



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*



JMS 2022

L'expérimentation TZCLD et l'enquête QVT

L'introduction d'une variable endogène inobservable modifie-t-elle les résultats ?

Félix Bonnetête (Dares)

Introduction

L'EXPÉRIMENTATION « TERRITOIRES ZÉRO
CHÔMEURS DE LONGUE DURÉE »

Les principes de l'expérimentation TZCLD

- Entrée en vigueur en janvier 2017, l'objectif de l'expérimentation « Territoires Zéro Chômeurs de Longue Durée » est d'**éradiquer le chômage de longue durée**.
- Embauche de personnes privées durablement d'emploi (**PPDE**) au sein d'entreprises qui leur sont dédiées : les entreprises à but d'emploi (les **EBE**).
- Les individus recrutés sont employés **en CDI au SMIC** à temps choisi.
- En théorie, le coût pour les finances publiques est **neutre** (selon les porteurs du projet).
- Les activités des EBE ne doivent **pas concurrencer** les activités existantes.

Le rôle du comité scientifique et de la Dares

- Un comité scientifique réalise une **évaluation de l'expérimentation** sur les territoires pilotes.
- La Dares, qui assure le secrétariat du comité, crée et met en place l'enquête « **Qualité de vie dans les territoires** ». Ses objectifs sont les suivants :
 - analyser les **trajectoires professionnelles et sociales des bénéficiaires** (individus recrutés en EBE) et mesurer l'impact du dispositif sur celles-ci (emploi, bien-être, santé, lien social) en les comparant à une population témoin ;
 - évaluer **l'impact de l'expérimentation sur les territoires de l'expérimentation**, en particulier une éventuelle modification du cadre et de la qualité de vie.

Les données utilisées

L'ENQUÊTE « QUALITÉ DE VIE DANS LES
TERRITOIRES »

L'enquête « Qualité de Vie dans les Territoires »

- 10 territoires **pilotes**. Des territoires **témoins** ont été choisis via une Analyse en Composantes Principales.
- **3 populations** interrogées:
 - Les bénéficiaires : salariés conventionnés d'une EBE.
 - Les éligibles : individus non salariés d'une EBE mais éligibles à une embauche. Dans les faits, il s'agit d'allocataires du RSA ou de personnes inscrites à Pôle Emploi depuis au moins un an.
 - Les résidents : personnes non éligibles à une embauche en EBE mais qui peuvent profiter des externalités positives de l'expérimentation.
- Une enquête **en 2 vagues** : une au T3 2018 et une au T4 2019 – T1 2020 (avant le 1er confinement) – Panel non cylindré

Méthodologie

3.1) IMPLÉMENTATION D'UN MODÈLE DE MATCHING

Matching sur score de propension (1/2)

- Réaliser un **matching sur score de propension** entre les bénéficiaires et les éligibles témoins pour constituer des **faux-jumeaux** avec lesquels une comparaison sera pertinente.
- Un matching sur des **périodes différentes** :

| Date d'observation des variables ayant servi à la constitution des faux-jumeaux | Date d'observation des variables ayant servi à l'analyse des trajectoires |
|---|---|
| Vague 1 | Vague 2 |

- But : Analyser les effets de l'expérimentation sur les bénéficiaires de vague 1 plus d'un an après leur première interrogation

Matching sur score de propension (2/2)

- Appariement réalisé via la méthode de **kernel** sur un grand nombre de **variables** comme l'âge, le sexe, le niveau de diplôme, le type de logement de résidence, le fait d'être en situation de handicap ou non...
- **Qualité de l'estimation satisfaisante :**
 - Hypothèse de support commun : validée
 - Équilibrage bloc par bloc : avant constitution des faux-jumeaux, un seul croisement est non équilibré.
 - Équilibrage général : après constitution des faux-jumeaux, il n'existe aucune différence significative à 5 % entre bénéficiaires et faux-jumeaux.

Méthodologie

3.2) SIMULATION D'UNE VARIABLE
INOBSERVABLE

Pourquoi simuler une variable inobservable ?

- Objectif : vérifier la crédibilité de l'**hypothèse d'indépendance conditionnelle** : « une fois que toutes les variables de contrôle observables possible ont été utilisées lors du calcul du score de propension, la sélection dans le traitement ne dépend pas des résultats attendus de ce dernier pour un individu donné ».
- Méthode : faire varier l'**intensité de la corrélation de l'inobservable** avec le traitement et la variable d'intérêt et déterminer dans quelle mesure **les résultats varient** aussi.
- La **motivation** est ici cette variable inobservable. On la construit artificiellement et on l'ajoute dans le calcul du score de propension.

Algorithme de simulation d'une inobservable : un exemple avec le fait d'être en emploi

- Sélection d'un couple de facteurs multiplicatifs ($x; y$)
- Pour un couple ($x; y$) donné, création d'une variable **seuil** égale à 0,8.
- Si l'individu n'est pas bénéficiaire alors $\text{seuil} = \frac{\text{seuil}}{x}$; si l'individu n'est pas en emploi alors $\text{seuil} = \frac{\text{seuil}}{y}$.
- Génération d'une variable u qui suit une loi uniforme entre 0 et 1.
- Création de la variable inobservable **inobs** telle que si $\text{seuil} > u$ alors **inobs**=1, sinon **inobs**=0.
- **inobs** est utilisée dans le calcul du score de propension. Le matching est réalisé.
- Pour un couple donné, obtention d'un ensemble de coefficients pour chacune des simulations de u . Ces coefficients correspondent à la différence moyenne entre bénéficiaires et faux-jumeaux pour la variable « être en emploi ».
- Pour chaque couple ($x; y$), il est donc possible d'aboutir à un coefficient moyen et un intervalle de confiance à 90 %.

Simulation d'une inobservable : un exemple fictif

- $(x;y) = (4;2)$.
- **seuil** = 0,8.

| Individu | Valeur de <i>seuil</i> après calcul | Valeur de <i>inobs</i> avec $u=0,3$ | Valeur de <i>inobs</i> avec $u=0,6$ |
|----------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| Bénéficiaire en emploi | 0,8 | 1 | 1 |
| Bénéficiaire pas en emploi | 0,4 | 1 | 0 |
| Eligible en emploi | 0,2 | 0 | 0 |
| Eligible pas en emploi | 0,1 | 0 | 0 |

Résultats

Conclusions du matching sans prise en compte de l'inobservable

- L'étude des différences pondérées de moyennes entre bénéficiaires et faux-jumeaux conclue à **des effets positifs et significatifs à 5 %** sur l'emploi, la santé ou le bien-être.
- Parmi eux, deux font ici l'objet d'une attention particulière :
 - Les bénéficiaires ont une probabilité **d'être en emploi supérieure de 36,8 points** à celle de leurs faux-jumeaux.
 - Les bénéficiaires ont une probabilité **de ne pas avoir de difficultés à se concentrer supérieure de 17,2 points** à celle de leurs faux-jumeaux.
- L'impact de l'introduction d'une inobservable est ici testé sur ces deux dimensions.

Résultats sur l'emploi

| Couples (x;y) | (1;1) | (2;2) | (3;3) | (4;4) | (5;5) | (6;6) |
|--------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Coefficient | 0,37 | 0,28 | 0,23 | 0,19 | 0,17 | 0,16 |
| IC90 | [0,36 ; 0,37] | [0,25 ; 0,31] | [0,18 ; 0,27] | [0,15 ; 0,24] | [0,12 ; 0,22] | [0,12 ; 0,21] |

- Plus les facteurs x et y augmentent, plus la valeur du coefficient diminue.
- Ce coefficient semble cependant atteindre un palier.
- Malgré une importante corrélation, les coefficients restent significatifs à 10 %.

Résultats sur les difficultés à se concentrer

| Couples (x;y) | (1;1) | (2;2) | (3;3) | (4;4) | (5;5) | (6;6) |
|--------------------|---------------|---------------|---------------|---------------|----------------|----------------|
| Coefficient | 0,17 | 0,10 | 0,06 | 0,04 | 0,02 | 0,01 |
| IC90 | [0,17 ; 0,18] | [0,07 ; 0,12] | [0,03 ; 0,09] | [0,00 ; 0,07] | [-0,01 ; 0,05] | [-0,02 ; 0,04] |

- Constat moins net : le coefficient devient non significatif à 10 % à partir de certains niveaux de corrélation.
- Pour des niveaux de corrélation inférieurs mais tout de même élevés, certains coefficients restent significatifs à 10 %.

Conclusion, limites et prolongements

Conclusions, limites et prolongements

- Introduire une variable endogène inobservable simulée n'a que peu ou pas d'impact sur les résultats sur **l'emploi** ; c'est moins le cas pour ce qui est de **la difficulté à se concentrer**.
- Des prolongements à cette méthode peuvent être envisagés :
 - Augmenter le **nombre de simulations** pour chaque couple $(x;y)$ ou augmenter le **nombre de couples** considérés.
 - Raffiner le fonctionnement même de l'algorithme de création de la variable endogène en se plaçant dans le cas **continu** plutôt que **binaire**.
 - Plutôt que de se focaliser sur une seule variable de bien-être, construire **un indicateur de santé mentale** à partir de l'ensemble des variables du questionnaire portant sur ce sujet.



RÉPUBLIQUE
FRANÇAISE

*Liberté
Égalité
Fraternité*



Merci de votre attention !