

**Analyse comparative entre
les sexes plus (ACS Plus)
de certains programmes
du marché du travail**

Présentation pour l'atelier sur
l'évaluation quantitative de l'impact
Mars 2024

Plan de présentation

- Contexte
- Sources de données et indicateurs
- Méthodologie
- Résultats
- Conclusion
- Annexes
- Références



Contexte de l'étude

EDSC utilise des méthodes d'appariement pour évaluer l'efficacité de ses programmes du marché du travail:

- Les techniques d'appariement sont robustes, mais ne fournissent que les impacts moyens.
- Impossible d'estimer la distribution des impacts du programme parmi les participants.
- Difficile de mener des analyses par sous-groupes sur différents facteurs d'identité croisés.

Des développements récents en apprentissage automatique ont été appliqués à l'évaluation de programmes du marché du travail en Europe (Belgique et Suisse):

- L'apprentissage automatique a été utilisé pour estimer les impacts incrémentiels granulaires au niveau individuel, permettant de mettre en évidence « ce qui fonctionne pour qui » (Athey & Wager, 2018 ; Lechner, 2019).
- L'apprentissage automatique causal pour l'évaluation de la formation en Belgique (Lechner, 2020)



Objectif de l'étude

- Tester l'efficacité d'une nouvelle méthode d'apprentissage automatique pour estimer les impacts incrémentiels de programmes en fonction de différents facteurs d'identité croisés de l'ACS Plus.
- Examiner deux programmes actifs du marché du travail :
 - Ententes sur le développement du marché du travail (EDMT); et
 - Fonds d'intégration pour les personnes handicapées (FIPH).

Qu'est-ce que l'Analyse comparative entre les sexes plus (ACS Plus)?

- Un processus analytique utilisé pour évaluer l'expérience de différentes femmes, hommes et personnes de diverses identités de genre en ce qui concerne les politiques, les programmes et les initiatives.
- Le 'Plus' de l'ACS Plus reconnaît que l'ACS va au-delà des différences biologiques (sexe) et socioculturelles (genre).

Source : Femmes et Égalité des genres Canada¹

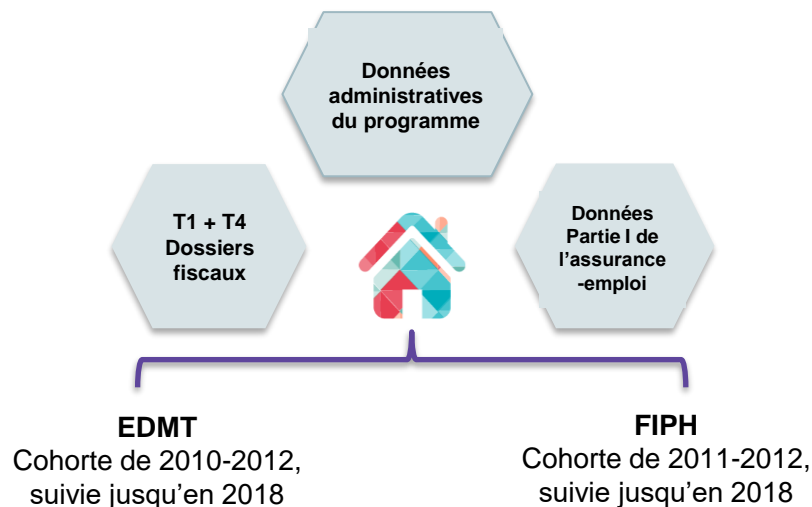


¹ <https://femmes-egalite-genres.canada.ca/fr/analyse-comparative-entre-sexes-plus/est-analyse-comparative-entre-sexes-plus.html>

Data sources: Labour Market Program Data Platform

Cette étude utilise des ensembles intégrés de riches données administratives.

- L'étape initiale consiste à nettoyer les enregistrements en double, créer des équivalents de plan d'action (ÉPA) à partir des interventions du programme et construire la base de données finale à partir des données du programme, des données de l'ARC et des données de la Partie I de l'assurance-emploi.



Deux groupes d'intérêt pour produire les impacts incrémentaux :

- **Groupes de participants:** Individus ayant participé aux EDMT et FIPH
- **Groupes de contrôle:** Individus similaires n'ayant pas participé aux EDMT et FIPH:
 - Pour EDMT: Prestataires d'assurance-emploi actifs qui n'ont pas participé à l'EDMT.
 - Pour FIPH: Les personnes handicapées qui ne sont pas des prestataires d'assurance-emploi et qui ont pris le service d'aide à l'emploi.



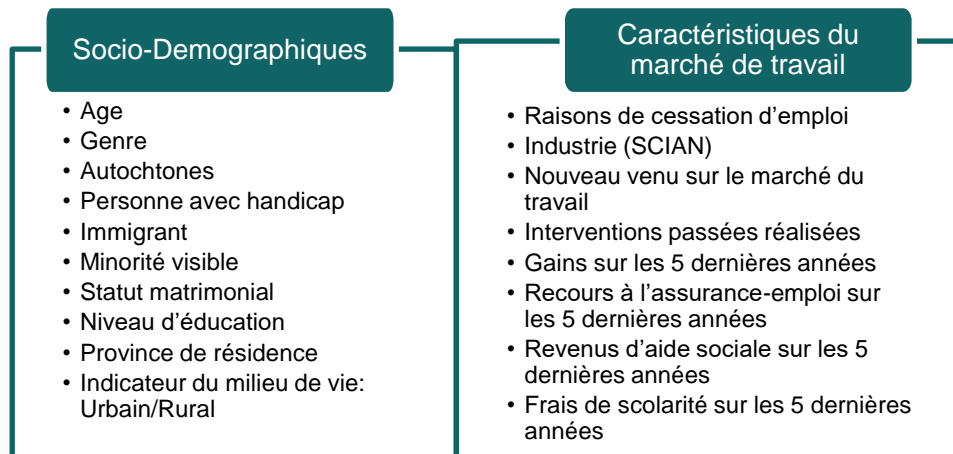
VARIABLES DE CONTRÔLE ET INDICATEURS

Les principaux indicateurs sont la moyenne sur 5 ans post-programme de :

- Incidence d'emploi (pp.)
- Revenus d'emploi (\$)
- Dépendance au soutien du revenu (pp.)

* Les observations présentant de l'information manquante pour les indicateurs de résultat ont été exclues pour garantir le bon fonctionnement de l'algorithme choisi.

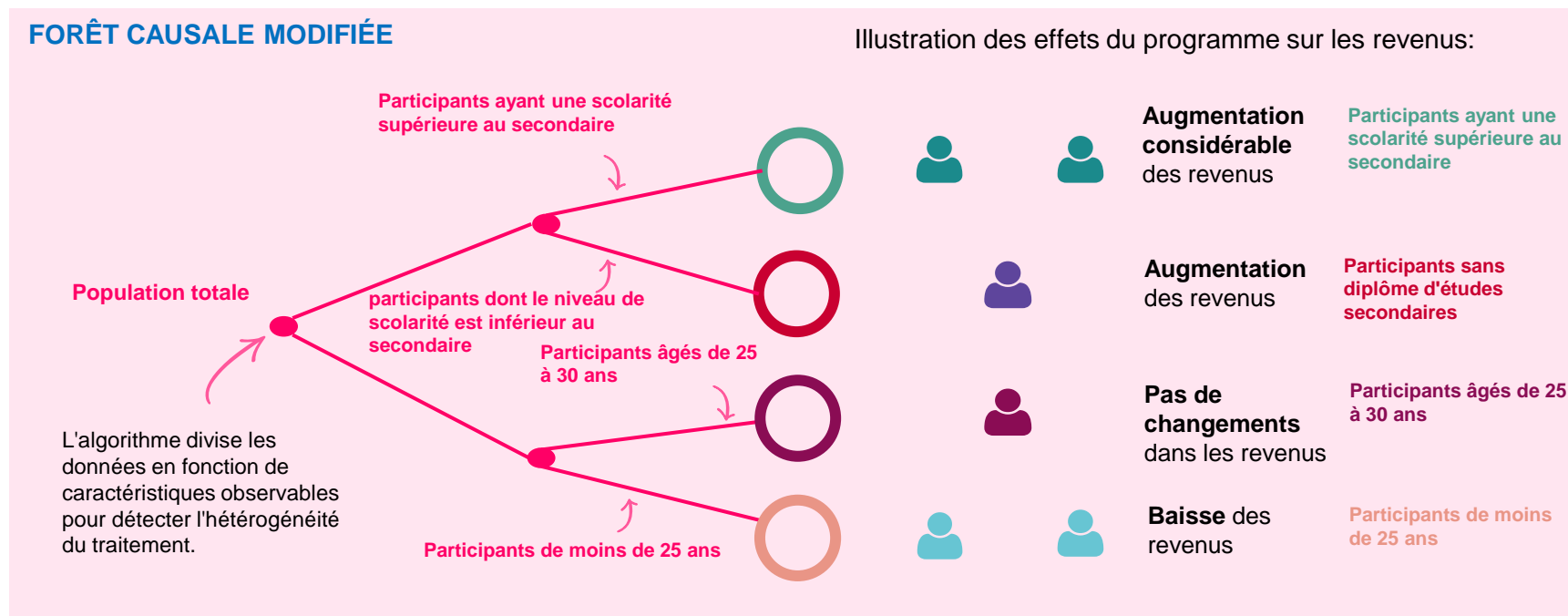
Plus de 75+ variables utilisées pour construire la matrice des covariables:



Méthodologie: Les forêts causales modifiées (FCM)

La présente étude utilise les forêts causales modifiées (FCM):

- Il s'agit d'un algorithme d'apprentissage automatique causal supervisé qui construit un ensemble d'arbres décorrélés, apprend les caractéristiques à partir des données et estime les impacts du programme avec R (Lechner, 2019).



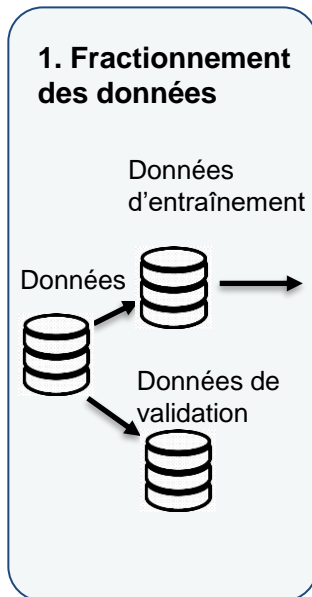
L'inférence causale

- L'impact différentiel fait référence à l'impact ou au changement mesurable d'un résultat qui peut être attribué à une intervention ou à un traitement spécifique, souvent évalué en comparant les résultats entre un groupe de traitement qui a reçu le traitement et un groupe témoin qui ne l'a pas reçu.
- Wager & Athey (2018) ont proposé les forêts causales avec des inférences statistiques valides basées sur la méthode traditionnelle d'apprentissage automatique, la forêt aléatoire, proposée par Briedman (2001).
- Lechner (2019) a modifié les termes d'erreur des forêts causales pour fournir des estimateurs plus impartiaux et plus granulaires.



Étape 1: Fractionnement des données

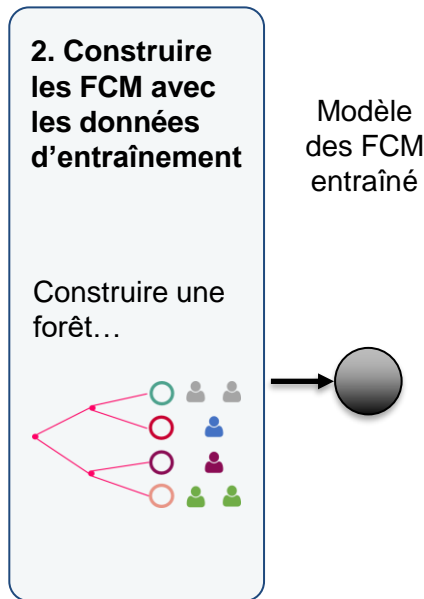
Les données administratives du programme liées sont réparties de manière aléatoire en:



- **50% Données d'entraînement:** pour entraîner l'algorithme
- **50% Données de validation:** pour estimer les effets en appliquant les FCM entraînées

Cela vise à éviter que les FCM ne performant de manière excessive sur les données qu'elles ont déjà "vues". En lui présentant un nouvel ensemble de données, les données de test, nous nous assurons de "généraliser" les FCM.

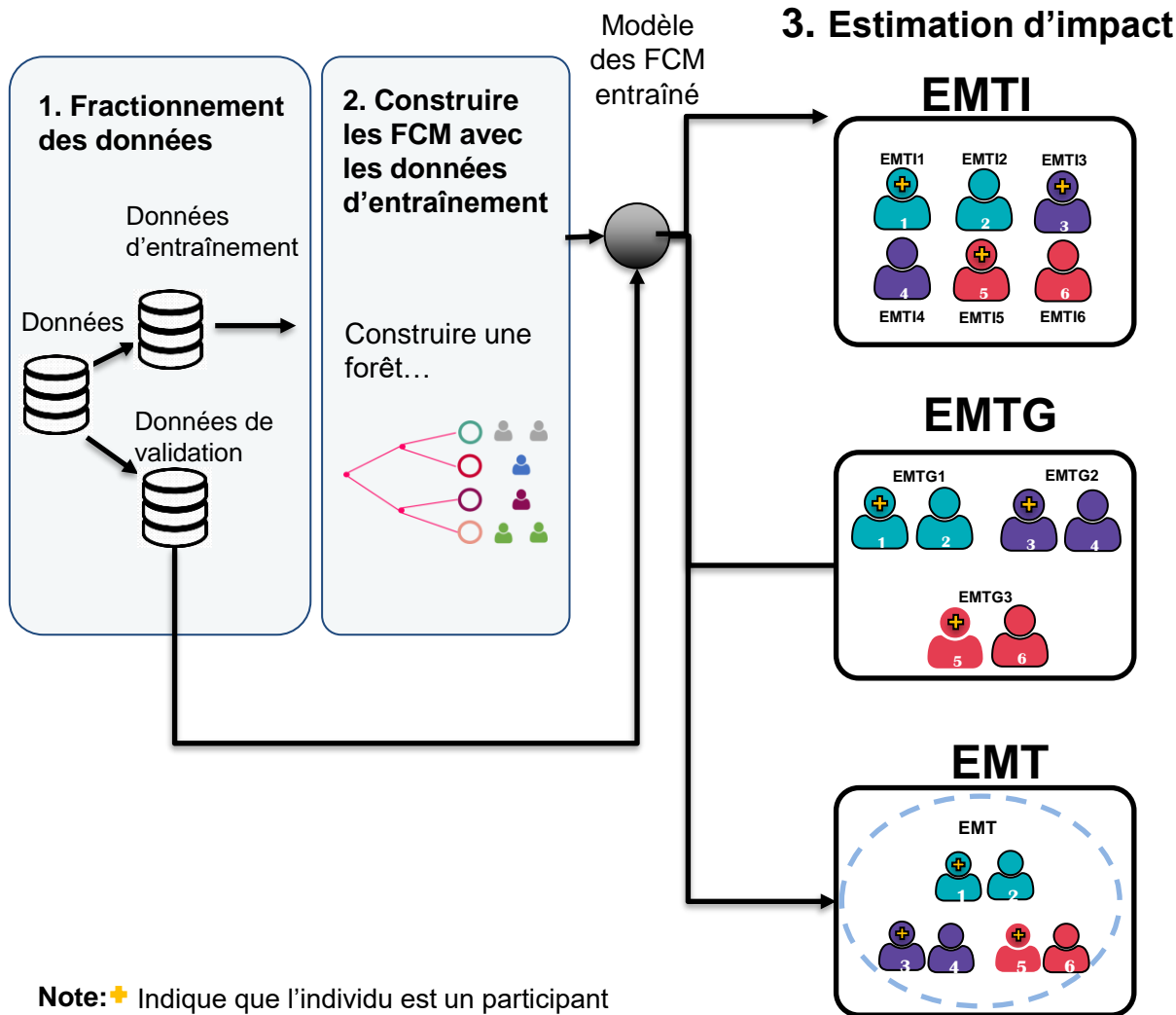
Étape 2: Construire les FCM avec les données d'entraînement



Lors de l'entraînement des FCM, l'algorithme divise lui-même les données en deux parties pour construire les forêts:

- **Données d'entraînement:** pour apprendre comment effectuer des divisions
- **Données honnêtes:** pour estimer les effets du traitement parmi les individus appartenant à la même division

Étape 3: Estimation d'impact



Effet moyen de traitement ***Nouveau*** individuel (EMTI)

Mesure l'impact d'un programme sur un individu ayant un ensemble de caractéristiques ou un profil donné. Représente l'impact causal du programme au niveau de granularité le plus fin.

Effet moyen de traitement ***Nouveau*** de groupe (EMTG)

L'EMTG peut être estimé en agrégeant et en pondérant les EMTI sur des sous-groupes spécifiques. Contrairement aux analyses de sous-groupes traditionnelles, les EMTG peuvent être comparées entre les groupes.

Effet moyen de traitement (EMT)

Représente l'impact moyen du programme sur la population.



Étape 4: Effectuer des tests de signification pour la différence entre les sexes pour les sous-groupes

- L'algorithme des FCM détermine si deux Effets Moyens du Traitement Global (EMTG) sont statistiquement significativement différents l'un de l'autre.
- Si c'est le cas, nous transmettons l'Effet Moyen du Traitement Global (EMTG) à l'étape suivante pour une investigation approfondie.

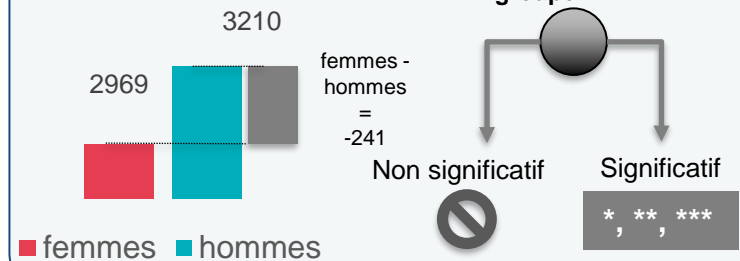
4. Effectuer des tests de signification pour la différence entre les sexes pour les sous-groupes

4.1 Utiliser l'impact moyen de groupe du programme

4.2 Prendre la différence de genre

4.3 Effectuer des tests de signification pour la différence entre les sexes au sein de chaque sous-groupe

Exemple illustratif:



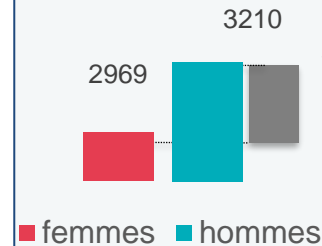
Étape 5: Équilibrage d'entropie

- Nous estimons et équilibrons les entropies sur les Effets Moyens du Traitement Individualisés (EMTI) de sorte que les caractéristiques des individus masculins puissent être similaires à celles des individus féminins dans les données.
- L'équilibrage est effectué sur le contrôle du genre.
- Cela nous permet de démontrer comment les effets différeraient si les hommes et les femmes avaient des caractéristiques similaires.

4. Effectuer des tests de signification pour la différence entre les sexes pour les sous-groupes

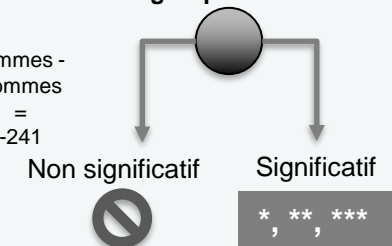
4.1 Utiliser l'impact moyen de groupe du programme

Exemple illustratif:



4.2 Prendre la différence de genre

4.3 Effectuer des tests de signification pour la différence entre les sexes au sein de chaque sous-groupe



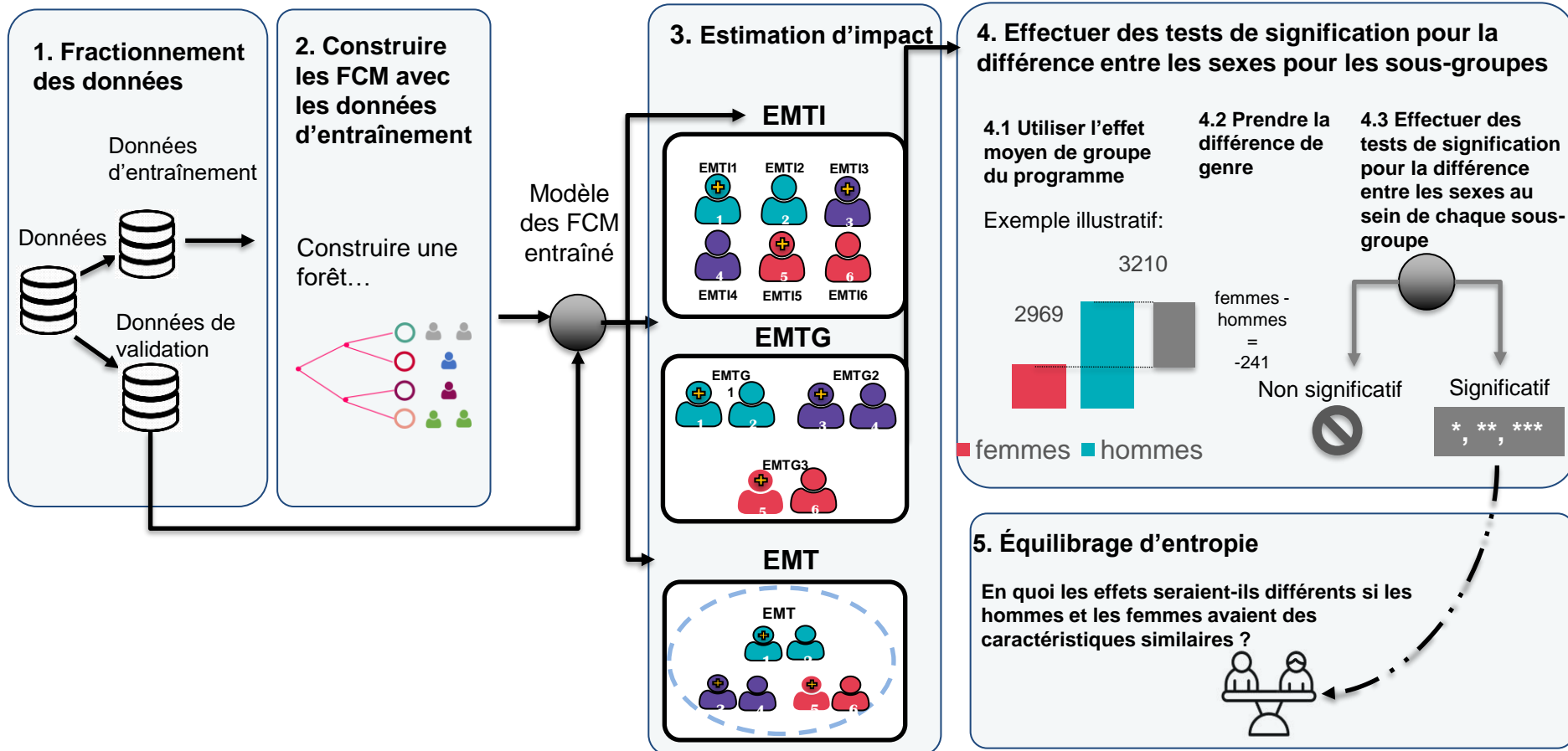
5. Équilibrage d'entropie

En quoi les effets seraient-ils différents si les hommes et les femmes avaient des caractéristiques similaires ?



Méthodologie: Récapitulatif

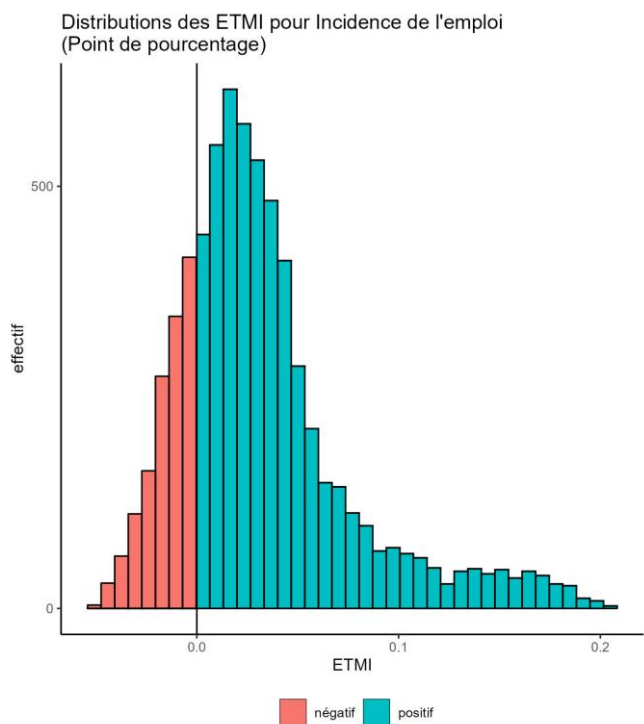
En utilisant les résultats de la méthode des FCM, la méthodologie comprend des tests de signification et un équilibrage d'entropie pour évaluer les différences de genre.



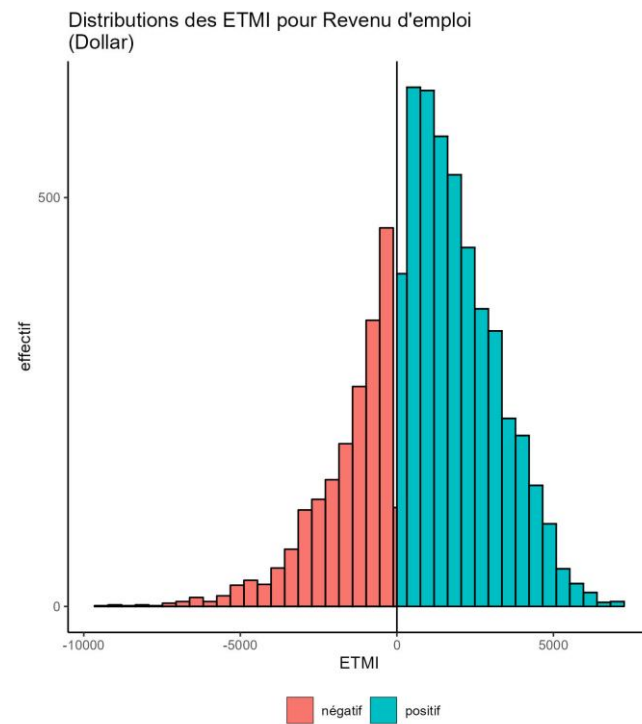
Exemples de résultats – Subventions salariales ciblées (SSC), EDMT

Distribution des ETMI

- Les impacts différentiels ont révélé qu'il y a une hétérogénéité limitée dans les impacts du programme.
- La majorité des prestataires actifs d'assurance-emploi qui ont participé au SSC en ont bénéficié.
- Les résultats indiquent que:
 - 79 % des participants ont connu une augmentation de l'incidence de l'emploi
 - 70 % des participants ont augmenté leurs revenus d'emploi



78.95 % des données sont supérieures à 0

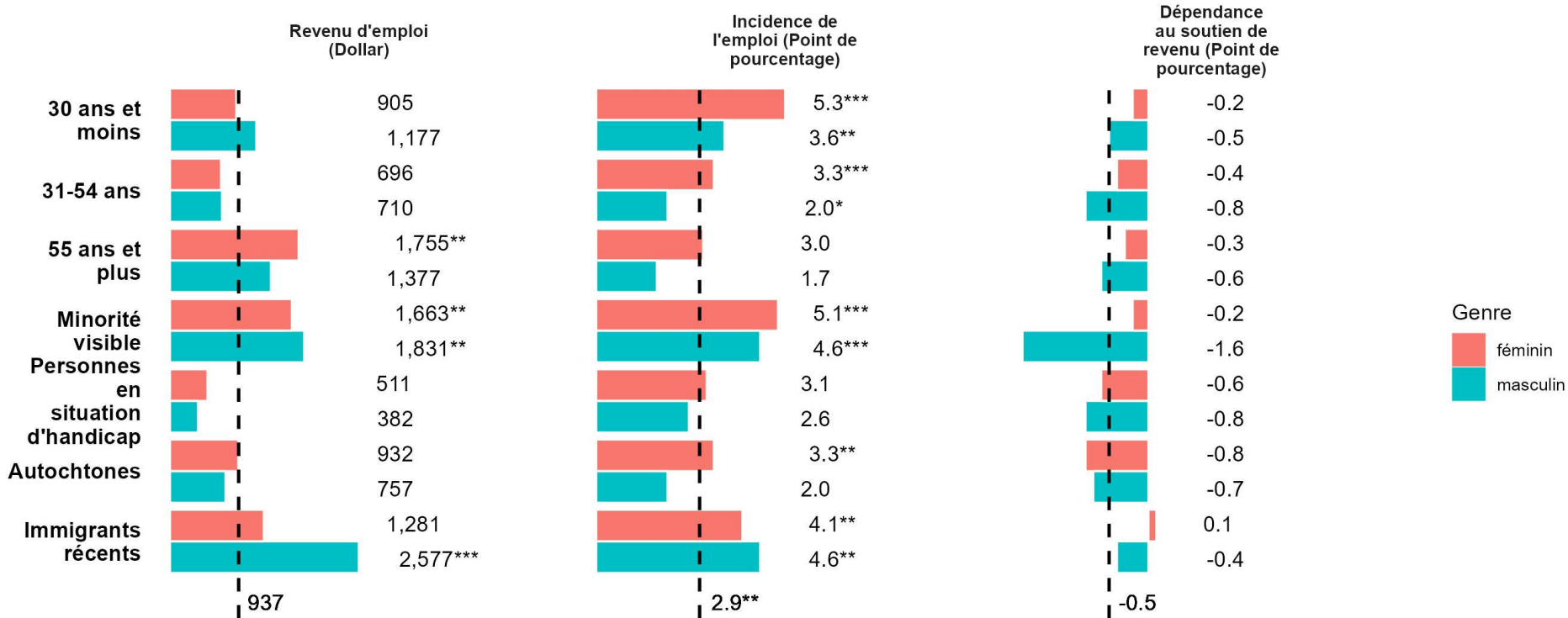


70.08 % des données sont supérieures à 0



Exemples de résultats – Subventions salariales ciblées (SSC), EDMT

Impacts différentiels par sexe et par autres sous-groupes, période de 5 ans après la participation, moyennes annuelles



- Dans l'ensemble, tous les sous-groupes ont augmenté leur incidence d'emploi. Deux groupes ont connu une augmentation plus importante de leur incidence d'emploi et de leurs revenus d'emploi:

- Participants féminins et masculins appartenant à des minorités visibles
- Immigrants récents de sexe masculin

Notation pour les seuils de signification

*** Seuil de 1%

** Seuil de 5%

* Seuil de 10%

L'effet moyen global du traitement sur les participants, annoté par la ligne pointillée.

Exemple d'Équilibrage d'Entropie - Subventions salariales ciblées, EDMT

Dans l'ensemble, nous n'avons observé aucune différence de genre dans les impacts du programme.

- Les résultats initiaux pour les subventions salariales suggéraient des différences de genre entre les hommes et les femmes récemment immigrés.
 - Les hommes ont augmenté leurs revenus d'emploi de 1 296 \$ de plus que les femmes (statistiquement significatif à 1 %), ce qui suggère une différence dans les impacts du programme.
- Cependant, après avoir contrôlé pour leurs caractéristiques sociodémographiques, les différences sont devenues non statistiquement significatives, suggérant qu'il n'y a pas de différence dans l'impact du programme.

	Revenus d'emploi (en dollars)
Sans contrôle des caractéristiques sociodémographiques	-1,296**
Lorsque les hommes ont des caractéristiques sociodémographiques similaires à celles des femmes	-328

Notation pour les seuils de signification: *** seuil de 1%, ** seuil de 5%, * seuil de 10%.



Conclusion de l'étude

La méthode d'apprentissage automatique a réussi à produire des résultats robustes pour les principales interventions du programme

- Dans l'ensemble, les résultats sont en cohérents avec les évaluations précédentes et offrent un nouveau niveau de détail pour examiner les impacts du programme à travers une lentille ACS Plus.
- Les résultats peuvent aider à comprendre la distribution des impacts sur différents groupes et à éclairer le développement des politiques tout en soutenant la conception des programmes du point de vue de "ce qui fonctionne le mieux pour qui".

Dans le cadre des futurs cycles d'évaluation :

- Les résultats de l'apprentissage automatique pourraient fournir une nouvelle ligne de preuves pour explorer les impacts différenciés sur les sous-groupes lorsque cela est possible.
- Des recherches qualitatives complémentaires et des analyses seraient nécessaires pour contextualiser ces résultats. Cela pourrait être réalisé dans le cadre des futurs cycles d'évaluation spécifiques aux programmes.



Limites

- Cette étude est limitée aux informations disponibles dans les données administratives :
 - Le sexe biologique a été utilisé comme un substitut du genre et des données n'étaient pas disponibles pour certains facteurs d'identité liés à l'ACS Plus.
- Des différences préexistantes pourraient exister entre les participants et les non-participants qui n'ont pas été mesurées lors du processus d'appariement :
 - Par exemple : capacité, santé et motivation à rechercher un emploi.
- Les résultats ne sont pas directement comparables entre les programmes :
 - Cette analyse a utilisé des groupes de comparaison construits par l'intervention du programme.
- L'étude ne capture pas la participation à plusieurs interventions :
 - En utilisant des EPA, l'analyse a attribué la plus longue intervention comme l'intervention principale dans l'unité d'analyse.



Perspectives

- Exploration de l'utilisation de la méthode des FCM dans le cadre d'évaluations à venir pour les programmes du marché du travail.
 - Lorsque seuls des ensembles de données plus petits sont disponibles, la méthode d'appariement traditionnelle restera la méthode privilégiée pour mener des analyses d'impact net.
- Poursuite de la collaboration avec le Professeur Lechner sur les moyens de mesurer l'effet du genre et d'autres facteurs d'identité interconnectés.
- Partage de notre expérience avec d'autres concernant des études exploratoires en apprentissage automatique.



Annexe A: Cadre des résultats potentiels

Pour un ensemble d'individus i.i.d (indépendants et identiquement distribués) $i = 1, \dots, n$, nous considérons le triplet (X_i, Y_i, D_i) , composé

- d'une covariable X_i
- d'un résultat Y_i
- d'une attribution du traitement D_i

Soit D le traitement qui peut prendre un nombre connu de M valeurs entières différentes de 0 à $M - 1$. Le résultat d'intérêt (potentiel) qui se réalise sous le traitement d est désigné par Y^d .

L'objectif est de trouver $EMTI(m, l; x, \Delta) = E(Y^m - Y^l | Z = z, D \in \Delta)$

$EMTI(m, l; x, \Delta)$ mesure l'impact du traitement m par rapport *au traitement* l pour des unités ayant des caractéristiques x qui appartiennent au groupe de traitement Δ , où Δ représente l'ensemble de tous les traitements d'intérêt.



Annexe B: Hypothèses d'identification

$$\{Y^0, \dots, Y^m, \dots, Y^{M-1}\} \perp\!\!\!\perp D \mid X = x, \quad \forall x \in \chi; \quad (CIA)$$

$$0 < P(D = d \mid X = x) = p_d(x), \quad \forall x \in \chi, \forall d \in \{0, \dots, M-1\}; \quad (CS)$$

$$Y = \sum_{d=0}^{M-1} \mathbb{1}(D = d) Y^d; \quad (SUTVA)$$

$$X^d = X, \quad \forall d \in \{0, \dots, M-1\}. \quad (EXOG)$$

- **Hypothèse d'indépendance conditionnelle (CIA):** Aucune caractéristique autre que X n'influence conjointement le traitement et les résultats potentiels dans la plage d'intérêt (χ).
- **Hypothèse de support commun (CS):** Chaque valeur dans χ permet l'observation de tous les traitements.
- **Hypothèse de la stabilité de la valeur du traitement unitaire (SUTVA):** La valeur observée du traitement est indépendante de l'allocation du traitement des autres individus, excluant les effets de débordement et de taille du traitement.
- **Hypothèses d'exogénéité (EXOG):** Les valeurs observées de X ne dépendent pas du statut du traitement, excluant ainsi tout effet causal de D sur X .



Annexe C: Trouver un estimateur pour l'Effet Moyen du Traitement Individualisé (EMTI)

Lorsque toutes les hypothèses d'identification sont respectées, l'EMTI peut également être exprimé comme suit :

$$EMTI(m, l; x) = \mu_{m(x)} - \mu_{l(x)}; \forall x \in \mathcal{X}, \forall m \neq l \in \{0, \dots, M - 1\}$$

En désignant les attentes conditionnelles de Y données X dans la sous-population $D = d$ par $\mu_d(x)$

- Un estimateur facile à mettre en œuvre implique d'estimer séparément deux attentes conditionnelles à l'aide d'outils d'apprentissage automatique standard, puis de prendre la différence.
 - L'inconvénient de cette approche est que les méthodes standard d'apprentissage automatique privilégient la maximisation de la puissance prédictive hors échantillon pour chaque estimateur séparément.
 - L'utilisation de méthodes telles que les Forêts Aléatoires peut entraîner une variabilité dans les effets du traitement estimés, surtout lorsque les caractéristiques sont fortement prédictives de Y , mais les effets du traitement sont relativement constants.
 - Des parts de traitement inégales* peuvent également poser un problème, la forêt pour un traitement étant plus fine que pour l'autre en raison de différences de tailles d'échantillons.
- Une approche alternative consiste à utiliser les mêmes règles de division pour les deux sous-échantillons, à estimer $\mu_m(x)$ et $\mu_l(x)$ séparément, puis à trouver une règle de division plausible pour une forêt 'jointe'.



*: la proportion d'individus recevant un traitement

Références

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523), 1228-1242.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260.

Lechner, M. (2019). Modified causal forests for estimating heterogeneous causal effects.

Cookiecutter Data Science

A logical, reasonably standardized, but flexible project structure for doing and sharing data science work.

<https://drivendata.github.io/cookiecutter-data-science/>

Athey S. Solving Heterogeneous Estimating Equations Using Forest Based Algorithms

<https://www.youtube.com/watch?v=CPz0HdUM3dE>

