

# Université d'été en Sciences Sociales Tany Vao 2022 - Madagascar Essais randomisés contrôlés (ERC/RCT)

Florent Bédecarrats, Jeanne de Montalembert,  
Marc Bouvier, Marin Ferry, Kenneth Hougbedji

Université de Toliara, Octobre 2022



# Présentation de la méthode des ERC

## Les ERC: une méthode d'évaluation expérimentale

Le problème principal soulevé par le modèle de Rubin lors de l'évaluation des effets d'une intervention est l'impossibilité d'observer  $E(Y_{0i} | T_i = 1)$

- Problème du **contrefactuel**: Quel aurait été le **résultat** observé sur les unités traitées **en l'absence du traitement**?
  - Enjeu: Trouver un contrefactuel valide / pertinent
  - Menaces: Eligibilité au programme / auto-sélection  $\implies$  biais entre traités et non-traités
  - Une solution: Les essais randomisés contrôlés

# Présentation de la méthode des ERC

## Pourquoi la randomisation?

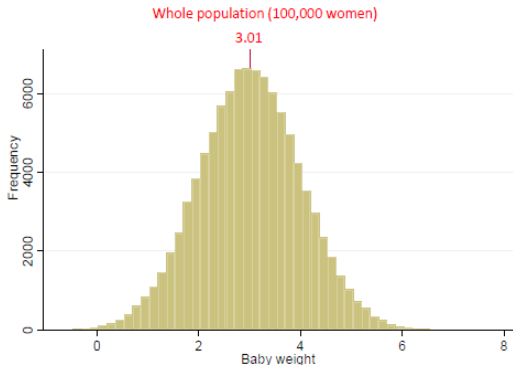
Randomisation: Le fait de participer au projet ou non est défini de **façon aléatoire** (exemple = à l'aide d'une loterie)

- Deux groupes sont aléatoirement choisis au sein d'une population "homogène":
  - Un groupe bénéficie de l'intervention (vaccin, prêt, formation, status d'aire protégée) = **groupe de traitement**
  - L'autre groupe ne bénéficie pas de l'intervention = **groupe de contrôle**
- L'assignation aléatoire permet d'obtenir deux groupes probablement très similaires (dispersion des observables et inobservables statistiquement identiques)... **si les groupes sont suffisamment importants!**

# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- Sélection aléatoire d'un nombre important d'individus (unités) d'une population → distrib. sous-groupe = distrib. de la pop. tot.

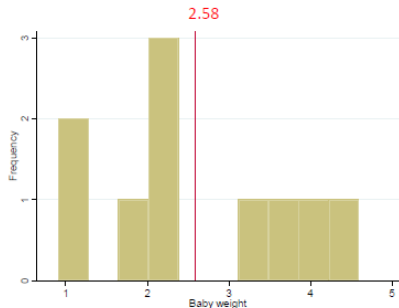
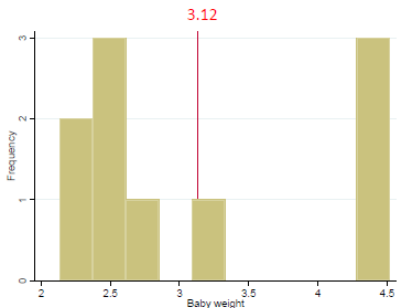


# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- Sélection aléatoire d'un nombre important d'individus (unités) d'une population → distrib. sous-groupe = distrib. de la pop. tot.

2 groups of 10 random observations

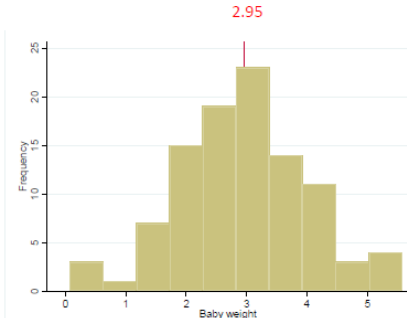
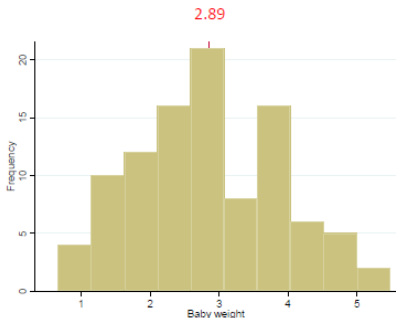


# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- Sélection aléatoire d'un nombre important d'individus (unités) d'une population → distrib. sous-groupe = distrib. de la pop. tot.

2 groups of 100 random observations

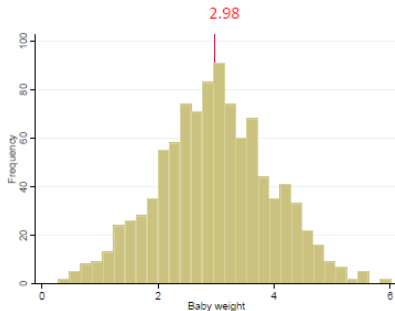
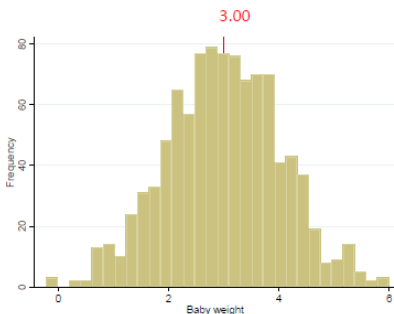


# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- Sélection aléatoire d'un nombre important d'individus (unités) d'une population → distrib. sous-groupe = distrib. de la pop. tot.

2 groups of 1,000 random observations



# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- En s'appuyant sur le tirage aléatoire de deux sous-échantillons de la population cible, nous devrions observer une distribution similaire des caractéristiques observables et inobservables:

$$E(X|T = 1) = E(X|T = 0) = E(X)$$

- De même en l'absence du programme, la distribution des variables de résultats devrait être identique (absence de biais de sélection)

- Ainsi le groupe de contrôle devient un **contrefactuel pertinent**. L'hypothèse suivante devient beaucoup plus réaliste:

$$E(Y_0|T = 0) = E(Y_0|T = 1) = E(Y_0)$$



# Présentation de la méthode des ERC

## Pourquoi la randomisation?

- **Attention!** L'hypothèse faite sur la validité du contrefactuel n'est valable qu'en moyenne (impossible de dire que  $X_i(1) = X_i(0)$  et que  $Y_i(1) = Y_i(0)$ )
- Nous pouvons uniquement essayer de mesurer l'**effet moyen** du programme qui est obtenu par:

$$E(Y_i | T_i = 1) - E(Y_i | T_i = 0)$$

# Présentation de la méthode des ERC

## Estimation de l'effet moyen du traitement

- Différence de moyennes:  $E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) \longrightarrow$

$$\bar{Y}_1 - \bar{Y}_0$$

▶ exemple

- Régression linéaire:  $Y_i = \alpha + \beta T_i + \epsilon_i$  avec

$\beta$  l'effet moyen du traitement

$\alpha$  la moyenne de  $Y_i$  pour les non-traités

$\epsilon$  le terme d'erreur

▶ exemple

# Processus du tirage aléatoire

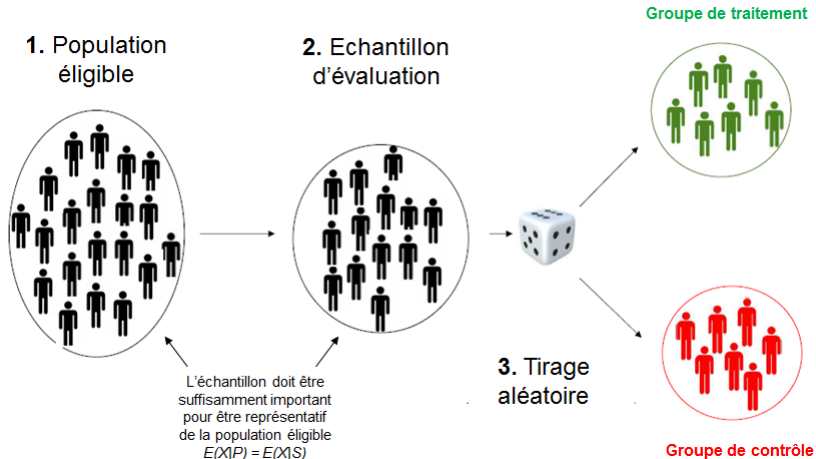
## Comment randomiser?

Quel est le processus à suivre pour “randomiser” un projet/programme/intervention? Plusieurs étapes:

1. Identification des **unités éligibles** au programme (individus, ménages, écoles, villages, aires géographiques): les unités pour lesquelles nous voulons connaître l'impact du programme
2. Sélection de l'**échantillon soumis à l'évaluation** (extraction de la population d'intérêt)
3. **Assignment aléatoire** du traitement (quel groupe bénéficie de l'intervention?)

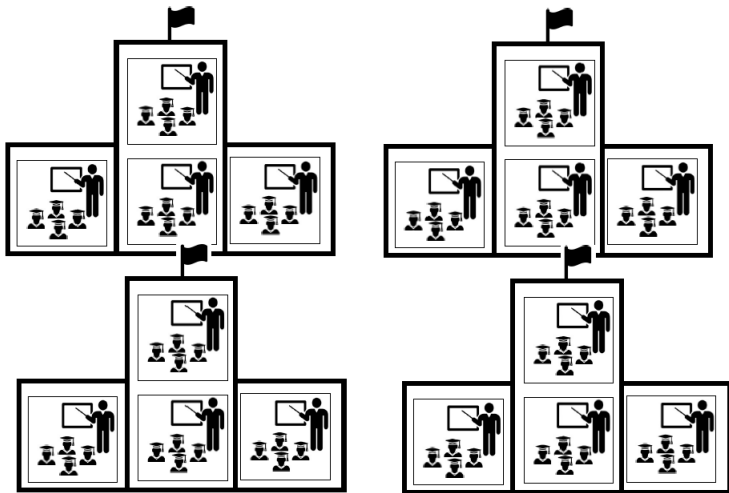
# Processus du tirage aléatoire

Comment randomiser?



# Processus du tirage aléatoire

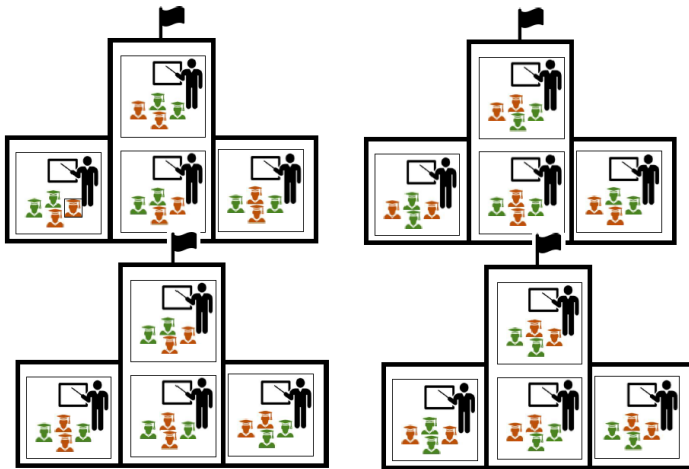
A quel niveau faut-il effectuer le tirage?



Source: Marine de Talancé (DEIPM UGE)

# Processus du tirage aléatoire

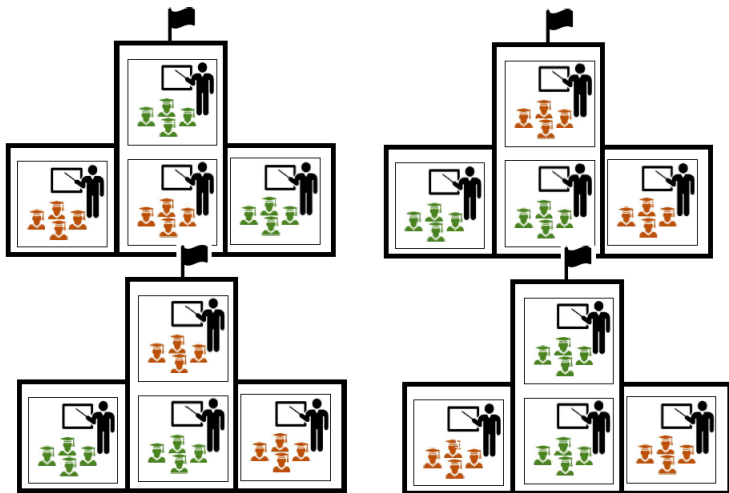
A quel niveau faut-il effectuer le tirage?



Source: Marine de Talancé (DEIPM UGE)

# Processus du tirage aléatoire

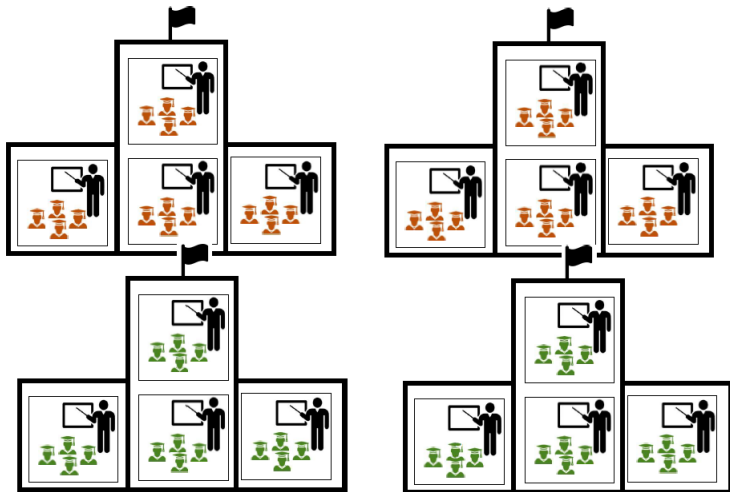
A quel niveau faut-il effectuer le tirage?



Source: Marine de Talancé (DEIPM UGE)

# Processus du tirage aléatoire

A quel niveau faut-il effectuer le tirage?



Source: Marine de Talancé (DEIPM UGE)



# Processus du tirage aléatoire

A quel niveau faut-il effectuer le tirage?

Le niveau du tirage dépend de celui de l'intervention: le niveau de randomisation ne peut être inférieur au niveau auquel est assigné le programme

La randomisation n'est pas la panacée: sa pertinence dépend du type d'intervention → **arbitrages**:

- Un niveau trop agrégé de l'assignation (régions, pays, provinces) diminue la puissance statistique et augmente le risque de facteurs confondants
- Assigner une intervention à un niveau très désagrégé (individu, ménage) augmente les risques de contamination du groupe de contrôle

# Généralisation des résultats d'ERC

## Validité interne

**Validité interne:** La différence observée sur les variables de résultats entre traités et non-traités est attribuable à l'intervention et à l'intervention seulement

- Validité interne  $\Leftarrow$  Assignation aléatoire du traitement
  - Le groupe de contrôle est similaire au groupe de traitement (en moyenne) avant l'intervention
  - Une fois l'intervention commencée, les deux groupes sont exposés au même environnement et aux mêmes évolutions temporelles
  - Groupe de contrôle: bonne représentation de ce qu'il serait arrivé aux "traités" en l'absence de traitement
  - La différence de résultats entre traités et non-traités après l'intervention est donc due à cette dernière

# Généralisation des résultats d'ERC

## Validité externe

**Validité externe:** Les résultats de l'évaluation sont généralisables à l'ensemble des unités éligibles (dans le temps, dans d'autres contextes)

- La validité externe des ERC est faible. En effet ils conduisent à l'obtention de résultats...
  - Qui sont spécifiques à un pays, une population spécifique (celle de l'étude), à un moment donné
  - Qui ne peuvent être extrapolés à d'autres endroits du monde et sur d'autres populations
  - Qui nous permettent néanmoins de comprendre les effets d'une intervention dans un contexte spécifique

# Problèmes soulevés et limites des ERC

## 1. Enjeux éthiques

- La sélection aléatoire conduit à refuser le traitement à un ensemble de personnes à l'instant T
- Questionnable lorsque les interventions évaluées par ERC portent sur les besoins humains
- Les ERC comme étalon-or de l'évaluation des politiques de développement?  $\implies$  Décrédibilisation des autres approches évaluatives
- Tout ne peut pas être testé avec les ERC

# Problèmes soulevés et limites des ERC

1. Enjeux éthiques
2. Enjeux statistiques: être sûr de pouvoir détecter un effet
  - Les enquêtes sur un grand nombre d'individus, et la mise en place de l'attribution aléatoire de l'intervention sont coûteuses
  - Arbitrage entre précision de l'évaluation et moyens financiers

## Problèmes soulevés et limites des ERC

1. Enjeux éthiques
2. Enjeux statistiques: être sûr de pouvoir détecter un effet
3. Contamination et externalités: menaces sur la validité interne
  - Certains non-traités peuvent en réalité bénéficier de manière indirecte de l'intervention (ex: Kremer & Miguel, 2004)
  - L'intervention peut avoir des conséquences sur des facteurs environnants qui vont impacter les variables de résultats (changement de prix)
  - Ajustement des comportements (Hawtorne effect, John Henry effect, Découragement)

# Problèmes soulevés et limites des ERC

1. Enjeux éthiques
2. Enjeux statistiques: être sûr de pouvoir détecter un effet
3. Contamination et externalités: menaces sur la validité interne
4. Validité externe
  - Replications des ERC dans d'autres contextes et identification des raisons pour lesquelles les effets changent
  - Les résultats sur un sous-échantillon de la population d'un pays sont-ils généralisables à l'ensemble de la population?

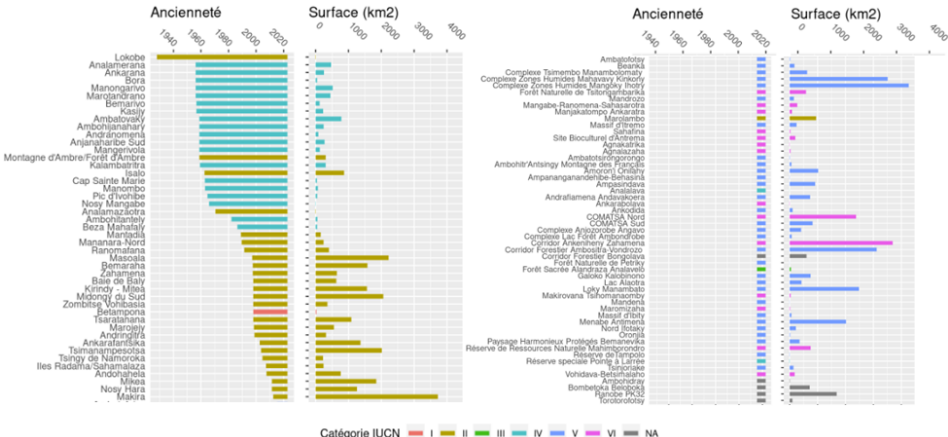
Une application à notre thématique

**À vous de jouer !**



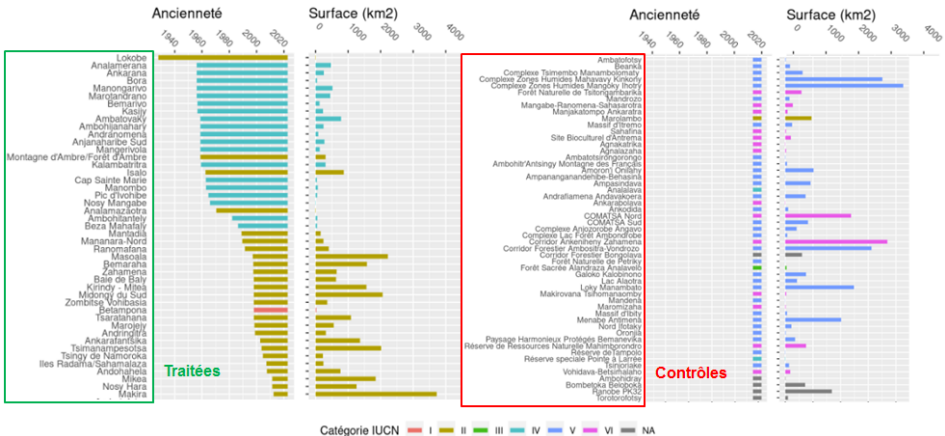
# Aires protégées et déforestation

Estimation de l'impact des AP sur la déforestation via les ERC



# Aires protégées et déforestation

Estimation de l'impact des AP sur la déforestation via les ERC



Contrôles

Catégorie IUCN I II III IV V VI NA

# Aires protégées et déforestation

Estimation de l'impact des AP sur la déforestation via les ERC

## Moyennes après intervention

(exercice : "comme si" c'était une RCT)

Groupe	Controle	Traitement	t-test
Déforestation 2006-2016 (%)	8.50	3.91	0.01

Source : Association Vahatra et Carvalho et al. 2018

# Aires protégées et déforestation

Estimation de l'impact des AP sur la déforestation via les ERC

## Moyennes après intervention

(exercice : "comme si" c'était une RCT)

p-value

Groupe	Contrôle	Traitement	t-test
Déforestation 2006-2016 (%)	8.50	3.91	0.01

Source : Association Vahatra et Carvalho et al. 2018

# Aires protégées et déforestation

Estimation de l'impact des AP sur la déforestation via les ERC

## Moyennes après intervention

(exercice : "comme si" c'était une RCT)

Groupe	Controle	Traitement	t-test
Déforestation 2006-2016 (%)	8.50	3.91	0.01

Source : Association Vahatra et Carvalho et al. 2018

p-value

p-value < 0.1 →  
différence significative

▶ back

## Déforestation 2006-2016 (%)

(exercice : "comme si" c'était une RCT)

<b>term</b>	<b>estimate</b>	<b>std.error</b>	<b>statistic</b>	<b>p.value</b>
(Intercept)	8.499417	1.207649	7.037989	0.0000000003414988
GroupeTraitement	-4.592813	1.745413	-2.631362	0.0099718430305533

Source : Association Vahatra et Carvalho et al. 2018

Moyenne groupe de  
contrôle

## Déforestation 2006-2016 (%)

(exercice : "comme si" c'était une RCT)

term	estimate	std.error	statistic	p.value
(Intercept)	8.499417	1.207649	7.037989	0.0000000003414988
GroupeTraitement	-4.592813	1.745413	-2.631362	0.0099718430305533

Source : Association Vahatra et Carvalho et al. 2018

**Effet du traitement**

## Déforestation 2006-2016 (%)

(exercice : "comme si" c'était une RCT)

<b>term</b>	<b>estimate</b>	<b>std.error</b>	<b>statistic</b>	<b>p.value</b>
(Intercept)	8.499417	1.207649	7.037989	0.0000000003414988
GroupeTraitement	-4.592813	1.745413	-2.631362	0.0099718430305533

Source : Association Vahatra et Carvalho et al. 2018



Effet du traitement

## Déforestation 2006-2016 (%)

(exercice : "comme si" c'était une RCT)

term	estimate	std.error	statistic	p.value
(Intercept)	8.499417	1.207649	7.037989	0.0000000003414988
GroupeTraitement	-4.592813	1.745413	-2.631362	0.0099718430305533

$$\bar{Y}_1 - \bar{Y}_0 = 3.91 - 8.50 = 4.59$$

**Effet du traitement**

## Déforestation 2006-2016 (%)

(exercice : "comme si" c'était une RCT)

term	estimate	std.error	statistic	p.value
(Intercept)	8.499417	1.207649	7.037989	0.0000000003414988
GroupeTraitement	-4.592813	1.745413	-2.631362	0.0099718430305533

$$\bar{Y}_1 - \bar{Y}_0 = 3.91 - 8.50 = 4.59$$

P-value &lt; 0.1 → Effet significatif

▶ back