2010 (%)	0.03	(0.10)	0.04	(0.12)	-0.01**
Autre statut depuis 2010 (%)	0.51	(0.43)	0.48	(0.41)	0.03*

Suite sur la page suivante

Tableau 4 – suite

	Moyenne	Ecart-type	Moyenne	Ecart-type	Différence
<i>Secteur d'activité</i>					
Secteur: Services	0.94	(0.24)	0.93	(0.26)	0.01
Secteur: non-Services	0.06	(0.24)	0.07	(0.26)	-0.01
<i>Nationalité</i>					
UE: Français	0.06	(0.25)	0.07	(0.25)	-0.00
UE: Allemand	0.01	(0.11)	0.01	(0.10)	0.00
UE: Anglais	0.03	(0.17)	0.02	(0.14)	0.01
UE: Portugais	0.04	(0.21)	0.11	(0.31)	-0.07***
UE: Espagnol	0.03	(0.17)	0.05	(0.22)	-0.02***
UE: Italien	0.05	(0.22)	0.09	(0.28)	-0.04***
UE: Pays de l'Est	0.11	(0.31)	0.10	(0.30)	0.00
UE: Autres	0.03	(0.17)	0.04	(0.20)	-0.01
Anciennes républiques soviétiques	0.13	(0.33)	0.06	(0.24)	0.06***
Afrique: Français	0.09	(0.29)	0.07	(0.25)	0.02**
Afrique: Anglais	0.01	(0.11)	0.01	(0.12)	-0.00
Afrique: Portugais	0.01	(0.08)	0.01	(0.11)	-0.01
Afrique: Autres	0.08	(0.27)	0.05	(0.22)	0.03***
USA-Océanie	0.02	(0.15)	0.02	(0.14)	0.00
Amerique du Sud: Portugais	0.03	(0.18)	0.06	(0.23)	-0.02***
Amerique du Sud: Espagnol	0.04	(0.20)	0.03	(0.18)	0.01
Asie: Sud	0.10	(0.30)	0.08	(0.27)	0.02
Asie: Reste	0.07	(0.25)	0.06	(0.25)	0.00
Pays Balkan	0.05	(0.22)	0.05	(0.22)	0.00
Autres (inconnu)	0.00	(0.04)	0.00	(0.04)	-0.00
<i>Education</i>					
Education primaire	0.01	(0.10)	0.02	(0.14)	-0.01**
Education secondaire	0.11	(0.31)	0.19	(0.40)	-0.09***
Education tertiaire	0.80	(0.40)	0.69	(0.46)	0.10***
Education inconnue	0.09	(0.28)	0.09	(0.29)	-0.00
<i>Titre de séjour</i>					
Visa: Autorisation	0.38	(0.48)	0.47	(0.50)	- 0.09***
Visa: Carte Bleue	0.08	(0.28)	0.04	(0.21)	0.04***
Visa: Regroupement familial (Lux.)	0.04	(0.19)	0.04	(0.19)	0.00
Visa: Regroupement familial (UE)	0.10	(0.30)	0.09	(0.29)	0.01
Visa: Regroupement familial (non-UE)	0.13	(0.34)	0.12	(0.33)	0.01
Visa: Privé	0.03	(0.17)	0.03	(0.18)	-0.00
Visa: Protection	0.03	(0.16)	0.03	(0.18)	-0.01
Visa: Emploi	0.19	(0.39)	0.15	(0.36)	0.04***
Visa: inconnu	0.02	(0.13)	0.02	(0.13)	0.00
<i>Année</i>					
Année 2017	0.20	(0.40)	0.22	(0.41)	-0.02
Année 2018	0.34	(0.47)	0.32	(0.47)	0.02
Année 2019	0.39	(0.49)	0.37	(0.48)	0.02
Année 2020	0.07	(0.25)	0.09	(0.29)	-0.02**
Observations	2067		2239		4306

Tableau 5: Déterminants de la probabilité d'obtenir le certificat

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Femme	0.047** [0.02,0.08]	0.045** [0.01,0.08]	0.045** [0.01,0.08]	0.042* [0.01,0.08]
<i>Age; référence: Age 35-39</i>				
Age 20-24	-0.153*** [-0.23,-0.08]	-0.156*** [-0.23,-0.08]	-0.154*** [-0.23,-0.08]	-0.158*** [-0.23,-0.08]
Age 25-29	-0.116*** [-0.16,-0.07]	-0.120*** [-0.17,-0.07]	-0.118*** [-0.17,-0.07]	-0.114*** [-0.16,-0.07]
Age 30-34	-0.031 [-0.07,0.01]	-0.033 [-0.08,0.01]	-0.032 [-0.08,0.01]	-0.032 [-0.08,0.01]
Age 40-44	0.047 [-0.01,0.10]	0.048 [-0.01,0.10]	0.048 [-0.00,0.10]	0.050 [-0.00,0.10]
Age 45-49	0.035 [-0.03,0.10]	0.034 [-0.03,0.10]	0.033 [-0.03,0.10]	0.034 [-0.03,0.10]
Age 50-54	0.068 [-0.02,0.16]	0.068 [-0.02,0.16]	0.066 [-0.02,0.15]	0.062 [-0.03,0.15]
Age 55-59	-0.033 [-0.14,0.07]	-0.023 [-0.13,0.09]	-0.026 [-0.14,0.08]	-0.039 [-0.15,0.07]
Age 60+	-0.046 [-0.17,0.07]	-0.062 [-0.18,0.06]	-0.068 [-0.19,0.05]	-0.072 [-0.19,0.05]
<i>Canton de résidence; référence: Lux.-Ville</i>				
Capellen	-0.003 [-0.07,0.07]	0.001 [-0.07,0.07]	0.002 [-0.07,0.07]	-0.006 [-0.07,0.06]
Clervaux	0.173 [-0.13,0.47]	0.175 [-0.11,0.46]	0.173 [-0.11,0.46]	0.179 [-0.11,0.47]
Diekirch	-0.014 [-0.12,0.09]	-0.012 [-0.12,0.09]	-0.012 [-0.12,0.09]	-0.015 [-0.12,0.09]
Echternach	0.051 [-0.07,0.17]	0.063 [-0.06,0.19]	0.062 [-0.06,0.19]	0.056 [-0.07,0.18]
Esch-sur-Alzette	-0.012 [-0.05,0.03]	0.001 [-0.04,0.04]	0.001 [-0.04,0.04]	0.005 [-0.04,0.05]
Grevenmacher	0.031 [-0.06,0.12]	0.038 [-0.05,0.13]	0.041 [-0.05,0.13]	0.051 [-0.04,0.14]
Luxembourg-Campagne	-0.036 [-0.08,0.01]	-0.036 [-0.08,0.01]	-0.036 [-0.08,0.01]	-0.034 [-0.08,0.01]
Mersch	-0.005 [-0.10,0.09]	-0.002 [-0.09,0.09]	-0.002 [-0.09,0.09]	0.004 [-0.09,0.10]
Redange	0.194* [0.00,0.38]	0.203* [0.01,0.40]	0.208* [0.01,0.40]	0.211* [0.02,0.40]
Remich	0.016 [-0.10,0.14]	0.025 [-0.09,0.14]	0.026 [-0.09,0.14]	0.027 [-0.09,0.15]
Vianden	-0.149 [-0.49,0.19]	-0.163 [-0.51,0.19]	-0.153 [-0.50,0.20]	-0.129 [-0.47,0.21]
Wiltz	-0.244*** [-0.37,-0.11]	-0.218*** [-0.35,-0.09]	-0.215** [-0.34,-0.09]	-0.209** [-0.34,-0.08]
Résidence depuis 2010 (années)	-0.009** [-0.02,-0.00]	-0.007* [-0.01,-0.00]	-0.006 [-0.01,0.00]	-0.005 [-0.01,0.00]
<i>État civil; référence: Marié</i>				
Célibataire	-0.013 [-0.06,0.03]	-0.011 [-0.06,0.04]	-0.010 [-0.06,0.04]	-0.008 [-0.05,0.04]
Séparé/divorcé/veuf	-0.046 [-0.13,0.03]	-0.037 [-0.12,0.04]	-0.037 [-0.12,0.04]	-0.040 [-0.12,0.04]
Etat civil inconnu	-0.084	-0.076	-0.074	-0.077

Suite sur la page suivante



Tableau 5 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Taille du ménage	[-0.22,0.06] 0.007	[-0.22,0.06] 0.008	[-0.21,0.07] 0.008	[-0.21,0.06] 0.010
Né à Luxembourg	[-0.01,0.02] 0.376* [0.00,0.75]	[-0.01,0.03] 0.359* [0.00,0.72]	[-0.01,0.02] 0.360* [0.00,0.72]	[-0.01,0.03] 0.360 [-0.00,0.72]
<i>Nationalité regroupée; référence: UE: Français</i>				
UE: Allemand	0.058 [-0.09,0.21]	0.050 [-0.10,0.20]	0.050 [-0.10,0.20]	
UE: Anglais	0.071 [-0.04,0.18]	0.057 [-0.06,0.17]	0.057 [-0.06,0.17]	
UE: Portugais	-0.150*** [-0.23,-0.07]	-0.129** [-0.21,-0.05]	-0.127** [-0.21,-0.05]	
UE: Espagnol	-0.140** [-0.23,-0.05]	-0.144** [-0.24,-0.05]	-0.145** [-0.24,-0.05]	
UE: Italien	-0.107** [-0.19,-0.03]	-0.094* [-0.18,-0.01]	-0.093* [-0.17,-0.01]	
UE: Pays de l'Est	0.019 [-0.06,0.09]	0.018 [-0.06,0.09]	0.017 [-0.06,0.09]	
UE: Autres	-0.051 [-0.15,0.05]	-0.051 [-0.15,0.05]	-0.052 [-0.15,0.05]	
Anciennes républiques soviétiques	0.233*** [0.14,0.32]	0.230*** [0.14,0.32]	0.228*** [0.14,0.32]	
Afrique: Français	0.199*** [0.11,0.29]	0.204*** [0.11,0.30]	0.204*** [0.11,0.30]	
Afrique: Anglais	0.078 [-0.07,0.22]	0.072 [-0.07,0.22]	0.074 [-0.07,0.22]	
Afrique: Portugais	0.052 [-0.12,0.22]	0.076 [-0.10,0.25]	0.074 [-0.10,0.25]	
Afrique: Autres	0.216*** [0.12,0.31]	0.218*** [0.12,0.32]	0.216*** [0.12,0.31]	
USA-Océanie	0.049 [-0.08,0.18]	0.043 [-0.09,0.17]	0.042 [-0.09,0.17]	
Amerique du Sud: Portugais	-0.012 [-0.11,0.09]	-0.005 [-0.11,0.10]	-0.006 [-0.11,0.10]	
Amerique du Sud: Espagnol	0.152** [0.04,0.26]	0.150** [0.04,0.26]	0.149** [0.04,0.26]	
Asie: Sud	0.113* [0.02,0.20]	0.113* [0.02,0.20]	0.110* [0.02,0.20]	
Asie: Reste	0.102* [0.01,0.20]	0.101* [0.01,0.20]	0.098* [0.00,0.19]	
Pays Balkan	0.160** [0.06,0.26]	0.163** [0.06,0.27]	0.162** [0.06,0.27]	
Autres (inconnu)	0.056 [-0.38,0.49]	0.053 [-0.37,0.48]	0.049 [-0.38,0.47]	
<i>Statut professionnel; référence: Salarié</i>				
Revenu minimum garanti		-0.130 [-0.27,0.00]	-0.158* [-0.30,-0.02]	-0.155* [-0.29,-0.02]
Chômage		-0.131* [-0.24,-0.02]	-0.129* [-0.24,-0.02]	-0.131* [-0.24,-0.02]
Co-assuré		-0.072 [-0.15,0.00]	-0.100* [-0.18,-0.02]	-0.105* [-0.19,-0.02]
Autres (inactifs, fonc. EU, ...)		-0.036 [-0.11,0.04]	-0.058 [-0.14,0.02]	-0.058 [-0.14,0.02]
Fonctionnaire		-0.193	-0.196	-0.216

Suite sur la page suivante

Tableau 5 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Indépendant		[-0.46,0.08] -0.055 [-0.16,0.05]	[-0.47,0.08] -0.064 [-0.17,0.05]	[-0.50,0.07] -0.062 [-0.17,0.05]
<i>Caractéristiques de l'emploi et historique d'activité</i>				
Salaire réel horaire (dernier)		0.000 [-0.00,0.00]	0.000 [-0.00,0.00]	0.000 [-0.00,0.00]
Ouvrier		-0.101*** [-0.15,-0.05]	-0.100*** [-0.15,-0.05]	-0.102*** [-0.15,-0.05]
Sans CDI		-0.043 [-0.11,0.02]	-0.035 [-0.10,0.03]	-0.037 [-0.10,0.03]
Mesure d'activation		0.033 [-0.15,0.22]	0.034 [-0.15,0.22]	0.058 [-0.13,0.24]
Sans salaire depuis 2010		0.029 [-0.03,0.09]	0.023 [-0.04,0.08]	0.025 [-0.03,0.08]
Autre statut depuis 2010 (%)		0.037 [-0.04,0.11]	0.019 [-0.06,0.10]	0.019 [-0.06,0.10]
Allocation chômage depuis 2010 (%)		-0.056 [-0.27,0.16]	-0.059 [-0.27,0.15]	-0.051 [-0.27,0.16]
<i>Sceteur d'activité; référence: Sect: Services</i>				
Sect: Autres			0.005 [-0.07,0.08]	-0.002 [-0.08,0.07]
Sans secteur d'emploi			0.056 [-0.01,0.12]	0.056 [-0.01,0.13]
<i>Nationalité détaillée; référence: France</i>				
Afghanistan				-0.137 [-0.39,0.12]
Albanie				0.211* [0.04,0.38]
Algérie				0.121 [-0.09,0.33]
Allemagne				0.025 [-0.13,0.18]
Belgique				-0.125 [-0.28,0.03]
Bosnie-Herzégovine				0.188* [0.02,0.35]
Brésil				-0.030 [-0.14,0.08]
Bulgarie				0.036 [-0.13,0.20]
Cap-Vert				0.016 [-0.17,0.20]
Cameroun				0.139 [-0.02,0.29]
Chine				0.079 [-0.03,0.19]
Croatie				0.011 [-0.17,0.20]
Egypte				0.163 [-0.03,0.36]
Erythrée				-0.125 [-0.50,0.25]
Espagne				-0.165***

Suite sur la page suivante

Tableau 5 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Etats-Unis				[-0.26,-0.07] 0.013
Grèce				[-0.12,0.15] -0.041
Hongrie				[-0.18,0.09] -0.005
Inde				[-0.17,0.16] 0.083
Iran				[-0.02,0.18] 0.105
Iraq				[-0.10,0.31] 0.154
Italie				[-0.09,0.40] -0.112*
Kosovo				[-0.20,-0.03] 0.036
Liban				[-0.21,0.29] 0.406**
Macédoine				[0.15,0.67] 0.195
Maroc				[-0.09,0.48] 0.181**
Maurice				[0.05,0.32] 0.158
Monténégro				[-0.03,0.35] 0.104
Nigéria				[-0.06,0.27] -0.162
Pakistan				[-0.37,0.05] 0.178
Pays-Bas				[-0.08,0.43] -0.103
Philippines				[-0.38,0.17] 0.017
Pologne				[-0.14,0.18] -0.055
Portugal				[-0.16,0.05] -0.147***
Roumanie				[-0.23,-0.06] 0.071
Royaume-Uni				[-0.03,0.18] 0.035
Russie				[-0.08,0.15] 0.249***
Serbie				[0.15,0.35] 0.079
Syrie				[-0.07,0.23] 0.004
Tunisie				[-0.20,0.20] 0.298***
Turquie				[0.15,0.45] 0.233***
Ukraine				[0.11,0.35] 0.185**
Afrique-Anglophone				[0.06,0.31] 0.130

Suite sur la page suivante

Tableau 5 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Afrique-Francophone				[-0.04,0.30] 0.142*
Afrique-Lusophone				[0.01,0.27] 0.206
Amérique				[-0.19,0.60] 0.124*
Asie de l'Est				[0.01,0.24] -0.035
Asie du Sud/ Sud-Est				[-0.30,0.23] 0.144
UE10 Pays de l'Est				[-0.01,0.29] -0.075
Europe Autres				[-0.20,0.05] 0.502***
UE15 Autres				[0.39,0.61] -0.278***
Moyen-Orient Afrique Nord				[-0.43,-0.13] 0.194*
Océanie				[0.02,0.37] 0.143
Anciennes républiques soviétiques				[-0.23,0.51] 0.091
Autres				[-0.06,0.24] -0.010 [-0.41,0.39]
<i>Education; référence: Education tertiaire</i>				
Education primaire	-0.181** [-0.30,-0.06]	-0.154* [-0.27,-0.03]	-0.150* [-0.27,-0.03]	-0.152* [-0.27,-0.03]
Education secondaire	-0.117*** [-0.16,-0.07]	-0.089*** [-0.14,-0.04]	-0.089*** [-0.14,-0.04]	-0.088*** [-0.14,-0.04]
Education inconnue	-0.034 [-0.09,0.02]	-0.021 [-0.08,0.03]	-0.021 [-0.08,0.03]	-0.024 [-0.08,0.03]
<i>Titre de séjour; référence: Visa: Autorisation</i>				
Visa: Carte Bleue	-0.003 [-0.08,0.08]	-0.010 [-0.09,0.07]	-0.009 [-0.09,0.08]	-0.014 [-0.10,0.07]
Visa: Regroupement familial (Lux.)	-0.099* [-0.20,-0.00]	-0.097 [-0.20,0.00]	-0.096 [-0.19,0.00]	-0.078 [-0.18,0.02]
Visa: Regroupement familial (UE)	-0.113** [-0.19,-0.04]	-0.110** [-0.18,-0.04]	-0.110** [-0.18,-0.04]	-0.107** [-0.18,-0.03]
Visa: Regroupement familial (non-UE)	-0.111** [-0.18,-0.04]	-0.110** [-0.18,-0.04]	-0.110** [-0.18,-0.04]	-0.104** [-0.18,-0.03]
Visa: Privé	-0.114* [-0.21,-0.01]	-0.119* [-0.22,-0.02]	-0.121* [-0.22,-0.02]	-0.110* [-0.21,-0.01]
Visa: Protection	-0.166** [-0.28,-0.05]	-0.119 [-0.24,0.01]	-0.124 [-0.25,0.00]	-0.009 [-0.18,0.16]
Visa: Emploi	-0.018 [-0.08,0.04]	-0.018 [-0.08,0.04]	-0.017 [-0.08,0.05]	-0.011 [-0.07,0.05]
Visa: inconnu	0.013 [-0.10,0.13]	0.009 [-0.11,0.13]	0.009 [-0.11,0.13]	0.016 [-0.10,0.13]
<i>Année; référence: Année 2017</i>				
Année 2018	0.010 [-0.03,0.05]	0.008 [-0.03,0.05]	0.008 [-0.03,0.05]	0.014 [-0.03,0.06]
Année 2019	0.018	0.013	0.012	0.014

Suite sur la page suivante

Tableau 5 – suite

	Modèle 1	Modèle 2	Modèle 3	Modèle 4
Année 2020	[-0.02,0.06] -0.061 [-0.12,0.00]	[-0.03,0.05] -0.066* [-0.13,-0.00]	[-0.03,0.05] -0.066* [-0.13,-0.00]	[-0.03,0.06] -0.064* [-0.13,-0.00]
Constante	0.504*** [0.41,0.60]	0.501*** [0.40,0.60]	0.501*** [0.40,0.60]	0.511*** [0.41,0.61]
Observations	4306	4306	4306	4306
Notes: 95% intervalles de confiance entre parenthèses; * $p < 0.05$ , ** $p < 0.01$ , *** $p < 0.001$				

## Références

Abadie, Alberto, 2005, Semiparametric Difference-in-Differences Estimators, *Review of Economic Studies*, 72, issue 1, pages 1-19.

Callaway, Brantly & Sant'Anna, Pedro H.C., 2021, Difference-in-Differences with multiple time periods, *Journal of Econometrics* vol. 225(2), pages 200-230.

Goodman-Bacon, Andrew, 2021, Difference-in-differences with variation in treatment timing, *Journal of Econometrics*, vol. 225(2), pages 254-277.

Heckman, James J. & Ichimura, Hidehiko & Todd, Petra, 1998, Matching As An Econometric Evaluation Estimator, *Review of Economic Studies*, vol. 65(2), pages 261-294.

Sant'Anna, Pedro H.C. & Zhao, June, 2020, Doubly robust difference-in-differences estimators, *Journal of Econometrics*, vol. 219(1), pages 101-122.

Wooldridge, Jeffrey M., 2013, *Introductory Econometrics: A Modern Approach*. 5th Edition, South-Western Pub, Mason.

# APPENDICE TECHNIQUE

Cet appendice technique donne d'abord une liste des variables utilisées dans le rapport ainsi que leurs définitions. Ensuite, nous détaillons le traitement des individus n'ayant aucune affiliation visible avec la sécurité sociale sur la période d'analyse. Enfin, nous détaillons l'estimateur doublement robuste DiD de Sant'Anna et Zhao (2020).

## **A.1 Définition des variables**

Nous définissons dans cette section toutes les variables (*en italique*) utilisées dans le rapport.

### **Variables démographiques**

- Sexe : *Femme* prend la valeur de 1 si l'individu est de sexe féminin (et 0 si l'individu est de sexe masculin). L'individu de référence est un homme.
- Âge: Pour des raisons d'anonymisation des données individuelles, la population analysée est regroupée dans les neuf catégories d'âge suivantes : *Age 20-24*, *Age 25-29*, *Age 30-34*, *Age 35-39*, *Age 40-44*, *Age 45-49*, *Age 50-54*, *Age 55-59*, *Age 60+*. Le groupe d'âge de référence dans nos régressions est *Age 35-39*.
- Lieu de résidence : Pour des raisons d'anonymisation des données individuelles, le canton de résidence est utilisé comme unité géographique. Le Luxembourg est administrativement divisé en douze cantons mais la base de données mise à notre disposition sépare le canton de Luxembourg entre la commune de Luxembourg-Ville (*Luxembourg-Ville*) et le reste des communes du canton (*Luxembourg-Campagne*). Les treize variables définissant le lieu de résidence sont: *Capellen*, *Clervaux*, *Diekirch*, *Echternach*, *Esch-sur-Alzette*, *Grevenmacher*, *Luxembourg-Campagne*, *Luxembourg.-Ville*, *Mersch*, *Redange*, *Remich*, *Vianden*, *Wiltz*. L'individu de référence habite à Luxembourg-Ville.
- Durée de résidence au Luxembourg : Afin de contrôler pour la durée de résidence dans le pays, nous construisons la variable *Résidence depuis 2010 (années)*, qui mesure la résidence au Luxembourg, en nombre d'années, depuis janvier 2010. La décision de se concentrer sur le temps passé au Luxembourg depuis 2010 est justifiée par le fait que la structure des données disponibles ne permet pas de différencier les entrées dans la base IGSS par les résidents de celles effectuées par un travailleur frontalier (non-éligible au CAI) avant cette date. Notre approche permet donc d'éviter de comptabiliser comme durée de résidence une période qui serait en réalité une période de travail frontalier.

- L'état civil : Quatre statuts possibles définissent l'état civil de l'individu : *Célibataire*, *Marié*, *Separé/divorcé/veuf* et *Etat civil inconnu*. L'individu de référence est *Marié*.
- Taille du ménage : *Taille du ménage* reprend le nombre d'individus ayant un lien d'affiliation commun à la sécurité sociale. Cette mesure ne capture qu'imparfaitement les ménages de fait et les ménages fiscaux. Ainsi, deux partenaires qui résident ensemble sans être co-habitants légaux ou mariés ne seront pas considérés comme constituant un ménage.
- Lieu de naissance : La variable binaire *Né à Luxembourg* prend la valeur de 1 si l'individu est né au Luxembourg et 0 autrement. Nous n'avons néanmoins pas accès au pays de naissance précis des individus nés hors Luxembourg. L'individu de référence est né hors de Luxembourg.

## Le statut professionnel

- Statut professionnel – non employés : A côté des différentes variables décrivant les individus en emploi (détaillés ci-après), la base IGSS permet de définir d'autres statuts professionnels: *Revenu minimum garanti* prend la valeur 1 si l'individu perçoit l'allocation sociale minimale (le REVIS ou RMG précédemment). *Chômage* prend la valeur 1 si l'individu perçoit une allocation de chômage. *Co-assuré* prend la valeur de 1 si l'individu est affilié à la sécurité sociale par l'intermédiaire d'un assuré principal et ne perçoit pas d'autres transferts. Enfin, la catégorie *Autres (inactifs, fonc. EU, ...)* reprend les autres statuts que nous ne pouvons différencier avec les données disponibles. Cette catégorie « résiduelle » inclut les inactifs ne percevant aucune indemnisation, les fonctionnaires européens et les travailleurs d'institutions internationales ayant un système de sécurité sociale séparé. Les individus pensionnés sont également repris dans cette catégorie. Enfin, notons que, dans les données, certains individus peuvent avoir de multiples statuts (par exemple, *Revenu minimum garanti* et *Chômage*) au cours d'un mois. Afin de faciliter l'analyse et d'éviter la démultiplication de combinaisons de statuts (relativement rares dans les données), nous définissons un statut principal par individu avec la priorisation suivante : *Emploi* > *Chômage* > *Revenu minimum garanti* > *Co-assuré* > *Autres statuts*. Ainsi, si durant un mois nous observons que l'individu a perçu une allocation de chômage et le revenu minimum garanti, nous lui octroyons le statut *Chômage* (et donc les variables *Chômage* =1 et *Revenu minimum garanti* =0).
- Statut professionnel – employés : Trois variables binaires capturent les différents statuts des employés : *Salarié*, *Fonctionnaire* (fonction publique luxembourgeoise uniquement) et *Indépendant*. Plusieurs variables additionnelles capturent les



conditions d'emploi : *Sans CDI* prend la valeur de 1 si le travailleur n'est pas sous contrat à durée indéterminée ; *Ouvrier* prend la valeur de 1 si l'individu a un travail manuel. *Mesure d'activation* prend la valeur de 1 si le travailleur bénéficie d'une mesure d'activation. Enfin, plusieurs variables caractérisent l'historique professionnel de l'individu, observé depuis 2010. *Emploi depuis 2010* (%) mesure la proportion du temps (en pourcentage des mois) pendant lequel l'individu est observé comme étant employé dans la base IGSS depuis 2010 (ou depuis son entrée dans la base si elle est ultérieure). De façon similaire, *Allocation de chômage depuis 2010* (%) mesure la proportion du temps (en pourcentage des mois) pendant lequel l'individu est employé dans la base IGSS depuis 2010. *Autre statut depuis 2010* (%) mesure la proportion du temps (en pourcentage des mois) pendant lequel l'individu est observé sans emploi, allocation de chômage, co-affiliation ou revenu minimal dans la base IGSS depuis 2010. L'individu de référence est salarié avec CDI.

- Le dernier salaire horaire réel connu de l'individu, mesuré par la variable *Salaire réel horaire (dernier)*, donne une indication de son niveau de salaire réel horaire, ajusté pour l'inflation (en déflétant le salaire nominal par l'indice des prix à la consommation de 2015). Dans les régressions, cette variable prend la valeur de 0 si l'information est indisponible ou si la personne ne travaille pas. Par ailleurs, la variable binaire *Sans salaire depuis 2010* capture les individus pour lesquels aucun salaire n'a été observé depuis période 2010.

## **Le secteur d'activité**

- Le secteur d'activité est détaillé en vingt-et-un secteurs suivant la nomenclature NACE Rev. 2 : *Sect: Agriculture*, *Sect: Indust. Extractive*, *Sect: Indust. Manufact.*, *Sect: Prod./distr. électricité, gaz*, *Sect: Prod./distr. eau et assain.*, *Sect: Construction*, *Sect: Commerce et rép. Automobiles*, *Sect: Transports et entreposage*, *Sect: Hébergement et restauration*, *Sect: Information et communication*, *Sect: Act. financ. et d'assurance*, *Sect: Activités immobilières*, *Sect: Act. spéc., scienti. et techn.*, *Sect: Act. serv. Administratifs*, *Sect: Admin. Publique*, *Sect: Enseignement*, *Sect: Santé humaine, act. Sociale*, *Sect: Arts et act. Récréatives*, *Sect: Autres services*, *Sect: Act. ménage en tant qu'employeur*, *Sect: Act. extra-territoriales*. Les statistiques descriptives montrent la proportion d'employés par secteur conditionnel au fait que l'information soit disponible. Dans les régressions, la variable *Sans secteur d'emploi* prend en compte que cette information peut être manquante pour certains travailleurs mais capture également tous les individus non-employés (et donc sans secteur d'activité par définition). Dans l'analyse sur les déterminants de la signature du CAI, le secteur de référence est le secteur des activités financières et d'assurance (*Sect: Act. financ. et d'assurance*). Dans l'analyse sur les déterminants de l'accomplissement du CAI, nous agrégeons les secteurs d'activité en *Sect : Services*, et *Sect : Non-Services*

et *Sans secteur d'emploi* étant donné la taille plus limitée de la population d'analyse. Dans ce cas, le secteur de référence est le secteur *Sect : Services*.

### **Effet calendrier**

- Afin de contrôler pour des fluctuations liées à l'environnement socio-économique général, nous ajoutons un effet fixe par année dans nos régressions, qui permet de contrôler pour des éléments communs aux signataires endéans une même année. Selon le modèle estimé, jusqu'à 6 années sont prises en compte à travers les variables : *Année 2017, Année 2018, Année 2019, Année 2020, Année 2021* et *Année 2022*. L'année 2017 sert de référence dans toutes les analyses.

### **Variables définissant la nationalité**

Deux approches différentes sont utilisées pour prendre en compte la nationalité.

- La nationalité regroupée: La première approche regroupe les nationalités en dix-neuf groupes géographiques et linguistiques, avec un accent sur les langues les plus communément utilisées au Luxembourg : *UE: Français, UE: Allemand, UE: Anglais, UE: Portugais, UE: Espagnol, UE: Italien, UE: Pays de l'Est, UE: Autres, Anciennes républiques soviétiques, Afrique: Français, Afrique: Anglais, Afrique: Portugais, Afrique: Autres, USA-Océanie, Amérique du Sud: Portugais, Amérique du Sud: Espagnol, Asie: Sud, Asie: Reste, Pays Balkan, Autres (inconnu)*. La répartition des pays en ces dix-neuf groupes est détaillée dans la troisième colonne du Tableau A1 ci-dessous. Dans les analyses basées sur les modèles 1 à 3, la nationalité de référence est le groupe *UE : Français* (constitué de la Belgique et la France).
- La nationalité détaillée : Dans le modèle 4, nous contrôlons pour la nationalité individuelle détaillée. Nous avons accès aux quarante nationalités principales des signataires des CAI. Pour des raisons d'anonymisation, l'IGSS a regroupé les restantes nationalités en groupes géographiques et linguistiques que nous avons définis. Notre approche a été de créer des groupes relativement homogènes au niveau des continents ainsi que les langues parlées du pays, avec un accent sur les langues les plus représentées au Luxembourg. Les nationalités disponibles et la classification des autres pays dans les groupes agrégés sont détaillées dans la deuxième colonne du Tableau A1 ci-dessous. Dans les analyses basées sur les modèles 4, la nationalité de référence est la *française*.

### **Variables additionnelles dans la base de donnée MIFA**

- L'éducation : Les signataires déclarent leur niveau d'éducation lors de la signature du CAI. Nous regroupons cette information en quatre variables qui définissent le

niveau d'éducation le plus élevé fréquenté par chaque signataire (sans qu'un diplôme n'ait nécessairement été obtenu): *Education primaire* (qui reprend l'éducation primaire et préscolaire), *Education secondaire*, *Education tertiaire* et *Education inconnue*. Dans l'analyse des déterminants de la conclusion du CAI, l'éducation tertiaire sert de niveau de référence.

- Le titre de séjour : Les signataires indiquent également leur titre de séjour. Etant donné le faible nombre de certains titres dans la base de donnée MIFA, nous les regroupons en neuf catégories principales : *Visa: Autorisation* (citoyens UE), *Visa: Carte Bleue*, *Visa: Regroupement familial (Lux)*, *Visa: Regroupement familial (UE)*, *Visa: Regroupement familial (non-UE)*, *Visa: Vie Privé* (liens personnels avec un résidant hors catégories éligibles au regroupement familial ou ressources propres suffisantes), *Visa: Protection*, *Visa: Emploi* (reprend les visas pour les salariés et les indépendants) et *Visa: inconnu*. Dans l'analyse des déterminants de la conclusion du CAI, l'autorisation de séjour de référence est le Visa : Autorisation.

## **A.2 Individus ayant un statut particulier**

Parmi les signataires présents dans la base de données de l'IGSS, 296 individus (soit environ 4.5%) n'ont aucun lien observable avec la sécurité sociale du Grand-Duché sur la période d'observation. Ces individus n'ont pas été soumis aux cotisations sociales ni été enregistrés comme employé, co-assuré, ou bénéficiaire de transferts sociaux (chômage, revenu minimum garanti...). Ils ont donc des statuts particuliers (fonctionnaires des institutions européennes ou autres institutions internationales), sont des personnes pensionnées sur l'ensemble de la période, ou ont été absent du pays sur l'ensemble de la période d'analyse. Nous leur attribuons un statut professionnel « autre ». Bien qu'ils soient un sous-groupe particulier, nous les gardons dans l'analyse afin de ne pas réduire davantage la taille de notre échantillon de signataires. Nous leur attribuons un statut professionnel « autre ». Dans la population des non-signataires, près de 22% des individus ont ce statut « autre » sur l'ensemble de la période d'analyse. Afin d'éviter cette sur-représentation d'un groupe très particulier dans le groupe des non-signataires, nous les supprimant de la population étudiée. Enfin, nous contrôlons pour la part du temps passé dans ce statut « autre » pour tous les individus de la population étudiée. Nous montrons dans le Tableau A2 ci-dessous que les résultats principaux sont inchangés si nous excluons les 296 individus signataires concernés de l'analyse.

## **A.3 L'estimateur doublement robuste DiD de Sant'Anna et Zhao (2020)**

Le CDiD dit doublement robuste de Sant'Anna et Zhao (2020) s'applique dans un cadre où l'on compare le groupe traité avant et après un traitement à un moment donné au groupe de contrôle. L'ATT est donné par l'expression suivante :

$$\widehat{ATT} = E \left[ \left( \frac{D}{E[D]} - \frac{\frac{\widehat{p}(X)(1-D)}{1-\widehat{p}(X)}}{E\left[\frac{\widehat{p}(X)(1-D)}{1-\widehat{p}(X)}\right]}} \right) \left( Y_{after} - Y_{before} - \widehat{m}(X) \right) \right]$$

Cet estimateur combine deux approches différentes. La première approche est le « modèle de régression de résultat » (outcome regression model) proposée par Heckman et al. (1997). Cette approche se base sur deux étapes. Premièrement, une estimation de la tendance temporelle de Y pour des individus du groupe de contrôle ayant des caractéristiques observées X,  $\widehat{m}(X) = E[Y_{after} - Y_{before} | X, 1 - D = 1]$ . Deuxièmement, la tendance estimée est extrapolée aux unités traitées ayant des X identiques. Cette méthode, qui apparait dans la deuxième parenthèse de l'ATT, exige une spécification correcte du modèle prédisant l'évolution de la variable d'intérêt, Y, compte tenu des caractéristiques observables X. Ainsi, si on suppose une relation linéaire entre Y et le salaire d'un individu alors que la relation réelle est quadratique, l'estimation de  $\widehat{m}$  ne représentera pas une tendance commune correcte et l'ATT sera biaisée.

La deuxième approche CDiD intègre la pondération inverse de la probabilité (IPW - inverse probability weighting) à l'estimateur de différence dans les différences. Comme l'a montré Abadie (2005), l'estimateur CDiD semi-paramétrique peut être mis en œuvre

en repondérant les unités de contrôle par un « score de propension »  $\frac{\widehat{p}(X)(1-D)}{1-\widehat{p}(X)}$ , qui est un ratio basé sur la probabilité d'être traité en fonction de X.<sup>1</sup> Grâce à cette repondération, les unités de contrôle ressemblant le plus aux unités traitées recevront un poids plus élevé et, en moyenne, le groupe de contrôle repondéré partagera les mêmes caractéristiques que les unités traitées, si le modèle de score de propension a été correctement spécifié. L'hypothèse de tendance parallèle est alors imposée au groupe de contrôle repondéré. Cette approche apparaît dans la première parenthèse dans l'ATT.

L'ATT obtenu avec le CDiD de Sant'Anna et Zhao (2020) adapté à un cadre d'évaluation échelonnée par Callaway et Sant'Anna (2021) est écrit :

$$\widehat{ATT}(d, t) = E \left[ \left( \frac{D_d}{E[D_d]} - \frac{\frac{\widehat{p}_d(X)C}{1-\widehat{p}_d(X)}}{E\left[\frac{\widehat{p}_d(X)C}{1-\widehat{p}_d(X)}\right]}} \right) \left( Y_t - Y_{d-\delta-1} - \widehat{m}_{d,t,\delta}(X) \right), \right]$$

<sup>1</sup> Il est d'usage de normaliser les poids par leurs espérances pour améliorer les propriétés des estimateurs dans les petits échantillons.

**Tableau A1: Catégorisation des pays**

<b>Nationalité</b>	<b>Nationalité - détaillée</b>	<b>Nationalité - regroupée</b>
Afghanistan	Afghanistan	Asie: Sud
Afrique du Sud	Afrique-Anglophone	Afrique: Anglais
Albanie	Albanie	Pays Balkans
Algérie	Algérie	Afrique: Français
Allemagne	Allemagne	UE: Allemand
Andorre	Europe Autres	UE: Autres
Angola	Afrique-Lusophone	Afrique: Portugais
Antigua-et-Barbuda	Amériques	Amerique du Sud: Espagnol
Arabie saoudite	Moyen-Orient Afrique Nord	Afrique: Autres
Argentina	Amériques	Amerique du Sud: Espagnol
Arménie	Moyen-Orient Afrique Nord	Afrique: Autres
Aruba	Amériques	Amerique du Sud: Espagnol
Australie	Océanie	USA-Océanie
Autres	Autres	Autres
Autriche	UE15 Autres	UE: Autres
Azerbaïdjan	Ex-URSS	Ex-URSS
Bahamas	Amériques	Amerique du Sud: Espagnol
Bahreïn	Moyen-Orient Afrique Nord	Afrique: Autres
Bangladesh	Asie du Sud/ Sud-Est	Asie: Reste
Barbade	Amériques	Amerique du Sud: Espagnol
Bélarus	Ex-URSS	Ex-URSS
Belgique	Belgique	UE: Français
Belize	Amériques	Amerique du Sud: Espagnol
Bénin	Afrique-Francophone	Afrique: Français
Bermudes	Amériques	Amerique du Sud: Espagnol
Bhoutan	Asie du Sud/ Sud-Est	Asie: Reste
Bolivie (Etat plurinational de)	Amériques	Amerique du Sud: Espagnol
Bosnie-Herzégovine	Bosnie-Herzégovine	Pays Balkans
Botswana	Afrique-Anglophone	Afrique: Anglais
Brésil	Brésil	Amerique Sud: Portugais
Brunéi Darussalam	Asie du Sud/ Sud-Est	Asie: Reste
Bulgarie	Bulgarie	UE: Pays de l'Est
Burkina Faso	Afrique-Francophone	Afrique: Français
Burundi	Afrique-Francophone	Afrique: Français
Cabo Verde	Cabo Verde	Afrique: Portugais
Cambodge	Asie du Sud/ Sud-Est	Asie: Reste
Cameroun	Cameroun	Afrique: Français

Canada	Amériques	Amerique du Sud: Espagnol
Chili	Amériques	Amerique du Sud: Espagnol
Chine	Chine	Asie: Reste
Chypre	UE15 Autres	UE: Autres
Colombie	Amériques	Amerique du Sud: Espagnol
Comores	Afrique-Francophone	Afrique: Français
Congo	Afrique-Francophone	Afrique: Français
Costa Rica	Amériques	Amerique du Sud: Espagnol
Côte d'Ivoire	Afrique-Francophone	Afrique: Français
Croatie	Croatie	UE: Autres
Cuba	Amériques	Amerique du Sud: Espagnol
Curaçao	Amériques	Amerique du Sud: Espagnol
Danemark	UE15 Autres	UE: Autres
Djibouti	Moyen-Orient Afrique Nord	Afrique: Autres
Dominique	Amériques	Amerique du Sud: Espagnol
Egypte	Egypte	Afrique: Autres
El Salvador	Amériques	Amerique du Sud: Espagnol
Emirats arabes unis	Moyen-Orient Afrique Nord	Afrique: Autres
Equateur	Amériques	Amerique du Sud: Espagnol
Erythrée	Erythrée	Afrique: Autres
Espagne	Espagne	UE: Espagnol
Estonie	UE10 Pays de l'Est	UE: Pays de l'Est
État de Palestine	Moyen-Orient Afrique Nord	Afrique: Autres
Etats-Unis d'Amérique	Etats-Unis d'Amérique	USA-Océanie
Ethiopie	Afrique-Anglophone	Afrique: Anglais
Fidji	Océanie	USA-Océanie
Finlande	UE15 Autres	UE: Autres
France	France	UE: Français
Gabon	Afrique-Francophone	Afrique: Français
Gambie	Afrique-Anglophone	Afrique: Anglais
Géorgie	Ex-URSS	Ex-URSS
Ghana	Afrique-Anglophone	Afrique: Anglais
Grèce	Grèce	UE: Autres
Grenada	Amériques	Amerique du Sud: Espagnol
Groenland	Amériques	Amerique du Sud: Espagnol
Guatemala	Amériques	Amerique du Sud: Espagnol
Guernesey	Europe Autres	UE: Autres
Guinée	Afrique-Francophone	Afrique: Français
Guinée	Afrique-Francophone	Afrique: Français

Guinée équatoriale	Afrique-Francophone	Afrique: Français
Guinée-Bissau	Afrique-Lusophone	Afrique: Portugais
Guyana	Amériques	Amerique du Sud: Espagnol
Haiti	Amériques	Amerique du Sud: Espagnol
Honduras	Amériques	Amerique du Sud: Espagnol
Hong Kong	Asie de l'Est	Asie: Reste
Hongrie	Hongrie	UE: Pays de l'Est
Île de Man	Europe Autres	UE: Autres
Îles Caïmanes	Amériques	Amerique du Sud: Espagnol
Îles Mariannes septentrionales	Asie du Sud/ Sud-Est	Asie: Reste
Îles Marshall	Océanie	USA-Océanie
Îles Salomon	Océanie	USA-Océanie
Îles Vierges britanniques	Amériques	Amerique du Sud: Espagnol
Inde	Inde	Asie: Sud
Indonésie	Asie du Sud/ Sud-Est	Asie: Reste
Iran (République islamique d')	Iran (République islamique d')	Asie: Sud
Iraq	Iraq	Afrique: Autres
Irlande	UE15 Autres	UE: Autres
Islande	Europe Autres	UE: Autres
Israël	Moyen-Orient Afrique Nord	Afrique: Autres
Italie	Italie	UE: Italien
Jamaïque	Amériques	Amerique du Sud: Espagnol
Japon	Asie de l'Est	Asie: Reste
Jersey	Europe Autres	UE: Autres
Jordanie	Moyen-Orient Afrique Nord	Afrique: Autres
Kazakhstan	Ex-URSS	Ex-URSS
Kenya	Afrique-Anglophone	Afrique: Anglais
Kirghizistan	Ex-URSS	Ex-URSS
Kiribati	Océanie	USA-Océanie
Kosovo	Kosovo	Pays Balkans
Koweït	Moyen-Orient Afrique Nord	Afrique: Autres
Lesotho	Afrique-Anglophone	Afrique: Anglais
Lettonie	UE10 Pays de l'Est	UE: Pays de l'Est
Liban	Liban	Afrique: Autres
Libéria	Afrique-Anglophone	Afrique: Anglais
Libye	Moyen-Orient Afrique Nord	Afrique: Autres
Liechtenstein	Europe Autres	UE: Autres
Lituanie	UE10 Pays de l'Est	UE: Pays de l'Est

Macao	Asie de l'Est	Asie: Reste
Macédoine	Macédoine	Pays Balkans
Madagascar	Afrique-Francophone	Afrique: Français
Malaisie	Asie du Sud/ Sud-Est	Asie: Reste
Malawi	Afrique-Anglophone	Afrique: Anglais
Maldives	Asie du Sud/ Sud-Est	Asie: Reste
Mali	Afrique-Francophone	Afrique: Français
Malte	UE15 Autres	UE: Autres
Maroc	Maroc	Afrique: Français
Maurice	Maurice	Afrique: Français
Mauritanie	Afrique-Francophone	Afrique: Français
Mexique	Amériques	Amerique du Sud: Espagnol
Micronésie (Etats fédérés de)	Océanie	USA-Océanie
Moldavie	Ex-URSS	Ex-URSS
Monaco	Europe Autres	UE: Autres
Mongolie	Asie de l'Est	Asie: Reste
Monténégro (le)	Monténégro (le)	Pays Balkans
Montserrat	Amériques	Amerique du Sud: Espagnol
Mozambique	Afrique-Lusophone	Afrique: Portugais
Myanmar	Asie du Sud/ Sud-Est	Asie: Reste
Namibie	Afrique-Anglophone	Afrique: Anglais
Nauru	Océanie	USA-Océanie
Népal	Asie du Sud/ Sud-Est	Asie: Reste
Nicaragua	Amériques	Amerique du Sud: Espagnol
Niger	Afrique-Francophone	Afrique: Français
Nigéria	Nigéria	Afrique: Anglais
Nioué	Océanie	USA-Océanie
Norvège	Europe Autres	UE: Autres
Nouvelle-Calédonie	Océanie	USA-Océanie
Nouvelle-Zélande	Océanie	USA-Océanie
Oman	Moyen-Orient Afrique Nord	Afrique: Autres
Ouganda	Afrique-Anglophone	Afrique: Anglais
Ouzbékistan	Ex-URSS	Ex-URSS
Pakistan	Pakistan	Asie: Sud
Palaos	Océanie	USA-Océanie
Panama	Amériques	Amerique du Sud: Espagnol
Papouasie-Nouvelle-Guinée	Océanie	USA-Océanie
Paraguay	Amériques	Amerique du Sud: Espagnol
Pays-Bas	Pays-Bas	UE: Autres
Pérou	Amériques	Amerique du Sud: Espagnol
Philippines	Philippines	Asie: Reste



Pologne	Pologne	UE: Pays de l'Est
Polynésie française	Océanie	USA-Océanie
Portugal	Portugal	UE: Portugais
Qatar	Moyen-Orient Afrique Nord	Afrique: Autres
République centrafricaine	Afrique-Francophone	Afrique: Français
République de Corée	Asie de l'Est	Asie: Reste
République dém. pop. Lao	Asie du Sud/ Sud-Est	Asie: Reste
République démocratique du Congo	Afrique-Francophone	Afrique: Français
République dominicaine	Amériques	Amerique du Sud: Espagnol
République pop. dém. de Corée	Asie de l'Est	Asie: Reste
Roumanie	Roumanie	UE: Pays de l'Est
Royaume-Uni	Royaume-Uni	UE: Anglais
Russie	Russie	Ex-URSS
Rwanda	Afrique-Francophone	Afrique: Français
Sahara occidental	Moyen-Orient Afrique Nord	Afrique: Autres
Sainte-Lucie	Amériques	Amerique du Sud: Espagnol
Saint-Kitts-et-Nevis	Amériques	Amerique du Sud: Espagnol
Saint-Marin	Europe Autres	UE: Autres
Saint-Vincent-et-les Grenadines	Amériques	Amerique du Sud: Espagnol
Samoa	Océanie	USA-Océanie
Sao Tomé-et-Principe	Afrique-Lusophone	Afrique: Portugais
Sénégal	Afrique-Francophone	Afrique: Français
Serbie (la)	Serbie (la)	Pays Balkans
Serbie-et-Monténégro	Serbie-et-Monténégro	Pays Balkans
Seychelles	Afrique-Anglophone	Afrique: Anglais
Sierra Leone	Afrique-Anglophone	Afrique: Anglais
Singapour	Asie du Sud/ Sud-Est	Asie: Reste
Slovaquie	UE10 Pays de l'Est	UE: Pays de l'Est
Slovénie	UE10 Pays de l'Est	UE: Pays de l'Est
Somalie	Afrique-Anglophone	Afrique: Anglais
Soudan	Afrique-Anglophone	Afrique: Anglais
Soudan du Sud	Afrique-Anglophone	Afrique: Anglais
Sri Lanka	Asie du Sud/ Sud-Est	Asie: Reste
Suède	UE15 Autres	UE: Autres
Suisse	Europe Autres	UE: Autres
Suriname	Amériques	Amerique du Sud: Espagnol
Swaziland	Afrique-Anglophone	Afrique: Anglais
Syrie	Syrie	Afrique: Autres
Tadjikistan	Ex-URSS	Ex-URSS

Tanzanie	Afrique-Anglophone	Afrique: Anglais
Tchad	Afrique-Francophone	Afrique: Français
Tchécoslovaquie	UE10 Pays de l'Est	UE: Pays de l'Est
Tchéquie	UE10 Pays de l'Est	UE: Pays de l'Est
Thaïlande	Asie du Sud/ Sud-Est	Asie: Reste
Timor-Leste	Asie du Sud/ Sud-Est	Asie: Reste
Togo	Afrique-Francophone	Afrique: Français
Tonga	Océanie	USA-Océanie
Trinité-et-Tobago	Amériques	Amerique du Sud: Espagnol
Tunisie	Tunisie	Afrique: Français
Turkménistan	Ex-URSS	Ex-URSS
Turquie	Turquie	Afrique: Autres
Tuvalu	Océanie	USA-Océanie
Ukraine	Ukraine	Ex-URSS
URSS	Ex-URSS	Ex-URSS
Uruguay	Amériques	Amerique du Sud: Espagnol
Vanuatu	Océanie	USA-Océanie
Vatican	Europe Autres	UE: Autres
Venezuela	Amériques	Amerique du Sud: Espagnol
Viet Nam	Asie du Sud/ Sud-Est	Asie: Reste
Yémen	Moyen-Orient Afrique Nord	Afrique: Autres
Yougoslavie	Yougoslavie	Pays Balkans
Zambie	Afrique-Anglophone	Afrique: Anglais
Zimbabwe	Afrique-Anglophone	Afrique: Anglais

Tableau A 2: Déterminants de la probabilité d'obtenir le certificat

	Probabilité de signer le CAI	Probabilité d'obtenir le certificat
Femme	0.006*** [0.00,0.01]	0.041* [0.01,0.08]
<i>Age; référence: Age 35-39</i>		
Age 20-24	-0.015*** [-0.02,-0.01]	-0.168*** [-0.25,-0.09]
Age 25-29	-0.002 [-0.00,0.00]	-0.117*** [-0.17,-0.07]
Age 30-34	0.007*** [0.00,0.01]	-0.037 [-0.08,0.01]
Age 40-44	-0.005*** [-0.01,-0.00]	0.047 [-0.01,0.10]
Age 45-49	-0.010*** [-0.01,-0.01]	0.010 [-0.06,0.08]
Age 50-54	-0.012*** [-0.01,-0.01]	0.041 [-0.05,0.13]
Age 55-59	-0.011*** [-0.01,-0.01]	-0.029 [-0.14,0.09]
Age 60+	-0.011*** [-0.01,-0.01]	-0.095 [-0.23,0.05]
<i>Canton de résidence; référence: Lux.-Ville</i>		
Capellen	-0.005*** [-0.01,-0.00]	0.004 [-0.07,0.08]
Clervaux	-0.014*** [-0.02,-0.01]	0.181 [-0.10,0.47]
Diekirch	-0.009*** [-0.01,-0.01]	-0.012 [-0.12,0.09]
Echternach	-0.006*** [-0.01,-0.00]	0.056 [-0.07,0.18]
Esch-sur-Alzette	-0.007*** [-0.01,-0.01]	0.006 [-0.04,0.05]
Grevenmacher	-0.004* [-0.01,-0.00]	0.056 [-0.04,0.15]
Luxembourg-Campagne	0.004** [0.00,0.01]	-0.037 [-0.08,0.01]
Mersch	-0.006*** [-0.01,-0.00]	-0.022 [-0.12,0.07]
Redange	-0.012*** [-0.01,-0.01]	0.211* [0.02,0.40]
Remich	-0.006*** [-0.01,-0.00]	0.030 [-0.09,0.15]
Vianden	-0.012*** [-0.02,-0.01]	-0.122 [-0.46,0.22]
Wiltz	-0.012*** [-0.01,-0.01]	-0.227*** [-0.35,-0.10]
Résidence depuis 2010 (années)	-0.002*** [-0.00,-0.00]	-0.005 [-0.01,0.00]
<i>État civil; référence: Marié</i>		
Célibataire	-0.009*** [-0.01,-0.01]	-0.002 [-0.05,0.05]
Séparé/divorcé/veuf	-0.002* [-0.00,-0.00]	-0.049 [-0.13,0.03]
Etat civil inconnu	-0.019***	-0.078

Suite sur la page suivante

Tableau A 2 – suite

	Probabilité de signer	Probabilité de conclure
Taille du ménage	[-0.02,-0.02] -0.001***	[-0.21,0.06] 0.009
Né à Luxembourg	[-0.00,-0.00] -0.007***	[-0.01,0.03] 0.340
	[-0.01,-0.01]	[-0.02,0.70]
<i>Statut professionnel; référence: Salarié</i>		
Revenu minimum garanti	-0.005	-0.161*
	[-0.01,0.00]	[-0.30,-0.02]
Chômage	-0.001	-0.129*
	[-0.00,0.00]	[-0.24,-0.02]
Co-assuré	0.013***	-0.111*
	[0.01,0.02]	[-0.20,-0.02]
Autres (inactifs, fonc. EU, ...)	-0.016***	-0.058
	[-0.02,-0.01]	[-0.14,0.02]
Fonctionnaire	0.001	-0.218
	[-0.01,0.01]	[-0.49,0.06]
Indépendant	0.007***	-0.061
	[0.00,0.01]	[-0.17,0.05]
<i>Caractéristiques de l'emploi et historique d'activité</i>		
Salaire réel horaire (dernier)	-0.000	0.000
	[-0.00,0.00]	[-0.00,0.00]
Ouvrier	-0.006***	-0.104***
	[-0.01,-0.00]	[-0.16,-0.05]
Sans CDI	0.001	-0.037
	[-0.00,0.00]	[-0.10,0.03]
Mesure d'activation	-0.005	0.064
	[-0.01,0.00]	[-0.12,0.25]
Sans salaire depuis 2010	-0.012***	0.023
	[-0.01,-0.01]	[-0.04,0.08]
Autre statut depuis 2010 (%)	0.014***	0.017
	[0.01,0.02]	[-0.07,0.10]
Allocation chômage depuis 2010 (%)	-0.000	-0.051
	[-0.00,0.00]	[-0.27,0.16]
<i>Secteur d'activité</i>		
	<i>réf.: Sect: Act. financ. &amp; assur.</i>	<i>réf.: Sect: Services</i>
Sect: Agriculture	-0.012***	
	[-0.02,-0.01]	
Sect: Indust. extractive	0.005	
	[-0.01,0.02]	
Sect: Indust. manufact.	0.000	
	[-0.00,0.00]	
Sect: Prod./distr. électricité, gaz	0.014	
	[-0.01,0.03]	
Sect: Prod./distr. eau et assain.	-0.001	
	[-0.01,0.01]	
Sect: Construction	-0.006***	
	[-0.01,-0.00]	
Sect: Commerce et rép. automobiles	-0.005***	
	[-0.01,-0.00]	
Sect: Transports et entreposage	-0.004*	
	[-0.01,-0.00]	
Sect: Hébergement et restauration	-0.007***	
	[-0.01,-0.00]	

Suite sur la page suivante

Tableau A 2 – suite

	Probabilité de signer	Probabilité de conclure
Sect: Information et communication	0.014*** [0.01,0.02]	
Sect: Activités immobilières	-0.006* [-0.01,-0.00]	
Sect: Act. spéc., scient. et techn.	-0.005** [-0.01,-0.00]	
Sect: Act. serv. administratifs	-0.003* [-0.01,-0.00]	
Sect: Admin. publique	-0.003 [-0.01,0.00]	
Sect: Enseignement	0.001 [-0.00,0.01]	
Sect: Santé humaine, act. sociale	-0.006*** [-0.01,-0.00]	
Sect: Arts et act. récréatives	-0.014*** [-0.02,-0.01]	
Sect: Autres services	-0.005* [-0.01,-0.00]	
Sect: Act. ménage en tant qu'employeur	-0.007*** [-0.01,-0.00]	
Sect: Act. extra-territoriales	-0.007 [-0.02,0.00]	
Sans secteur d'emploi	-0.005** [-0.01,-0.00]	0.065 [-0.01,0.14]
Sect: Autres		-0.000 [-0.08,0.08]
<i>Nationalité détaillée; référence: France</i>		
Afghanistan	0.023** [0.01,0.04]	-0.148 [-0.41,0.11]
Albanie	0.066*** [0.05,0.08]	0.212* [0.04,0.38]
Algérie	0.069*** [0.05,0.09]	0.095 [-0.12,0.31]
Allemagne	0.002* [0.00,0.00]	0.034 [-0.12,0.19]
Belgique	0.001 [-0.00,0.00]	-0.111 [-0.27,0.05]
Bosnie-Herzégovine	0.029*** [0.02,0.04]	0.186* [0.02,0.35]
Brésil	0.079*** [0.07,0.09]	-0.028 [-0.14,0.08]
Bulgarie	0.034*** [0.02,0.05]	0.039 [-0.13,0.21]
Cap-Vert	0.019*** [0.01,0.02]	0.014 [-0.17,0.20]
Cameroun	0.104*** [0.08,0.13]	0.138 [-0.02,0.29]
Chine	0.048*** [0.04,0.06]	0.080 [-0.03,0.19]
Croatie	0.032*** [0.02,0.04]	0.013 [-0.18,0.21]
Egypte	0.187*** [0.14,0.23]	0.154 [-0.04,0.35]
Erythrée	0.004 [-0.00,0.01]	-0.142 [-0.53,0.24]

Suite sur la page suivante

Tableau A 2 – suite

	Probabilité de signer	Probabilité de conclure
Espagne	0.022*** [0.02,0.03]	-0.194*** [-0.30,-0.09]
Etats-Unis	0.039*** [0.03,0.05]	0.031 [-0.11,0.17]
Grèce	0.018*** [0.01,0.02]	-0.072 [-0.21,0.07]
Hongrie	0.022*** [0.01,0.03]	-0.006 [-0.18,0.17]
Inde	0.120*** [0.11,0.13]	0.089 [-0.01,0.19]
Iran	0.072*** [0.05,0.09]	0.101 [-0.10,0.31]
Iraq	0.053*** [0.04,0.07]	0.140 [-0.11,0.39]
Italie	0.015*** [0.01,0.02]	-0.116** [-0.20,-0.03]
Kosovo	0.018*** [0.01,0.03]	0.033 [-0.21,0.28]
Liban	0.077*** [0.05,0.11]	0.405** [0.14,0.67]
Macédoine	0.051*** [0.03,0.07]	0.193 [-0.09,0.48]
Maroc	0.089*** [0.07,0.10]	0.172* [0.04,0.31]
Maurice	0.068*** [0.05,0.09]	0.160 [-0.03,0.35]
Monténégro	0.018*** [0.01,0.02]	0.125 [-0.05,0.30]
Nigéria	0.117*** [0.08,0.15]	-0.164 [-0.37,0.05]
Pakistan	0.150*** [0.11,0.19]	0.178 [-0.08,0.43]
Pays-Bas	0.005*** [0.00,0.01]	-0.091 [-0.36,0.18]
Philippines	0.109*** [0.08,0.13]	0.019 [-0.14,0.18]
Pologne	0.029*** [0.02,0.03]	-0.055 [-0.17,0.06]
Portugal	0.011*** [0.01,0.01]	-0.136** [-0.22,-0.05]
Roumanie	0.028*** [0.02,0.03]	0.066 [-0.04,0.17]
Royaume-Uni	0.018*** [0.01,0.02]	0.042 [-0.08,0.17]
Russie	0.157*** [0.14,0.17]	0.263*** [0.16,0.37]
Serbie	0.036*** [0.03,0.04]	0.072 [-0.08,0.23]
Syrie	0.041*** [0.03,0.05]	-0.008 [-0.21,0.20]
Tunisie	0.058*** [0.04,0.07]	0.299*** [0.14,0.45]
Turquie	0.182*** [0.16,0.21]	0.229*** [0.10,0.36]
Ukraine	0.076*** [0.06,0.09]	0.187** [0.06,0.31]

Suite sur la page suivante

Tableau A 2 – suite

	Probabilité de signer	Probabilité de conclure
Yugoslavie (inconnu)	0.003 [-0.00,0.01]	
Afrique-Anglophone	0.069*** [0.05,0.09]	0.135 [-0.04,0.32]
Afrique-Francophone	0.068*** [0.06,0.08]	0.137* [0.00,0.27]
Afrique-Lusophone	0.019** [0.01,0.03]	0.199 [-0.19,0.59]
Amérique	0.088*** [0.08,0.10]	0.125* [0.01,0.24]
Asie de l'Est	0.018** [0.00,0.03]	-0.036 [-0.30,0.23]
Asie du Sud/ Sud-Est	0.045*** [0.03,0.06]	0.135 [-0.02,0.29]
UE10 Pays de l'Est	0.020*** [0.01,0.03]	-0.046 [-0.18,0.09]
Europe Autres	0.003 [-0.00,0.01]	0.507*** [0.40,0.62]
UE15 Autres	-0.000 [-0.00,0.00]	-0.270** [-0.43,-0.11]
Moyen-Orient Afrique Nord	0.115*** [0.08,0.15]	0.208* [0.03,0.39]
Océanie	0.026* [0.00,0.05]	0.146 [-0.23,0.52]
Anciennes républiques soviétiques	0.120*** [0.10,0.14]	0.090 [-0.06,0.24]
Autres	0.052** [0.02,0.08]	-0.019 [-0.42,0.38]
<i>Education; référence: Education tertiaire</i>		
Education primaire		-0.132* [-0.25,-0.01]
Education secondaire		-0.087*** [-0.13,-0.04]
Education inconnue		-0.023 [-0.08,0.03]
<i>Titre de séjour; référence: Visa: Autorisation</i>		
Visa: Carte Bleue		-0.016 [-0.10,0.07]
Visa: Regroupement familial (Lux.)		-0.076 [-0.18,0.03]
Visa: Regroupement familial (UE)		-0.102** [-0.18,-0.03]
Visa: Regroupement familial (non-UE)		-0.104** [-0.18,-0.03]
Visa: Privé		-0.100 [-0.20,0.00]
Visa: Protection		0.004 [-0.17,0.18]
Visa: Emploi		-0.012 [-0.08,0.05]
Visa: inconnu		0.028 [-0.09,0.15]
<i>Année; référence: Année 2017</i>		

Suite sur la page suivante

Tableau A 2 – suite

	Probabilité de signer	Probabilité de conclure
Année 2018	0.014*** [0.01,0.02]	0.009 [-0.03,0.05]
Année 2019	0.019*** [0.02,0.02]	0.010 [-0.03,0.05]
Année 2020	-0.009*** [-0.01,-0.01]	-0.079* [-0.14,-0.02]
Année 2021	-0.001 [-0.00,0.00]	
Année 2022	0.002 [-0.00,0.00]	
Constante	0.033*** [0.03,0.04]	0.521*** [0.42,0.62]
Observations	268339	4123
Notes: 95% intervalles de confiance entre parenthèses; * $p < 0.05$ , ** $p < 0.01$ , *** $p < 0.001$		



